On discrete Morse theory in persistent topology

Dissertation
zur Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Julian Jürgen Paul Brüggemann
aus
Herne

Bonn 2024

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachterin: Dr. Viktoriya Ozornova
2. Gutachter: Prof. Dr. Carl-Friedrich Bödigheimer

Tag der Promotion: 20.06.2024
Erscheinungsjahr: 2024

Abstract

The main topic of this thesis is the interplay between discrete Morse theory and persistent topology: we apply methods from discrete Morse theory to filtered complexes, as they appear in Topological Data Analysis, and study certain persistent homotopy invariants of these complexes, namely merge trees and barcodes.

This thesis is subdivided into three related research projects. In the first two projects we investigate the inverse problem between discrete Morse functions on graphs and their induced merge trees. We generalize a construction by Johnson–Scoville that associates a chiral merge tree to any discrete Morse–Benedetti function on a tree. Using this generalized construction, we present a complete combinatorial description of the fiber of the "induced merge tree" map. Moreover, we find an inverse construction: we associates a discrete Morse–Benedetti function on a tree, respectively graph, to any merge tree, respectively generalized merge tree. In addition, we give a finite number of edit moves, which we call component-merge equivalences, that relate all elements of the fiber to each other. It turns out that for generalized merge trees, which also contain cycle birth information, our inverse construction cannot always produce a critical discrete Morse function on a simple graph. We present a complete criterion for when a generalized merge tree can be realized by a discrete Morse–Benedetti function on a simple graph, and when such a graph can be chosen to be planar. Furthermore, we describe an algorithm that uses the induced (generalized) merge tree in order to optimize discrete Morse functions on graphs.

In the third project, we develop models for parameter spaces of discrete Morse functions on CW complexes and parameter spaces of merge trees. We relate them to parameter spaces of smooth Morse functions on manifolds, parameter spaces of discrete Morse matchings on regular complexes and parameter spaces of barcodes. The project is motivated on one hand by Cerf's investigation of the space of smooth functions on a manifold that he used to investigate under which circumstances pseudo-isotopies are actually isotopies. On the other hand the third project is also motivated by the goal of the previous two projects: investigations of inverse problems in Topological Data Analysis. parameter spaces provide a more convenient framework for the analysis of inverse problems in persistent topology because they carry more structure to capture the information of the persistent invariants at hand. Except for the parameter spaces of smooth Morse functions, which belong to the realm of (Fréchet) manifolds, all the other mentioned parameter spaces are of a combinatorial nature, namely spaces that are in some way associated to hyperplane arrangements in vector spaces of discrete functions. After introducing these parameter spaces, we realize the following constructions between their corresponding objects of study as continuous maps between the parameter spaces: defining a discrete Morse function on the CW decomposition induced by a smooth Morse function on a manifold, the merge tree induced by a discrete Morse function, the Morse matching induced by a discrete Morse function, and the barcode induced by a merge tree. We also study some basic properties of these maps between parameter spaces.

Acknowledgements

It seems impossible to name all of the wonderful people who had a positive impact on me and my dissertation during my PhD studies. Thus, this section is only to be seen as an attempt to name as many as of them as possible.

First and foremost, I want to thank my primary advisor, Viktoriya Ozornova, who has been an incredible mentor. In particular, I want thank her for the many friendly conversations, in which she provided valuable feedback to my work, gave helpful career advice, and made great suggestions for good academic habits and practices. Moreover, I want to thank her for the extended freedom I had with respect to my research subject, which many scientists only experience at higher career stages.

I want to thank my secondary advisor, Carl-Friedrich Bödigheimer for the many helpful discussions, especially during his Forschungsseminar. I also want to thank him for his career advice and for challenging my views on several aspects of academia and life.

I want to thank both of my advisors for their great support in the different stages and situations of my PhD studies.

Next, I want to thank all the wonderful people , who accompanied me during my time in Bonn. Special thanks go to Florian Kranhold, Andrea Bianchi, Simona Veselá, Daniel Bermudez, Luuk Stehouwer, and all the other wonderful mathematicians at both, MPIM and University of Bonn, who share a lot of wonderful, but also some difficult, memories with me. I thank Max Planck Institute for Mathematics and University of Bonn for all the ressources and the great scientific environment, in which I conducted my research.

Moreover, I want to thank University of Bonn and MPIM for the great scientific environment in which I conducted most of my PhD research. Additionally, I thank Ruhr University Bochum, especially the topology group, for the great scientific environment, in which I spent the first year of my PhD studies. In particular, I thank Gerd Laures and Björn Schuster for the many helpful discussions and for introducing me to the field of topology during my bachelor and master studies.

I want to thank my coauthor, Nick Scoville, for the great collaboration that led to our joint article. Moreover, I want to thank him for inspiring me with his work to take the first step towards working in Topological Data Analysis.

I thank my family and friends for their support from the outside, especially during the most stressful times of my PhD. In particular, I want to thank my parents, who supported me substantially during the whole time.

In addition, I want to thank all my friends from dancing and scuba diving, who gave me a place to escape and helped me to maintain sanity.

Last, but not least, I want to give special thanks to all the great people who read drafts of parts of this thesis and provided valuable feedback.

Contents

CHAPTER I

Introduction

Persistent topology is a relatively new subfield of topology and geometry. It consists of the study of filtered, so called persistent, spaces and their topological properties and invariants. The term “persistent” refers to the study of filtrations of topological spaces and stems from the concept of persistent homology as it is used in Topological Data Analysis (TDA): one of the objects of study is for what range of parameters in a filtration a topological invariant “persists”, i.e. is non-trivial. The notions of “persistent homology” and “persistent Betti numbers” have been first introduced in the early 2000s, see Edelsbrunner, Letscher, and Zomorodian 2002 and Robins 2002. In modern language, one can consider a filtered space as a functor $X_\bullet\colon (\mathbb{R}, \leq) \to Top$. Then the persistent homology of $X_\bullet$ is the concatenation $H_n \circ X_\bullet$, where H_n is n-th ordinary homology for any $n \in \mathbb{N}$.[1] Then persistent Betti numbers, also referred to as barcodes, are the Betti numbers of persistent homology.

The foundation of persistent homology, namely parameter dependent homology, goes back to Vietoris' work on homology for metric spaces in Vietoris 1927. In that work, the basic idea was to use a metric and a real-valued filtration parameter in order to associate homology classes to a metric space.[2] In the meantime, these methods have been further developed and applied to hyperbolic geometry in Hausmann 1996 and Gromov 1987, §1.7, and then adapted to data analysis in Edelsbrunner, Letscher, and Zomorodian 2002/Robins 2002.

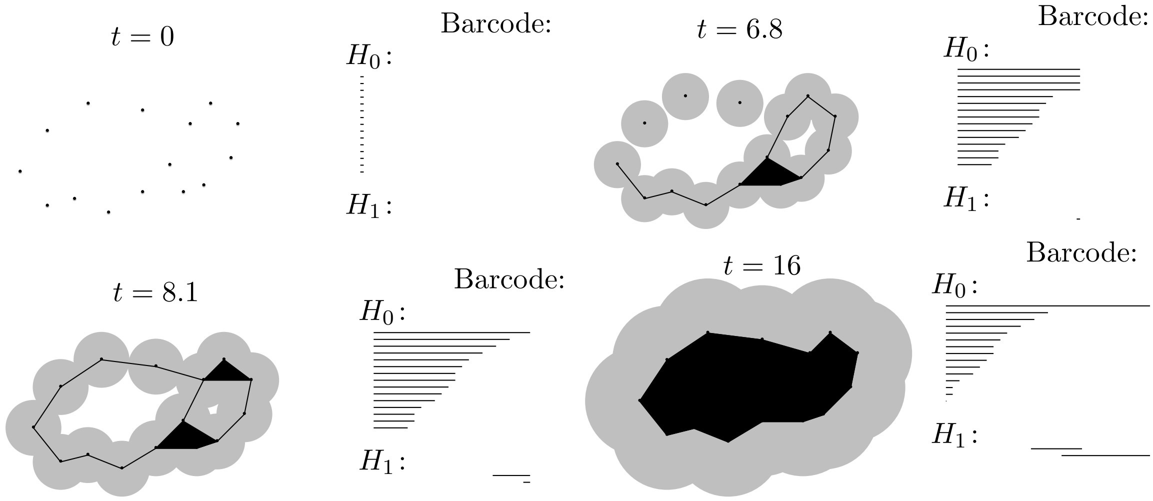


FIGURE 1. Four time steps of a Vietoris–Rips complex. The balls of radius t are gray and the Vietoris–Rips complex at time t is black. The values for t correspond to the displayed radii in points (length unit). Note that in this example, the Vietoris–Rips complexes are of dimension 2 at times $t = 6.8$ and $t = 8.1$ and of higher dimension at time $t = 16$. At $t = \infty$ the Vietoris–Rips complex is isomorphic to the standard 14-simplex.

[1] In this work, 0 is a natural number.

[2] Recall that singular homology had not yet been invented at the time.

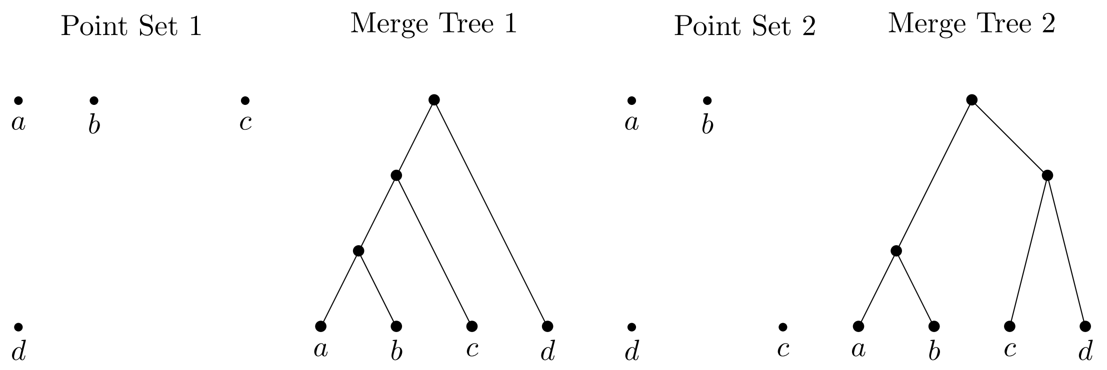


FIGURE 2. Two point sets and the merge trees of their corresponding Vietoris–Rips complexes. They both have the same 0-th barcode. The labels show which point corresponds to which leaf of the induced merge tree.

The reason why this relatively old concept is so useful in Topological Data Analysis is that in general it is not obvious how to associate a topological space to a data set in a meaningful way. One of today's standard approaches to this is the Vietoris–Rips complex. Considering a data set as a finite metric space allows to use an increasing real-valued parameter t to construct a persistent simplicial complex: the data points are the vertices and add higher simplices whenever their vertices have pairwise distance of at most t. See Figure 1 for a visualization. In Topological Data Analysis, it is not only relevant which non-trivial topological invariants a filtered space possesses but also for which range of parameters they are non-trivial, or in other words, how long they persist. In fact, in many applications, e.g. when using the Vietoris–Rips complex or the Čech complex, filtrations of the standard $n-1$ simplex are considered, where n is the number of data points. Hence, the filtration itself is usually more important than the, possibly homotopically trivial, space that is being filtered.

Thus, in Topological Data Analysis it is necessary to consider the filtration as part of the structure of a given space and not just as a tool for an investigation of spaces. Due to that line of reasoning, it is a current development that more and more authors refer to these concepts as ***persistence***, e.g. Botnan and Lesnick 2023, or ***persistent topology***, e.g. Ferri 2015, Elbers and Weygaert 2023, and Ghrist 2008, or ***persistent homotopy theory***, e.g. Jardine 2020. Thus, they slowly but steadily establish a new subfield of topology and geometry, which we will also call persistent topology throughout this thesis. Inside geometry and topologt, persistent topology somewhat lies in between combinatorial topology and homotopy theory. Persistent topology is different from homotopy theory in the sense that even though homotopy invariants are studied, the changes of these invariants along filtrations are studied rather than the invariants of the space being filtered. Persistent topology differs from combinatorial topology because more general filtrations than skeletons of CW complexes are considered. We give a more detailed introduction to persistent topology from a Topological Data Analysis point of view in Section 2.

Persistent topology, or more generally methods from Topological Data Analysis, have been used in a wide range of applications, mostly in other sciences. One of the concepts from persistent topology that is central to this thesis is the concept of merge trees. Merge trees are the persistent version of the path components of a filtered space. There are many different ways how to axiomatize and construct merge trees, some of which are explained in Section 2. The general idea of merge trees is to capture the hierarchical structure of the evolution of path components of a given filtered space. See Figure 2 for an example. The 0-th barcode factors over the merge tree via the "elder rule": We construct bars induced by paths from leaves

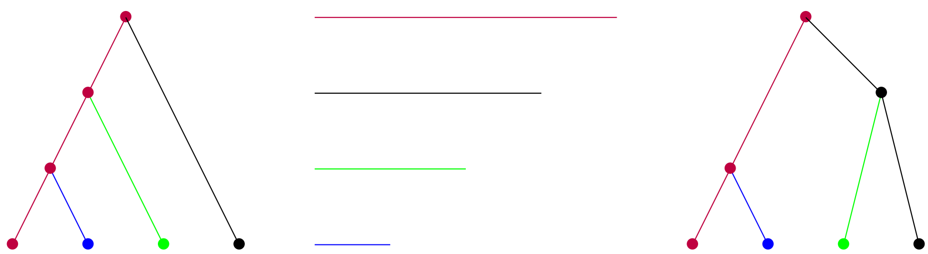


FIGURE 3. The barcode induced by the two merge trees in Figure 2 and how it can be seen within the merge trees via the elder rule. In this example, all choices of which path continues are non-canonical.

to inner nodes. Whenever two such paths meet, the older one with respect to the filtration value persists. If two paths of the same length meet, each choice leads to the same number of bars with the same birth and death time. See Figure 3 for an example. In Definition 5.30 we give a condition under which this choice can be made canonically. Although merge trees have originally been defined to approximate contour trees in computer graphics,[3] they have proven to be useful in their own right in applications, e.g. in Oesterling, Heine, G. H. Weber, Dimitry Morozov, et al. 2017, Baryshnikov 2019, Yan et al. 2019, Tralie et al. 2022, and Engelke et al. 2021.

One important family of problems in Topological Data Analysis and persistent topology is the family of inverse problems. Inverse problems investigate what kind of information causes certain invariants to distinguish objects and how different objects can be that induce the same value under these invariants. In a sense, inverse problems in TDA are similar to classification problems in topology, such as the classification of closed surfaces up to homeomorphism by Euler characteristic and orientability. A difference between inverse problems in TDA and classification problems in topology is that in TDA it is a priori usually not clear up to which notion of equivalence one wants to classify data sets, respectively the spaces/objects that represent data sets. Therefore, it is preferable to solve inverse problems in a more flexible way that allows to slightly change the involved invariants and the corresponding notion of equivalence while maintaining the classification result. Such approaches open up opportunities to modify and enhance topological invariants in order to adjust them to the problem at hand.

In visualization, more explicit instances of inverse problems are known as reconstruction problems. They answer the question of what information is needed to visualize a geometric object, i.e. to reconstruct it explicitly. This is also where the "ancestor" of the merge tree,[4] the contour tree, originates from: an embedded version of the Morse complex and the contour tree are sufficient to reconstruct[5] graphs of scalar functions on $\mathbb{R}^n$.

While we consider problems from persistent topology, our main tool for studying these problems is discrete Morse theory. Discrete Morse theory has been introduced in Forman 1998 as a combinatorial version of smooth Morse theory for a conjectured simpler proof of the s-cobordism theorem in the category of piecewise-linear manifolds. The basic idea of discrete Morse theory is to use certain well-behaved maps from abstract CW complexes to totally ordered sets in order to study the simple homotopy type and the combinatorial dynamics

[3] See Shinagawa, Kunii, and Kergosien 1991, Tarasov and Vyalyi 1998 and Carr, Snoeyink, and Axen 2003.

[4] See Carr, Snoeyink, and Axen 2003.

[5] See Shinagawa, Kunii, and Kergosien 1991, Kweon and Kanade 1994, Carr, Snoeyink, and van de Panne 2010 for Morse-theoretic techniques in graph reconstruction, and Heine et al. 2016 and Liu et al. 2016 for surveys on topological methods in visualization.

of said CW complexes. Discrete Morse functions have a combinatorial version of a gradient field associated to them, namely induced acyclic matchings on the Hasse diagrams of the corresponding face posets. The cells that are not matched by the combinatorial gradient field are referred to as critical, and play a role similar to critical points in smooth Morse theory. As in the smooth case, combinatorial gradient fields induce a combinatorial notion of flow lines that allow to compute combinatorial Morse homology. As explained in Mischaikow and Nanda 2013, combinatorial Morse homology is a more efficient approach to computing the homology of simplicial complexes than simplicial homology, given that one can provide a "good" discrete Morse function, i.e. one with as few critical cells as possible.

In Chapter II and Chapter III, we follow a convention first introduced by B. Benedetti to slightly misuse the term "discrete Morse function" in the sense that we use it for a certain notion of generic discrete Morse functions, which we call Morse–Benedetti functions in Chapter IV.[6]

Discrete Morse theory has many applications in both pure[7] and applied[8] mathematics. We further extend applications of discrete Morse theory to persistent topology in the three projects that make up this thesis:

1. Project Descriptions

The scientific work for this thesis has been conducted in three different projects, which resulted in three separate articles, which in turn appear as three different chapters in this thesis. Here, we present an overview over the three different projects.

1.1. On merge trees and discrete Morse functions on paths and trees.

Status of Publication.
This article is fully published as Brüggemann 2022.

Summary.
Inspired by the algorithmic construction of chiral merge trees[9] induced by discrete Morse–Benedetti[10] functions on trees and the open question about a possible representation of chiral merge trees by discrete Morse–Benedetti functions on paths raised in Johnson and Scoville 2022, we investigate the inverse problem between Morse–Benedetti functions on trees and their induced chiral merge trees. In order to do this, we introduce the notions of Morse orderings and Morse labelings as additional structure on merge trees. Moreover, we use words associated with paths in merge trees to construct two specific Morse orders that exist on any merge tree: the index Morse order Brüggemann 2022, Definition 3.3 and the sublevel-connected Morse order Brüggemann 2022, Definition 4.1. Then we use Morse labelings induced by Morse orders to construct critical discrete Morse functions on paths that represent given merge trees. In order to solve the inverse problem using these constructions, we further introduce two notions of equivalence between discrete Morse functions on trees: symmetry equivalence and component-merge equivalence. While symmetry equivalences are only generated by extensions of simplicial automorphisms of connected components of sublevel complexes, the more general component-merge equivalences also allow re-attachments of connecting edges

[6] In Chapter II and Chapter III only Morse–Benedetti functions appear, which is why the slight misuse of notation is forgivable. In Chapter IV on the other hand, both concepts appear and it is important to distinguish them there.

[7] See e.g. Arone and Brantner 2021 and Paolini and Salvetti 2021.

[8] See e.g. Delfinado and Edelsbrunner 1993, Delfinado and Edelsbrunner 1995, Robins, Wood, and Sheppard 2011, Dłotko and Wagner 2012, Mischaikow and Nanda 2013 and Maria and Schreiber 2023.

[9] Referred to just as merge trees in Johnson and Scoville 2022 and Brüggemann 2022/Chapter II

[10] Called a discrete Morse function in Johnson and Scoville 2022 and Brüggemann 2022/Chapter II

between connected components of sublevel complexes. Altogether, we obtain the following two main results, where DMF_P^{crit} denotes the set of critical discrete Morse functions (dMfs) on paths, MlT denotes the set of Morse-labeled merge trees (ML tree), Φ denotes discrete Morse function induced by a Morse labeled merge tree Brüggemann 2022, Definition 3.20, and $M(\_\,,\_)$ denotes the ML tree induced by a discrete Morse function Brüggemann 2022, Construction 2.14:

Theorem 5.4.
The induced labeled merge tree $M(\_\,,\_)$ and the induced dMf Φ define maps $M(\_\,,\_)\colon DMF_P^{crit} \leftrightarrow MlT\colon \Phi$ that are inverse to each other in the sense that:

(1) for a discrete Morse function (P, f) with only critical cells, the dMf $\Phi(M(P, f), \lambda_f)$ is symmetry-equivalent to (P, f), and
(2) for an ML tree (T, λ), the ML tree $M(\Phi T, f_\lambda)$ is isomorphic to (T, λ).

Theorem 5.6.
The induced labeled merge tree $M(\_\,,\_)$ and the induced dMf Φ define maps $M(\_\,,\_)\colon DMF_X^{crit} \leftrightarrow MlT\colon \Phi$ that are inverse to each other in the sense that:

(1) for any dMf (X, f) with only critical cells, the dMf $\Phi(M(X, f), \lambda_f)$ is component-merge-equivalent to (X, f), and
(2) for any ML tree (T, λ), the ML tree $M(\Phi T, f_\lambda)$ is isomorphic to (T, λ).

1.2. On cycles and merge trees.

Status of Publication and Author Statement.
This article is not yet accepted for publication. Its current preprint version is Brüggemann and Scoville 2023. The project "On cycles and merge trees" is collaborative work of the author of this thesis (referred to as the first author) and Nicholas A. Scoville (referred to as the second author). Both authors discussed all the contents of "On cycles and merge trees", and reviewed and edited all its written parts. The mathematical ideas, details, and initial formulations for the different sections have been developed separately. In particular, the first author is mainly responsible for sections 2, 3, 5 and 6. The second author is mainly responsible for section 4. The introduction was written jointly by both authors.

Summary. Continuing the investigation of the inverse problem between Morse–Benedetti functions and their induced merge trees, we consider the case of Morse–Benedetti functions on graphs. In order to find a complete description of the fiber, we enhance merge trees to what we call ***generalized merge trees***: a notion of merge trees that is able to capture cycle birth information via additional inner nodes. We extend the notions of Morse orderings and Morse labelings to generalized merge trees, as well as symmetry equivalences and component-merge equivalences to Morse–Benedetti functions on arbitrary finite graphs. As a result, we obtain the following:

Theorem 3.1.
Let DMF_{graphs}^{crit} denote the set of component-merge equivalence classes of discrete Morse functions with only critical cells on multigraphs. Let gMlT denote the set of isomorphism classes of generalized Morse labeled merge trees. Then the induced discrete Morse function Φ, see Definition 2.8, and the induced Morse labeled merge tree $M(\_\,,\_)$, see Definition 2.7, define maps $M(\_\,,\_)\colon DMF_{graphs}^{crit} \leftrightarrow gMlT\colon \Phi$ that are inverse of each other in the sense that:

(1) for any discrete Morse function (X, f) with only critical cells, the discrete Morse function $\Phi(M(X, f), \lambda_f)$ is cm-equivalent to (X, f), and
(2) for any generalized Morse labeled merge tree (T, λ), we have $M(\Phi T, f_\lambda) \cong (T, \lambda)$.

1. PROJECT DESCRIPTIONS

Moreover, we found a counting criterion for when exactly a generalized merge tree is represented by a critical discrete Morse–Benedetti function on a simple graph, and when such a simple graph can be embedded in the plane:

Theorem 4.1.
Let T be a generalized merge tree. Then there exists a simple graph X and a discrete Morse function $f\colon X \to \mathbb{R}$ such that $M(X, f) = T$ if and only if for every $c \in C(T)$,

$$|C(T(c_u))| < \frac{(\ell(c_u) - 2)(\ell(c_u) - 1)}{2}.$$

Furthermore, X can be made planar if and only if

$$|C(T(c_u))| < 2 \cdot \ell(c_u) - 5.$$

Moreover, using the induced generalized merge tree, we provide an algorithm to construct the spanning tree induced by a discrete Morse–Benedetti function on a graph. Then we use said spanning tree to optimize the Morse–Benedetti function by canceling pairs of critical cells.

1.3. On the parameter space of discrete Morse functions.

Status of Publication. This article is not yet accepted for publication. Its current preprint version is Brüggemann 2023.

Summary. This project is about describing the structure of the space of discrete Morse functions and how it relates to the spaces of smooth Morse functions, merge trees and barcodes. We reformulate the definitions of discrete Morse functions and Morse–Benedetti functions in terms of two hyperplane arrangements, the Morse arrangement[11] and the braid arrangement, on the space of discrete functions on a given CW complex. Using these hyperplane arrangements, we define the space of discrete Morse functions, the space of Morse–Benedetti functions and the parameter space of discrete Morse functions under shifting of function values. Moreover, we provide a framework for relating spaces of discrete Morse functions on different complexes to one another.

In addition, we explain how a result by Jean Cerf, namely Cerf 1970, Section 3.2, Definition 5, Lemma 1, can be interpreted and extended to maps between path components of the space of Morse–Smale functions on a manifold and the spaces of discrete Morse functions on CW decompositions induced by Morse–Smale functions. The main results of the comparison to smooth Morse theory of this work are:

Theorem 4.5.
Let f, g be Morse–Smale functions on M which are contained in the same path component $\mathcal{N}$ of the space of Morse–Smale functions $\mathcal{MS}(M)$ on M, i.e. $g \in \mathcal{N}(f) \subset \mathcal{MS}(M)$. Let ϕ_f, ϕ_g be the flows induced by f, g, respectively. Then there is a bijection $\psi\colon \operatorname{Cr}(f) \xrightarrow{\cong} \operatorname{Cr}(g)$ such that there is a flow line of ϕ_f connecting two critical points $c_1, c_2 \in \operatorname{Cr}(f)$ if and only if there is a flow line of ϕ_g connecting $\psi(c_1), \psi(c_2) \in \operatorname{Cr}(g)$.

Corollary 4.6.
For any compact finite-dimensional Riemannian manifold M, all Morse–Smale functions which are in the same path component of $\mathcal{MS}(M)$ induce the same CW decomposition M_f of M, up to cell equivalence. That is, there is an up to cell equivalence well-defined CW decomposition associated to any path component of the space of Morse–Smale functions on M.

Proposition 4.10.
Let $f, g \in \mathcal{MS}(M)$ be two Morse–Smale functions such that f arises from g by canceling

[11] See Definition 3.9.

a pair of critical points. Then the induced cellular decompositions M_f and M_g are simply homotopy equivalent by a simple collapse $M_g \to M_f$/ a simple extension $M_f \to M_g$ of the two cells which correspond to the canceled critical points of the corresponding sublevel complex. In particular, the map $M_f \to M_g$ is non-degenerate in the sense of Definition 3.21.

Corollary 4.11.
Let $\mathcal{N}_1 < \mathcal{N}_2 \in \pi_0(\mathcal{MS}_{smo}(M))$ be path components which are related by cancellations of critical points. Then $\mathcal{A}(M_{\mathcal{N}_1})$ canonically embeds into $\mathcal{A}(M_{\mathcal{N}_2})$ by Lemma 3.26. Moreover, we have an induced map $\mathcal{M}(M_f) \to \mathcal{M}(M_g)$ given by Definition 3.29.

We also recognize that the space of discrete Morse matchings $\mathfrak{M}(X)$[12] can be canonically identified as a subcomplex of the intersection poset $\mathcal{L}(X)$ of the Morse arrangement.

Proposition 5.3.
Let X be a CW complex and let $\mathcal{A}$ be the Morse arrangement on $\mathbb{R}^X$. Then there is a canonical embedding of posets $\mathfrak{M}(X) \subset \mathcal{L}(\mathcal{A})$.

Furthermore, we give a definition for spaces of merge trees and barcodes in a similar fashion and realize the merge tree induced by a discrete Morse function and the barcode induced by a merge tree as continuous maps between the corresponding spaces. In order to do this, we introduced metrics and hyperplane arrangements on the spaces of merge trees and barcodes that are closely related to the corresponding ones on the space of discrete Morse functions.

Proposition 5.16.
The Euclidean edit distance d from Definition 5.15 is a pseudo metric on the set of merge trees Mer, the set of strict merge trees $Mer_<$, and the set of well-branched merge trees Mer_{wb}.

Theorem 5.18.
Let X be a regular CW complex. Then the map $M\colon \mathcal{M}(X) \to Mer$ from Definition 5.8, which maps a discrete Morse function to its induced merge tree, is continuous.

Proposition 5.31.
The map $B\colon Mer_{wb} \to Bar$ from Definition 5.30, which maps a well-branched merge tree to its induced barcode, is continuous.

[12]See Definition 5.1.

2. From Topological Data Analysis to Persistent Topology

The purpose of this section is to shed some light on the so-called pipeline of Topological Data Analysis and to explain how it leads to the study of filtered, or as they are sometimes called in the field, persistent spaces. Before we get into the mathematical details, we want to explain the general concept of data analysis. We refer to Mathar et al. 2020 as a standard reference for data analysis.

Data analysis is the process of extracting information from data. In that sense, data and information are rather imprecise terms: basically anything from measured quantities to time-dependent data to images to collections of words, anything can be data. To perform mathematical data analysis, one first has to interpret the data in a mathematical way: for example, as a vector in Euclidean space, as a time series, as a digital image, or more abstractly as just a set. Depending on the type of data, it may seem more or less straightforward to decide which mathematical object is suitable for a given data set. In any case, the quality of this initial step of assigning an elementary mathematical object to data is on its own usually difficult to evaluate. Hence, we (and the literature) will assume that the data sets of interest already exist in some mathematically precise form. The evaluation whether an appropriate mathematical object was used to represent the data is usually done with statistical methods after the results of the mathematical data analysis have been interpreted.

Given a mathematical object that represents a data set, the next non-trivial question is how to choose a compatible mathematical model that is better suited for the underlying structure of the data set given by the application. By this we mean that the data points should be part of, or an approximation to, a richer mathematical structure, like a space. Probably the best-known concept in this direction is linear regression: we interpret the data points as a sampling of a line with an error given by a standard deviation. In Topological Data Analysis, we attempt to model data sets with more general spaces than lines, and use topological invariants of such spaces as features of the corresponding data sets. The idea is that data often has an inherent shape, which carries information that might be helpful for understanding the data set.

This approach is morally justified by the fact that many real-world phenomena in various sciences have been successfully described using mathematical equations, such as differential, polynomial, and linear equations. Since solutions of equations define geometric objects, such as manifolds and varieties, the topology of the solutions may contain desirable information and be worth studying.

Remark 2.1. Given the above reasoning, it may seem more immediate to apply geometric methods to data analysis. In fact, when applying Topological Data Analysis, one usually also applies geometric methods and preserves geometric information to some extent. It is mostly the use of topological invariants that justifies the term "Topological Data Analysis". The difference between topological and geometric data analysis is sometimes a fine one, which is why some authors might refer to the same techniques as either geometric or topological. The difference in opinion usually comes from focusing on a different part of the technique. The use of topological techniques instead of working purely with geometric techniques can be justified in (at least) two ways:

- In dimensions greater than 1, most geometric methods, e.g. parameterizations, become computationally expensive very quickly. This means that for high-dimensional large[13] data sets, we currently lack the computational power needed to access these data sets with geometric methods.
- Geometric methods would be able to capture all the information of a geometric model of a data set. Usually, some of the information contained in a data set is due to

[13] In terms of number of data points.

unwanted aspects of the data set. For example, from many data sets one can derive information about the process used to generate the data, simple enumerations of data points, or other circumstantial information which is not the information of concern. By applying topological methods, one can discard information in a controlled way, following the heuristics that unwanted information is more often contained in specific values of parameters and distances, and the information of interest is more likely to be contained in the general structure of the data set.

Once a model for a data set is fixed, one wants to compute features of the data set and interpret them in order to obtain some information about the data set. As mentioned before, we focus on topological spaces as models and topological invariants as features. See Chazal and Michel 2021 as a reference. After the features have been computed, one can either try to interpret them directly or apply statistical methods and/or machine learning. In practice, topological methods always compete with more direct applications of statistics and machine learning, mathematical models from other fields of mathematics (e.g. spectral methods for time series analysis) as well as direct interpretations of data sets. Topological Data Analysis usually outperform classical methods in cases where data sets either are already equipped with some geometric/topological meaning or are too large/high dimensional/unintuitive for classical techniques. See Carlsson 2009, Ferri 2015 and Carlsson and Vejdemo-Johansson 2022, Part III for surveys of applications of Topological Data Analysis. Moreover, it is often possible to combine topological methods with other techniques, which often yields even better results than any of the individual approaches on its own. See Ballester, Casacuberta, and Escalera 2024 and Hensel, Moor, and Rieck 2021 for applications of Topological Data Analysis in machine learning.

Figure 4 provides an overview of the pipeline of TDA.

There are several cases where techniques can only be evaluated as part of the complete pipeline by comparing input data and extracted information/interpretation. However, there is one particular step in the pipeline of Topological Data Analysis that can be investigated purely mathematically and independently of the other steps: the step between filtered spaces/complexes and their persistent invariants. Similar in spirit to classification results in topology, e.g. the classification of surfaces by orientability and Euler characteristic, one can investigate inverse problems and try to point out which (persistent) spaces are exactly distinguished by a certain (persistent) topological invariant.

Investigations of inverse problems lead to a better understanding of which properties of persistent spaces are detected by which invariants, which helps data analysts to choose the right invariants for different data sets. Choosing the right invariant is actually more difficult than it sounds: depending on the prior knowledge about a data set, the only possible approach may be to try out some techniques and choose different ones for a second analysis based on the results of the first investigation.

In mathematical terms, we can phrase the step between filtered, or persistent, spaces and persistent invariants as follows:

Definition 2.2. A ***persistent space*** is a family of topological spaces X_t indexed over a real parameter t, together with inclusion maps $X_t \hookrightarrow X_{t'}$ for filtration values $t \leq t' \in \mathbb{R}$.

For this work, the most important examples of persistent spaces are sublevel filtrations of CW complexes. In order to make sense of this, consider the following:

Definition 2.3. Let X be a topological space and let $f\colon X \to \mathbb{R}$ be a continuous function. The ***sublevel filtration*** induced by f is the persistent space $X_t := f^{-1}(-\infty, t]$.

In this work, we mostly consider sublevel filtration induced by (discrete) Morse functions.

It is also common to express the concept of persistent spaces as functors from the category $\mathbb{R}$, interpreted as a totally ordered set, to the category of topological spaces and embeddings.

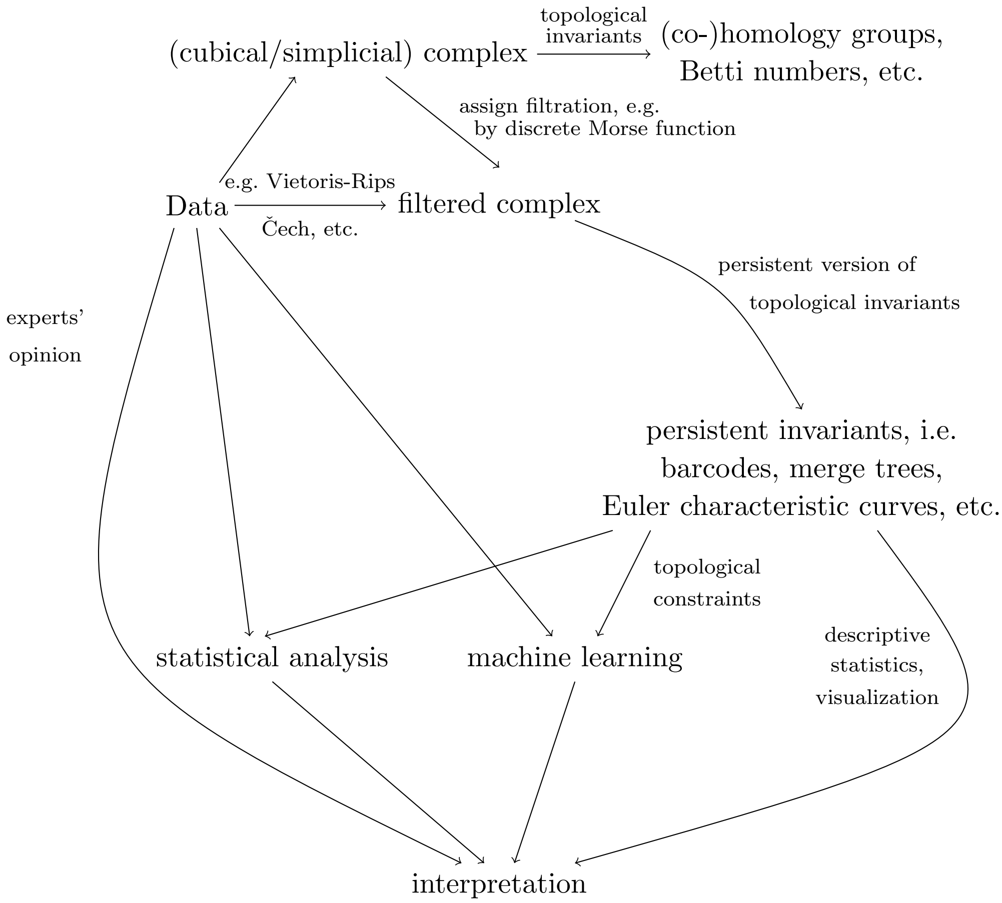


FIGURE 4. The pipeline of Topological Data Analysis.

From this point of view, many persistent invariants of persistent spaces can be interpreted as concatenations of persistent spaces with any functorial topological invariant. Thus, studying inverse problems means studying the fibres of such functors between categories of persistent spaces and persistent invariants, e.g. persistence modules. There are multiple works in the literature on investigations of various instances of inverse problems, see e.g. Curry 2019, Leygonie and Tillmann 2022, and Cyranka, Mischaikow, and Weibel 2020.

Since data analysis often requires adjustments of specific models, it is desirable to investigate inverse problems in a way that allows for similar adjustments, such that classification results can be preserved under such adjustments. That is, a solution to an inverse problem is more useful if it can be adapted directly to different mathematically precise versions of the same model.

Moreover, some persistent invariants factor over others, e.g. 0-barcodes factor over merge trees, Euler curves factor over the collection of barcodes of all dimensions. It is helpful when such factorizations are reflected in solutions to inverse problems because it allows data analysts to discard certain information of the data set in a controlled way. From a slightly different point of view, such solutions to inverse problems can tell in detail what kind of information might be lost by considering a coarser invariant, such as Euler curves, instead of a finer invariant, e.g. barcodes or (higher) merge trees. In some situations, data analysts may be forced to use coarser invariants due to the computational demands of finer invariants.

In this way, the purely mathematical investigation of inverse problems in persistent topology helps to improve decision making for the choice of persistent invariants of models

2. FROM TOPOLOGICAL DATA ANALYSIS TO PERSISTENT TOPOLOGY

of data. Moreover, solutions to inverse problems help to interpret the results of topological methods in data analysis.

3. Complexes

Before we discuss discrete Morse theory, we recall a few basic definitions from combinatorial topology and point out some subtle details between different notions of complexes. That is, we want to make clear under which conditions abstract complexes already contain the same homotopical information as their geometric counterparts, even though geometric complexes allow for more morphisms than abstract complexes. As a standard reference for more details, we refer to Lundell and Weingram 1969. We start with the most elementary concept of a complex, namely that of cell complexes:

Definition 3.1 (Lundell and Weingram 1969, Definition 1.1). Let X be a set. A ***cell structure*** on X is a pair $(X, \mathfrak{X})$ where $\mathfrak{X}$ is a collection of set maps of closed Euclidean cells $D^n \subset \mathbb{R}^n, n \in \mathbb{N}$, into X satisfying the following conditions:

(i) If $\varphi\colon D^n \to X$ belongs to $\mathfrak{X}$, then φ is injective on the interior $\operatorname{int}(D^n)$ of D^n.
(ii) The images $\{\varphi(\operatorname{int}(D^n)) | \varphi \in \mathfrak{X}, n \in \mathbb{N}\}$ partition X, that is, they are disjoint and have union X.
(iii) If φ has domain D^n, then $\varphi(\partial D^n) \subset X^{(n-1)} := \bigcup_{\substack{\psi\colon D^k \to X,\\ \psi \in \mathfrak{X}, k<n}} \psi(\operatorname{int}(D^k))$.

We call $\varphi \in \mathfrak{X}$ a ***characteristic map*** for the n-***cell*** $\sigma^{(n)} := \varphi(D^n) \subset X$. We say that a cell σ is a ***face*** of a cell τ if $\sigma \subset \tau$. The set $X^{(n-1)}$ in (iii) is called the ***n-1 skeleton*** of $(X, \mathcal{X})$.

We say that two cell structures $\mathfrak{X}, \mathfrak{X}'$ on X are ***strictly equivalent*** if there is a bijection $\mathfrak{X} \leftrightarrow \mathfrak{X}'$ such that characteristic maps that correspond to each other are related by a reparametrization of their domain D^n.

A ***cell complex*** is a pair $(X, \mathcal{X})$, where X is a set and $\mathcal{X}$ is a strict equivalence class of cell structures. A characteristic map is called ***regular*** if it is a bijection. A cell complex is called ***regular*** if all characteristic maps are regular.

Remark 3.2. The interior of D^0 is again D^0. This is in particular important for property (ii). Moreover, condition (iii) ensures that higher dimensional cells are always attached to lower dimensional cells in the filtration induced by the n-skeleta.

Note that up to this point, we have only considered sets without any topology. Although there are several ways to topologize cell complexes[14] we will focus exclusively on CW complexes and special cases of them. Before continuing with topology, we recall the concept of face posets, which can also be defined in a purely combinatorial way:

Definition 3.3. Let X be a cell complex. The ***face poset*** $F(X)$ of X is as a set $F(X) := \{\sigma | \sigma \text{ is a cell of } X\}$ and has relations $\sigma \leq \tau$ whenever σ is a face of τ.

Considering a cell complex to be a colimit of cells, the face poset represents the diagram over which the colimit is defined. However, this diagram does not necessarily know the attaching maps, and, therefore, neither the cell structure. The importance of this missing information will become clearer in the case of CW complexes.

Motivated by the notion of face posets, the concept of abstract complexes has been established.

Definition 3.4 (Alexandrov 1956, Chapter IV, §1.7). An ***abstract complex*** is a partially ordered set K together with a strictly monotone function $\mathrm{d}\colon K \to \mathbb{N}$.

[14] See Lundell and Weingram 1969, Chapter 1.

Definition 3.5. Let P be a poset. The ***Hasse diagram*** $D(P)$ associated with P is the directed graph with elements of P as vertices and an arrow $a \to b$ whenever $a \geq b$ in P.[15] For cell complexes X, we shorten the notation of the Hasse diagram of the face poset of X to $D(X)$.

Example 3.6. We consider the following example Figure 5 of a CW complex and its face poset.

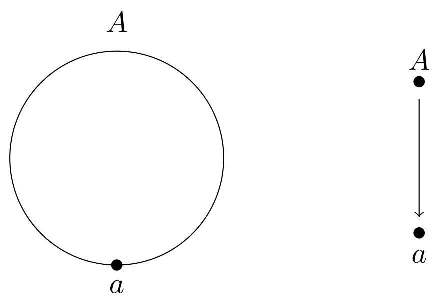


Figure 5. A non-regular cell structure on S^1 and the Hasse diagram of the corresponding face poset.

Remark 3.7. Face posets of cell complexes canonically carry the structure of abstract complexes by setting $\mathrm{d} := \dim$. Nonetheless, not every abstract complex is the face poset of a CW complex, e.g. three minimal elements that are all less than one maximal element. Moreover, various cell complexes may induce the same face poset: for example, there is a cell structure on $\mathbb{R}P^2$ that has the same face poset as D^2 with one 0-cell, one 1-cell and one 2-cell. Since the attaching map of the respective 2-cell is injective for D^2 but a non-trivial quotient map for $\mathbb{R}P^2$, the cell structures are not equivalent.

In computer image analysis, the concept of abstract complexes has found its way into applications in the context of digital topology.[16] As a result, it is not unusual to find only the term "complex" in the literature on computer images when authors want to refer to abstract complexes. Nevertheless, the term abstract complex has basically disappeared from the literature about pure topology due to the more topological concept of CW complexes, which we recall below, and the observation that finite abstract complexes carry the same information as finite topological spaces that satisfy the T_0 separation property. In Alexandroff 1937, basically all the necessary arguments are given for the proof that finite T_0 spaces with continuous maps and posets with order-preservings maps are isomorphic categories. In modern language, the isomorphisms of categories involved are called Alexandroff topology and specialization preorder. Moreover, it turns out that this isomorphism of categories extends to finite topological spaces and preordered sets.

In algebraic topology, the notion of CW complexes, originally introduced in Whitehead 1949, has become standard for cell decompositions of topological spaces.

Definition 3.8 (Lundell and Weingram 1969, Chapter II Definition 1.1). A Hausdorff space X is a ***CW complex*** with respect to a family of cells $\mathfrak{X}$ provided:

[15] It is a matter of convention whether the arrows in the Hasse diagram follow the lesser-equal relation or the greater-equal relation. We chose the greater-equal relation because then the arrows in the Hasse diagram will be consistent with combinatorial gradient flows of discrete Morse functions later on.

[16] See Klette and Rosenfeld 2004.

(i) the pair $(X, [\mathfrak{X}])$ is a cell complex such that each cell $\sigma \in \mathfrak{X}$ has a continuous characteristic map;
(ii) the space X has the weak topology with respect to $\mathfrak{X}$;[17]
(iii) the cell complex is closure finite.[18]

Remark 3.9. Definition 3.8 is, with minor reformulations made by the authors of Lundell and Weingram 1969, the original one presented in Whitehead 1949. There is a similar definition for CW complexes in the literature that uses skeletal filtrations instead of cell structures as presented above. It is argued in tom Dieck 2008, page 205, 8.2.6 and 8.3.8 that the two definitions are equivalent.

Similar to cell complexes, we also have the notion of regular CW complexes.

Definition 3.10 (Forman 1998, Definition 1.1). Let X be a CW complex and let σ be a face of a cell τ. We say σ is a ***regular face*** of τ if

(i) the characteristic map $\varphi\colon D^n \to \sigma$ is a homeomorphism, and
(ii) $\overline{\varphi^{-1}(\sigma)}$ is a closed n-ball.

A CW complex is called ***regular*** if all faces are regular.

For completeness, we also include the notion of an abstract simplicial complex.

Definition 3.11. An ***abstract simplicial complex*** is a set X of non-empty subsets of a set of vertices X_0[19] that is downward closed, i.e. if $\sigma \in X$, then all non-empty subsets of σ are also in X. Let X be an abstract simplicial complex and let $\sigma, \tau \in X$. We call σ and τ ***simplices*** of X and we call σ a ***face*** of τ if $\sigma \subset \tau$. A ***simplicial map*** $f\colon X \to Y$ between abstract simplicial complexes is a map $f\colon X_0 \to Y_0$ of the corresponding sets of vertices such that the images of the vertices of simplices span simplices.

Remark 3.12. An abstract simplicial complex is canonically an abstract complex, see Definition 3.4: the dimension of a simplex $\sigma = \{v_0, \ldots, v_n\}$ is n and the partial order is given by inclusion.

There is also a geometric version of simplicial complexes.

Definition 3.13. A ***geometric n-simplex*** σ is the convex hull of affinely independent $n+1$ points, called ***vertices*** of σ, in a Euclidean vector space. A ***face*** of a geometric simplex σ is the convex hull of any subset of σ's vertices. A ***geometric simplicial complex*** is a set X of geometric simplices such that

(i) every face of a simplex of X is also in X, and
(ii) the non-empty intersection of two simplices σ, τ of X is a common face of σ and τ.

A ***simplicial map*** $f\colon X \to Y$ between geometric simplicial complexes is a map of sets of simplices such that for every simplex σ the convex hull of the images of the vertices of σ spans a simplex of Y.

Remark 3.14. Geometric simplicial complexes are in particular regular cell complexes. The cell structure is given by the simplices. Abstract simplicial complexes are abstract complexes in a straightforward manner. Hence, as usual in the literature, we will sometimes use the terms abstract simplicial complexes and their face posets interchangeably when it is clear from the context that we are referring to simplicial complexes.

---

[17] That is, the topology coinduced by the characteristic maps.
[18] That is, for every point $x \in X$, the intersection of all subcomplexes of X that contain x is finite.
[19] The set X_0 can be thought of as the 0-skeleton of X, hence the notation.

The main purpose of this section so far is to elaborate on the following correspondence: for simplicial complexes, there are two constructions that are in a sense inverse to each other: the geometric realization and the induced abstract simplicial complex. The set of simplices of a geometric simplicial complex is an abstract simplicial complex. Conversely, any finite abstract simplicial complex can be realized as a simplicial subcomplex of a geometric $|X_0| - 1$ simplex. If one combines the two constructions, one obtains the original object back up to isomorphism. Moreover, the induced abstract simplicial complex and geometric realization define an equivalence of categories between abstract and geometric simplicial complexes and simplicial maps. Hence, we will sometimes just refer to simplicial complexes without specifying whether we mean abstract or geometric ones if the difference is not important for the matter at hand.

Even though this construction is pretty much straightforward, there is another construction that does a similar thing but is more suitable for generalization, namely the face poset and the order complex.

DEFINITION 3.15. Let P be a poset. The ***order complex***[20] of P is the simplicial complex that has the underlying set of P as its set of vertices and the n simplices are defined by chains of length $n + 1$ in P.

If one takes the order complex of the face poset of a simplicial complex X, one obtains the barycentric subdivision $\mathrm{Sd}(X)$ of X. In the case of regular CW complexes X, we call this construction the derived subdivision $\mathrm{Sd}(X)$.[21] In both cases, the respective order complex is homeomorphic to the initial complex X but the respective homeomorphism is in general not compatible with the cell structure in the sense that the barycentric subdivision might not be simply homotopy equivalent to X. For example, it is an open problem, called the Zeeman conjecture, whether the products of contractible 2-dimensional CW complexes with the interval are collapsible. But in Adiprasito and Benedetti 2019 the authors show that in certain situations a number of barycentric subdivisions make the product collapsible.

Example 3.16. We consider the following example of a regular CW complex and its derived subdivision in Figure 6. Faces in the CW decomposition on the left-hand-side correspond to elements in the Hasse diagram of the face poset in the middle. Chains of comparable elements in the face poset induce cells in the order complex on the right.

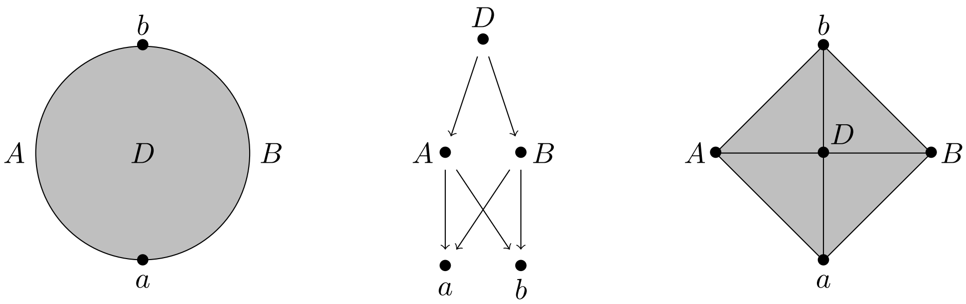


FIGURE 6. A regular CW decomposition of D^2, the Hasse diagram of the corresponding face poset, and the derived subdivision/order complex.

All in all, the face poset and the order complex allow us to move back and forth between posets/abstract complexes and regular CW/simplicial complexes without changing homotopy

[20]This construction is also referred to as the classifying space of the poset.

[21]We do not distinguish in notation between the derived and barycentric subdivision because the latter is a special case of the former and it is clear from the context which one is being applied.

type. This is particularly important in the setting of discrete Morse theory, where discrete Morse functions and combinatorial gradient fields are defined on face posets rather than on geometric complexes. As a consequence, applying discrete Morse theory is more convenient for regular CW complexes than for arbitrary CW complexes, although the theory is defined in the generality of CW complexes as we will see below. Nonetheless, it should be noted that this correspondence only holds at the level of objects, as there are more cellular maps between regular CW complexes than there are order preserving maps between their face posets, even up to homotopy.

An even stronger connection between combinatorial and topological methods is reflected in the equivalence of homotopy categories of simplicial sets and CW complexes induced by the singular simplicial set of a space and the geometric realization of simplicial sets[22]. However, this correspondence is not used in applied topology because one of the functors involved, namely taking the singular simplicial set, is not easily accessible by computational methods.

4. Discrete Morse Theory

Discrete Morse theory has been introduced by Forman in Forman 1998 as a potentially better suited framework for proving the s-cobordism theorem for piecewise-linear manifolds. In the original article, Forman introduced discrete Morse functions for arbitrary finite CW complexes:

Definition 4.1 (Forman 1998, Definition 2.1). Let X be a finite CW complex and let $F(X)$ be the face poset of X. A ***discrete Morse function*** on X is a function $f\colon F(X) \to \mathbb{R}$ such that for every p-dimensional cell $\alpha^{(p)} \in X$ we have

(1) $\#\{\beta^{(p+1)} \supset \alpha | f(\beta) \leq f(\alpha)\} \leq 1$,
(2) $\#\{\gamma^{(p-1)} \subset \alpha | f(\gamma) \geq f(\alpha)\} \leq 1$, and
(3) $f(\alpha) < f(\beta)$ whenever $\alpha \subset \beta$ is not regular.

Cells for which the inequalities in (1) and (2) are both strict are called ***critical***.

Remark 4.2. In order to avoid confusion, we remark that according to Forman a discrete Morse function on a finite CW complex is actually a function on its face poset. It is common practice in the literature to restrict to the setting of simplicial complexes (or cubical complexes), where one can consider abstract complexes directly without loss of information, and thus avoid this possible confusion. Considering abstract complexes, i.e. face posets, instead of geometric ones can be extended to regular CW complexes because, as argued in the previous section, the homotopy types of regular CW complexes are determined by their face posets. In the setting of regular CW complexes, condition (3) of Definition 4.1 can be dropped for obvious reasons.

Lemma 4.3 (Forman 1998, Lemma 2.5)**.** *The inequalities (1) and (2) of Definition 4.1 cannot both be equalities.*

Lemma 4.3 implies that every discrete Morse function $f\colon F(X) \to \mathbb{R}$ induces a partial matching on the Hasse diagram $D(X)$.

There is a generic version of discrete Morse functions, originally defined in Benedetti 2016 for simplicial complexes but the generalization to arbitrary CW complexes is straightforward.

Definition 4.4 (Benedetti 2016, Section 2.1). Let X be a CW complex. A ***Morse–Benedetti*** function on X is a function on the face poset $f\colon F(X) \to \mathbb{R}$ that fulfills for cells σ, τ that:

[22] See Goerss and Jardine 2009, Theorem 11.4 for a concise proof and Quillen 1976, Chapter 2 for the original introduction of the correspondence. There is an even stronger version of this result in terms of a Quillen equivalence of model categories, see Hovey 2007, Theorem 3.6.7, Theorem 2.4.23

Monotonicity: If $\sigma \subset \tau$, we have $f(\sigma) \leq f(\tau)$.
Semi-injectivity: $|f^{-1}(\{z\})| \leq 2$ for all $z \in \mathbb{R}$.
Generacity: If $f(\sigma) = f(\tau)$, then either $\sigma \subseteq \tau$ or $\tau \subseteq \sigma$ holds.
Regularity: If $\sigma \subset \tau$ is irregular, then $f(\sigma) < f(\tau)$.

It is straightforward to see that Morse–Benedetti functions are always discrete Morse functions and, thus, induce partial matchings on the Hasse diagram of the face poset. Critical cells of Morse–Benedetti functions are then exactly those cells that are unique preimages of values.

Basically every concept from smooth Morse theory has a corresponding concept in the discrete version: discrete Morse functions induce combinatorial gradient fields, which in turn induce gradient flow lines. These gradient flow lines can be used to define boundary maps for a chain complex generated by the critical cells of a given discrete Morse function. Furthermore, the usual theorems such as the weak and strong Morse inequalities hold and in the filtration induced by a discrete Morse function the homotopy type changes only at critical values. For an overview of the most important features, we refer to Forman 2002. In the remainder of this section, we will concentrate on the notions that are relevant to this thesis.

As mentioned before, discrete Morse functions induce partial matchings on face posets. We interpret such matchings as a combinatorial version of vector fields.

Definition 4.5 (Forman 2002, Definition 3.3). Let X be a CW complex and let $D(X)$ be the Hasse diagram of its face poset. A ***discrete vector field*** V on X is a partial matching on $D(X)$. That is, if $\prec$ denotes the cover relation[23] in $F(X)$, then V is a partition of the set of cells of X into singletons and pairs of cells that are related to one another by a cover relation.[24]

If f is a discrete Morse function on X, we call the discrete gradient field that is induced by f[25] the ***combinatorial gradient field*** of f.

Associated to any discrete vector field on a CW complex X, a V***-path*** on X is a sequence of cells

$$\alpha_0, \beta_0, \alpha_1, \beta_1, \ldots \alpha_l, \beta_l, \alpha_{l+1}$$

such that $(\alpha_i \prec \beta_i) \in V$ and $\alpha_{i+1} \prec \beta_i$ for all $i = 0, \ldots, l$.

This leads to the following useful characterization:

Theorem 4.6 (Forman 2002, Theorem 3.5)**.** *A discrete vector field V is the combinatorial gradient field of a discrete Morse function if and only if there are no non-trivial closed V-paths.*

Thus, combinatorial gradient fields of discrete Morse functions are the same as acyclic matchings on the Hasse diagram of the face poset. Hence, it is common to consider acyclic matchings instead of discrete Morse functions, which allows for a purely combinatorial treatment, as long as one works at least with regular CW complexes.

It is desirable to find discrete Morse functions, or acyclic matchings, that match as many cells as possible. Depending on the point of view, discrete Morse functions with few critical cells can be used either to simplify the CW complex directly or to simplify the chain complex induced by such a discrete Morse function. In both situations, discrete Morse theory helps to simplify computations, which allows for better efficiency.

Due to the discrete Morse inequalities, the number of critical cells cannot be less than the Betti numbers of the corresponding CW complex. In fact, depending on the CW complex at hand, it might be the case that only exist discrete Morse functions with strictly more critical cells than the sum of the Betti numbers exist. A discrete Morse function with as

[23] i.e. $\sigma \prec \tau \Leftrightarrow \sigma \subset \tau$ and there is no $\sigma' \in F(X)$ such that $\sigma \subset \sigma' \subset \tau$.
[24] In regular CW complexes this is equivalent to pairs of faces of codimension one.
[25] See Lemma 4.3.

few critical cells as possible for a given CW complex is called ***optimal***. Unfortunately, finding optimal discrete Morse functions is known to be NP-complete,[26] i.e. there are only algorithms with more than polynomial runtime that can guarantee finding optimal discrete Morse functions. This makes finding optimal discrete Morse functions unfeasible in many situations. Nevertheless, there are better algorithms in some special cases, e.g. an algorithm with linear run time for finding optimal discrete Morse functions on discrete 2-manifolds was presented in Lewiner, Lopes, and Tavares 2003a.

[26]See e.g. Hajebi and Javadi 2023.

CHAPTER II

On Merge Trees and Discrete Morse Functions on Paths and Trees

1. Summary

This article is published as Brüggemann 2022. It is the sole work of the author of this thesis, Julian Brüggemann.

This work addresses the inverse problem between discrete Morse functions on trees and their induced merge trees in the sense of Johnson and Scoville 2022. Although merge trees were originally defined as approximations to contour trees, they can be interpreted more algebraically as the persistent version of the path components invariant π_0. Therefore, the inverse problem between discrete Morse functions and their induced merge trees consists of studying the relation between filtrations induced by discrete Morse functions and all possibilities of evolution of connected components along such filtrations.

We consider intermediate steps of this inverse problem by inducing the notions of Morse-ordered merge trees, Definition 2.17, and Morse-labeled merge trees, Definition 2.19. These concepts rely on the specific version of merge trees used. The version of merge trees we use differs from other notion of merge trees used in the literature by the additional structure of chirality: For each inner node, the two child nodes have additional labels L and R. Johnson and Scoville 2022 uses chirality to define an algorithm for the merge tree induced by a discrete Morse function, see Johnson and Scoville 2022, Theorem 3.5. It turns out that this algorithm implicitly constructs certain labelings on induced merge trees, see Proposition 2.20, which we axiomatize as Morse labelings. Moreover, Morse labelings induce certain total orders on the nodes of merge trees, which we axiomatize as Morse orders, see Definition 2.17.

The use of chirality in this work is closely related to the Elder rule, see Curry, DeSha, et al. 2024, Example 2.14: the Elder rule constructs barcodes from merge trees. The idea is that whenever two branches of a merge tree meet, the elder one, i.e. the one with the smaller minimum, should continue. This way, the Elder rule provides a decomposition of merge trees into barcodes. It is common to use this decomposition as a rule how to draw merge trees: the bars of the corresponding barcode decomposition should be uninterrupted lines in the merge tree. The chirality of chiral merge trees also provides a rule for drawing them: starting whith the root, child nodes with chirality L go to the left and child nodes with chirality R go to the right. The construction of the induced merge tree given in Johnson and Scoville 2022 has the property that the two different drawing rules coincide.

We introduce two more important constructions: on one hand, we construct Morse orders on merge trees. See Definition 3.3 for the index Morse order and Definition Definition 4.1 for the sublevel connected Morse order. On the other hand, we construct representing discrete Morse functions on paths, see Definition 3.20. Due to the chirality as part of the structure of merge trees, we can use paths from arbitrary nodes to the root to define words, called path words, with letters L,R given by chirality of nodes the path passes. See Definition 3.1 for path words. We use these path words to define Morse orders, see Definition 3.3 and 4.1. Given a Morse order, one can define compatible Morse labelings using natural numbers in the Morse order. We then define the simplex orders on paths, see Definition 3.15, and on merge trees, see Definition 3.11, and establish an isomorphism of totally ordered sets between the two.

This allows us to push Morse labelings to paths, which creates discrete Morse functions, see Definition 3.20.

With this setup, we proof the main theorems in Section 5:

Theorem 5.4.
The induced labeled merge tree $M(\_,\_)$ *and the induced dMf* Φ *define maps* $M(\_,\_)\colon DMF_P^{crit} \leftrightarrow MlT\colon \Phi$ *that are inverse to each other in the sense that:*

(1) for a discrete Morse function (P, f) *with only critical cells, the dMf* $\Phi(M(P, f), \lambda_f)$ *is symmetry-equivalent to* (P, f)*, and*
(2) for an ML tree (T, λ)*, the ML tree* $M(\Phi T, f_\lambda)$ *is isomorphic to* (T, λ)*.*

Theorem 5.6.
The induced labeled merge tree $M(\_,\_)$ *and the induced dMf* Φ *define maps* $M(\_,\_)\colon DMF_X^{crit} \leftrightarrow MlT\colon \Phi$ *that are inverse to each other in the sense that:*

(1) for any dMf (X, f) *with only critical cells, the dMf* $\Phi(M(X, f), \lambda_f)$ *is component-merge-equivalent to* (X, f)*, and*
(2) for any ML tree (T, λ)*, the ML tree* $M(\Phi T, f_\lambda)$ *is isomorphic to* (T, λ)*.*

Here, DMF_P^{crit} denotes the set of critical discrete Morse functions (dMfs) on paths, MlT denotes the set of Morse-labeled merge trees (ML tree), Φ denotes discrete Morse function induced by a Morse labeled merge tree Brüggemann 2022, Definition 3.20, and $M(\_,\_)$ denotes the ML tree induced by a discrete Morse function Brüggemann 2022, Construction 2.14.

In section 6, we compare our results and used concepts to a similar framework used in Curry 2019 to solve a similar inverse problem between Morse-like functions on the interval and chiral merge trees. Furthermore, we discuss possible future directions and applications of this work at the end of Section 6.

2. Copyright and License Notice

On merge trees and discrete Morse functions on paths and trees
Author: Julian Brüggemann
Publication: Journal of Applied and Computational Topology
Publisher: Springer Nature
Date: Nov 26, 2022

Changes. Page numbers of the surrounding document, i.e this thesis, have been added at the bottom of every page.

3. Published Version

Journal of Applied and Computational Topology (2023) 7:103–138
https://doi.org/10.1007/s41468-022-00101-w

On merge trees and discrete Morse functions on paths and trees

Julian Brüggemann[1]

Received: 29 December 2021 / Revised: 29 June 2022 / Accepted: 7 September 2022 /
Published online: 26 November 2022

Abstract

In this work we answer an open question asked by Johnson–Scoville. We show that each merge tree is represented by a discrete Morse function on a path. Furthermore, we present explicit constructions for two different but related kinds of discrete Morse functions on paths that induce any given merge tree. A refinement of the used methods allows us to define notions of equivalence of discrete Morse functions on trees which give rise to a bijection between equivalence classes of discrete Morse functions and isomorphism classes of certain labeled merge trees. We also compare our results to similar ones from the literature, in particular to work by Curry.

Keywords Discrete Morse theory · Combinatorial algebraic topology · Merge trees

Mathematics Subject Classification 57Q70 (Primary) · 05C05 (Secondary) · 05C90

Contents

Julian Brüggemann
brueggemann@mpim-bonn.mpg.de

[1] Max Planck Institute for Mathematics, Bonn, Germany

Springer

1 Introduction

Discrete Morse theory is a combinatorial version of the classical smooth Morse theory. It was originally developed by Forman (1998).

In discrete Morse theory, topological properties of simplicial complexes X are analyzed by considering discrete Morse functions $f: X \to \mathbb{R}$. These topological properties can in turn be used to obtain cell decompositions of X with fewer cells. A good introduction to the topic is found in Forman (2001).

Merge trees are used in Morse theory in order to keep track of the development of connected components of sublevel sets $X_a := f^{-1}(-\infty, a]$ of a given Morse function $f: X \to \mathbb{R}$. Since the sublevel sets form a filtration of X, merge trees can be seen as a combinatorial description of the persistent connectivity of X. In particular, every branching in the induced merge tree $M(X, f)$ corresponds to a pair of connected components of a sublevel set $X_{a-\varepsilon}$ that merge to one connected component in a sublevel complex X_a of higher level.

Initially, merge trees were introduced to topological data analysis as an approximation to the Reeb graph, respectively contour tree, in Carr et al. (2003). The Reeb graph is a graph that keeps track of the connected components of level sets of any given filtered manifold. In applications, the data set is often interpreted as a sampling of the graph of a function rather than a more general manifold, which is why the Reeb graph is often actually a tree, the so-called contour tree.

Among other implementations of techniques of smooth Morse theory, computational methods for the Reeb graph have originally been introduced in Shinagawa et al. (1991) in order to handle surfaces embedded in 3D with the help of computers. Later on, several ways to compute and apply contour trees of data sets have been discussed in many articles, e.g. Kweon and Kanade (1994), van Kreveld et al. (1997), Tarasov and Vyalyi (1998), and Carr et al. (2003). Among other applications, merge trees have been used in visualization, e.g. in Oesterling et al. (2013), Weber et al. (2007), Oesterling et al. (2017), and Yan et al. (2019). Surveys about applications of merge trees and other concepts in visualization can be found in Heine et al. (2016) and Liu et al. (2016). Furthermore, a certain version of chiral merge trees has been used in Baryshnikov (2019) to analyze asymmetries of time series.

We focus on the more structural and theoretical side of merge trees, in particular the connection to discrete Morse theory. We consider a specific construction for merge trees induced by discrete Morse functions on trees which was introduced in Johnson and Scoville (2022). We use this construction to gain a better understanding of the set of discrete Morse functions, the set of merge trees, and the relationship between the two.

Similar work has been done in Curry (2019) for the relationship between Morse-like functions on the interval and a related version of merge trees. Furthermore, our work is in some sense similar to parts of Curry et al. (2021), with the difference being that the authors of Curry et al. (2021) consider the relationship between merge trees and their induced barcodes instead. It seems reasonable to adapt the following terms from both of these articles and say that we consider a different instance of the “fiber of the persistence map”, respectively a different instance of the “inverse problem” of the “persistence map”. These names refer to the fact that merge trees and barcodes are invariants of filtered spaces rather than just spaces. In this context, the name "persistent" became popular due to persistent homology as it appears in topological data analysis. Similar to the authors of Curry (2019) and Curry et al. (2021), we are interested in finding out what information is lost by considering our invariant at hand, the induced merge tree, instead of the given data, in our case a discrete Morse function, and how this information might be re-obtained. Such knowledge might be helpful to investigate the space of merge trees and the space of discrete Morse functions in future work. Furthermore, a good understanding of the fiber of the persistence map is useful for topological data analysis because it hints at features which might be lost due to the chosen invariant. Moreover, insights about the inverse problem might be helpful to enhance the chosen invariant in a way such that it preserves certain desired features of the data set.

We respond to an open question asked in Johnson and Scoville (2022) by showing that every merge tree is represented by a discrete Morse function (dMf). In particular, for any given merge tree we construct a dMf on a path as a representative of the isomorphism class defined by said merge tree:

Theorem 5.5 *Let T be a merge tree. Then there is a path P such that $T \cong M(P, f_{io}) \cong M(P, f_{sc})$ holds as merge trees where f_{io} denotes the induced index-ordered dMf* (Definition 3.20) *and f_{sc} denotes the sublevel-connected dMf* (Definition 4.5) *on P.*

In particular, the discrete Morse function from Theorem 5.5 can be chosen to be index-ordered (Definition 2.1) or sublevel-connected (Definition 2.31).

The main tool for the construction is the corresponding Morse order (Definition 3.3), that is, the index Morse order (Definition 3.3) or the sublevel-connected Morse order (Definition 2.31) on the nodes of a given merge tree T. The index Morse order defines leaf nodes to be strictly less than inner nodes. Among leaf nodes and among inner nodes, the index Morse order is defined by using a twisted version of length-lexicographical order on the set of path words (Lemma 3.16) that correspond to the respective nodes. The path words are defined by the chirality of the nodes of the shortest path from the root to the corresponding node. For the sublevel-connected Morse order (Definition 4.1) we do not artificially distinguish between leaf nodes and inner nodes.

We use the index Morse order (Definition 3.3) to define the index Morse labeling (Definition 3.7) on the nodes of T. Together with the simplex order (Definition 3.11), which establishes a correspondence (Remark 3.18) between the nodes of T and the simplices of a path P, the index Morse labeling defines the induced index-ordered discrete Morse function on said path P.

In Sect. 2.2 we introduce several kinds of equivalence relations on the sets of discrete Morse functions with only critical cells on paths and trees. These equivalence relations allow us to identify equivalence classes of discrete Morse functions with isomorphism classes of Morse labeled merge trees:

Theorem 5.4 *The induced labeled merge tree $M(\_,\_)$ and the induced dMf Φ define maps $M(\_,\_)\colon DMF_P^{crit} \leftrightarrow MlT : \Phi$ that are inverse to each other in the sense that:*

(1) *For a dMf (P, f) with only critical cells, the dMf $\Phi(M(P, f), \lambda_f)$ is symmetry-equivalent to (P, f), and*

(2) *For an Ml tree* (T, λ)*, the Ml tree* $M(\Phi T, f_\lambda)$ *is isomorphic to* (T, λ)*.*

Theorem 5.6 *The induced labeled merge tree* $M(\_, \_)$ *and the induced dMf* Φ *define maps* $M(\_, \_)\colon DMF_X^{crit} \leftrightarrow MlT : \Phi$ *that are inverse to each other in the sense that:*

(1) *For any dMf* (X, f) *with only critical cells, the dMf* $\Phi(M(X, f), \lambda_f)$ *is cm-equivalent to* (X, f)*, and*
(2) *For any Ml tree* (T, λ)*, the Ml tree* $M(\Phi T, f_\lambda)$ *is isomorphic to* (T, λ)*.*

The construction of the discrete Morse function induced by a Morse labeling is similar to the construction of functions on the interval in Curry (2019). In particular, Theorem 5.4 is very similar to the result (Curry 2019, Prop 6.11). The use of Morse labelings in this work basically plays the role of the function $\pi : T \to \mathbb{R}$ from Curry (2019). Moreover, the simplex order is almost the same as the use of chirality in Curry (2019, Lem 6.4). But in this work, Morse orders, and in turn Morse labelings, have to satisfy a certain compatibility with the chirality, that is, property (2) of Definition 2.17.

The notion of merge trees we use originates from Johnson and Scoville (2022, Def 5) and differs from the one used in Curry (2019): A priori, merge trees T in the sense of Johnson and Scoville (2022) do not carry a height function $T \to \mathbb{R}$ as part of their data. Instead, the two children of each node have a chirality assigned to them as part of the tree's data. This means that for any two child nodes of the same parent node, it is specified as part of data which is the right and which is the left child. This version of chirality is also canonically assigned to merge trees induced by discrete Morse functions. In contrast, the chirality of chiral merge trees in the sense of Curry (2019) arises from a chosen orientation on the interval. We obtain a similar correspondence between the chirality of merge trees and orientations on paths using the simplex order, Definition 3.11. Apart from these differences, the notion of merge trees in the sense of Johnson and Scoville (2022) is closely related to the one from Curry (2019). Chirality in the sense of Johnson and Scoville (2022) is a specific version of the notion of chirality used in Curry (2019). In order to see this, we show that the construction of the merge tree $M(X, f)$ induced by a discrete Morse function f in the sense of Johnson and Scoville (2022) can be modified, see Proposition 2.20, to obtain a function $T \to \mathbb{R}$ from f, similarly to Curry (2019). This gives rise to the notion of Morse labelings, Definition 2.19, and Morse orders, Definition 3.3. It turns out that the use of chirality in Johnson and Scoville (2022) assumes a certain compatibility between Morse orders and the simplex order, whereas the use of chirality in Curry (2019) does not. As a result, the induced merge tree in the sense of Curry (2019) distinguishes between symmetry equivalences, Definition 2.42, whereas the induced merge tree in the sense of Johnson and Scoville (2022) identifies symmetry-equivalent discrete Morse functions with each other, see Proposition 2.48. We discuss this in a bit more detail at the end of Sect. 5.

In said discussion, we mention the notion of CMl trees, see Definition 5.9, which is as objects basically the same as the notion of merge trees from Curry et al. (2021, Def 2.2). However, the notion of combinatorial equivalence of labeled merge trees from Curry et al. (2021, Def 2.6) corresponds to a non-chiral version of shuffle equivalence

Definition 2.24 rather than an equivalence of the induced persistent set as in Curry (2019, 5.1), which is more similar to an isomorphism of Ml trees Definition 2.23. The chiral merge trees from Baryshnikov (2019, Def 2.1) are as objects also very similar to the chiral merge trees from Curry (2019) and, thus, differ similarly from our notion of labeled merge trees.

The aforementioned versions of labeled merge trees all have in common that their labelings need to be compatible with some other data inherent to the merge tree. In contrast to that, the labelings from Yan et al. (2019) can be quite arbitrary and might even assign multiple labels to a single node. Hence, our Ml trees a priori seem to be unrelated to the notion of labeled merge trees from Yan et al. (2019).

2 Preliminaries

We consider discrete Morse functions (dMf) on trees. Recall that trees are finite acyclic simple graphs. Furthermore, simple graphs are 1-dimensional simplicial complexes. Where feasible, we introduce the preliminaries in the broader generality they are usually defined in, rather than in the lesser generality we actually need for this work. We adapt most notations and conventions from Johnson and Scoville (2022). For simplicity, we assume all trees in this article to be non-empty. Similar to Johnson and Scoville (2022), we assume the dMfs to fulfill certain generic properties. In detail this means the following:

Definition 2.1 Let X be a simplicial complex. A map $f: X \to \mathbb{R}$ is called a *discrete Morse function (dMf)* if it fulfills the following properties for any pair of simplices $\sigma, \tau \in X$:

(i) $\sigma \subseteq \tau \Rightarrow f(\sigma) \leq f(\tau)$ (weakly increasing)
(ii) f is at most $2-1$
(iii) $f(\sigma) = f(\tau) \Rightarrow (\sigma \subset \tau \vee \tau \subset \sigma)$ (matching)

Simplices on which f is $1-1$ are called *critical*. Simplices which belong to the preimage of the same value are called *matched*. The set of critical simplices is denoted by $\mathrm{Cr}(f)$. Values of critical simplices under f are called *critical values* of f. A dMf is called *index-ordered* if for arbitrary critical simplices σ, τ the following holds: If $\dim(\sigma)$ is smaller than $\dim(\tau)$, then $f(\sigma) < f(\tau)$ holds.

Remark 2.2 The definition given above is not the most general definition of dMfs but rather assumes several generic properties. This means that any dMf in the sense of Forman (1998) can be modified by a Forman equivalence (see Johnson and Scoville, 2022, Def 4.1) to fulfill these properties, that is, without changing the induced Morse matching. As usual in the context of dMfs, we write $f: X \to \mathbb{R}$ although the map f is actually defined on the face poset of X.

In Nanda et al. (2018), a similar generic property is used to analyze flow paths induced by dMfs. The notion of faithful dMfs as defined in Nanda et al. (2018, Def 2.9) is almost the same as index-ordered dMfs in this article. The only difference is that for index-ordered dMfs matched cells have the same value, whereas for faithful dMfs the values of the matched boundary simplices are higher than the values on the corresponding matched co-boundary simplices.

Furthermore, since simple graphs and in particular trees are examples of simplicial complexes, the definition of dMfs can be applied to them as well.

Remark 2.3 Because f maps $\mathrm{Cr}(f)$ injectively into a totally ordered set, it induces a total order on $\mathrm{Cr}(f)$. We refer to this induced order whenever we speak of simplices being *ordered by* f.

Notation 2.4 We use the following conventions regarding notation:

- We depict dMfs on graphs by labeling the graph with the values of the dMf.
- Let G be a graph and let v be a node of G. By $G[v]$ we denote the connected component of G which contains v.

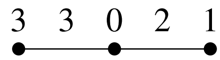
Definition 2.5
- Let X be a simplicial complex, $f: X \to \mathbb{R}$ a dMf and $a \in \mathbb{R}$. The *sublevel complex* of level a, denoted by X_a^f, is defined by $X_a^f := \{\sigma \in X \mid f(\sigma) \leq a\}$. If the referred dMf f is clear from the context, we drop the superscript f from the notation.
- The ordered critical values $c_0 < c_1 < \cdots < c_m$ induce a chain of sublevel complexes $X_{c_0}^f \subsetneq X_{c_1}^f \subsetneq \cdots \subsetneq X_{c_m}^f$. Within this chain, we refer by $X_{c_i - \varepsilon}^f$ to the complex that immediately precedes X_{c_i}.

Remark 2.6 The given definition of sublevel complexes differs from the standard one used in the literature. We make use of the fact that a dMf f being weakly increasing implies that X_a^f as defined above is already a subcomplex of X. If we wanted to consider the general definition of dMfs as introduced in Forman (1998), we would have to work with the smallest supercomplex of X_a^f in X instead. Taking the smallest supercomplex corresponds to additionally including all faces of simplices of X_a^f to make X_a^f a simplicial complex.

Lemma 2.7 *Let X be a finite simplicial complex and let $f: X \to \mathbb{R}$ be a dMf. Then f attains its minimum on a critical 0-simplex. Furthermore, the statement also holds for the restriction to any connected component of sublevel complexes.*

Sketch of Proof The statement follows by a proof by contradiction and properties (i) and (ii) of Definition 2.1. □

Remark 2.8 The analogous statement for the maximum of a dMf f on arbitrary 1-simplices is false, as the following example shows:



Here, the maximum is attained on a pair of matched simplices.

2.1 Merge trees

We briefly recapture preliminaries about merge trees as they are explained in Johnson and Scoville (2022). The basic idea is that merge trees keep track of the chronological development of the connected components of sublevel sets. We adapt the point of view of Johnson and Scoville (2022), that is, we consider them 'upside down'. Thus, the children will appear above their parent node. Afterwards we introduce additional structure that dMfs induce on their corresponding merge trees and consider notions of equivalence which arise from that structure.

Definition 2.9 (Merge Tree) A *merge tree* is a full rooted chiral binary tree T. In detail this means that T is a rooted tree fulfilling the properties of being binary and full, and that T has the extra datum of being chiral:

Full binary Each node of T has either zero or two children.

Chiral Each child node in T carries the extra datum, the so-called *chirality*, of being a left or a right child.

Morphisms of merge trees are morphisms of rooted binary trees which are compatible with the chirality.

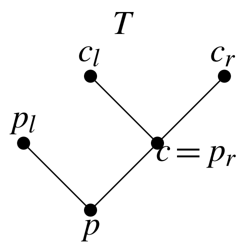
For rooted trees T we use the notions of *subtrees*, *ancestors* and *descendants* as they are commonly used in computer science.

Definition 2.10 For any node p of T, the *descendants* of p are defined inductively: A node c is a descendant of p if the parent node of c is a descendant of p or p itself.

A *subtree* of T is a subgraph of T that consists of exactly all of the descendants of some node p of T.

For a node p of T we call all nodes which lie on the shortest path between p and the root, including the root, the *ancestors* of p.

Notation 2.11 For an inner node c of T, we denote the left child of c with c_l and the right child of c with c_r. We illustrate this notation in the following example:



Remark 2.12 For full binary trees T with $i(T)$ inner nodes and $l(T)$ leaves it is a well-known result that $l(T) = i(T) + 1$ holds. It can be proved inductively.

Remark 2.13 The chirality of nodes will either be denoted by labels or indicated implicitly by embedding the merge tree on the page. Throughout the literature there are different notions of merge trees that are not always distinguished by name or notation. In this work, merge trees do not have explicit weights on edges. Moreover, merge trees in this work a priori do not carry a function to the real numbers. In that way, we distinguish between merge trees and Morse labeled merge trees which will be introduced in Definition 2.19. We use the chirality to obtain Morse labelings on unlabeled merge trees. This leads to Definition 3.3.

Construction 2.14 (Johnson and Scoville 2022, Thm 9) Let X be a tree and let $f: X \to \mathbb{R}$ be a dMf. The *merge tree induced by* f, denoted by $M(X, f)$ is constructed as follows:

Let $c_0 < c_1 < \cdots < c_m$ be the critical values of f that are assigned to 1-simplices. The associated merge tree $M(X, f)$ is constructed by induction over these critical values in descending order. Furthermore, we label the nodes of $M(X, f)$ in order to refer to them later. The label of a node n will be denoted by $\lambda(n)$.

For the base case we begin by creating a node $M(c_m)$ which corresponds to the critical 1-simplex in X labeled c_m and setting its label $\lambda(M(c_m))$ to (c_m, L).

For the inductive step, let $M(c_i)$ be a node of $M(X, f)$ that corresponds to a critical 1-simplex between two 0-simplices v and w. Define $\lambda_v := \max\{f(\sigma) | \sigma \in X_{c_i - \varepsilon}[v], \sigma \text{ critical}\}$ and $\lambda_w := \max\{f(\sigma) | \sigma \in X_{c_i - \varepsilon}[w], \sigma \text{ critical}\}$. Two child nodes of $M(c_i)$ are created, named n_{λ_v} and n_{λ_w}. Then label the new nodes $\lambda(n_{\lambda_v}) := \lambda_v$ and $\lambda(n_{\lambda_w}) := \lambda_w$. If $\min\{f(\sigma) | \sigma \in X_{c_i - \varepsilon}[v]\} < \min\{f(\sigma) | \sigma \in X_{c_i - \varepsilon}[w]\}$, we assign n_{λ_v} the same chirality (L or R) as $M(c_i)$ and give n_{λ_w} the opposite chirality. Continue the induction over the rest of the critical 1-simplices.

Remark 2.15 By construction, one of the following two cases holds for the labels λ_v and λ_w. They might be either critical values lower than c_i that are assigned to edges or critical values that are assigned to nodes. The two labels λ_v and λ_w do not necessarily belong to the same case.

Therefore, the nodes n_{λ_v} and n_{λ_w} will possibly be denoted as $M(c_j)$ and $M(c_k)$ for some $j, k < i$ in later steps of the induction. In particular, this means that the node which is considered in the first instance of the inductive step is c_m.

Remark 2.16 Although the induced merge tree as introduced above comes with a labeling, the labeling is not part of the data of the induced merge tree. This is one of the main differences between merge trees in Johnson and Scoville (2022) and the merge trees in Curry (2019).

We consider one way to keep some information provided by the induced labeling on $M(X, f)$.

Definition 2.17 Let T be a merge tree. We call a total order $\leq$ on the nodes of T a *Morse order* if it fulfills the following two properties for any subtree T' of T:

(1) The restriction $\leq_{|T'}$ attains its maximum on the root of T'.
(2) The restriction $\leq_{|T'}$ attains its minimum on the subtree with root p_l/p_r if L/R is the chirality of the root p of T'.

Moreover, we call a merge tree $(T, \leq)$ together with a Morse order $\leq$ a *Morse-ordered merge tree (Mo tree).*

Remark 2.18 Assuming property (2) of Definition 2.17 for every subtree T' with root p of T is equivalent to either of the following:

- For any subtree T' with root p of T, the minimum of $\leq_{|T'}$ has the same chirality as p.
- For any subtree T' with root p of T, all nodes on the shortest path between p and the minimum of $\leq_{|T'}$ have the same chirality as p.

The equivalence can be proved by an inductive argument over all nodes of the shortest path between p and the minimum.

Because Morse orders $\leq$ define in particular finite totally ordered sets, there are unique order-preserving isomorphisms $\lambda\colon (V(T), \leq) \xrightarrow{\cong} \{0, 1, \dots, i(T)+l(T)-1\} \subseteq \mathbb{N}_0 \subset \mathbb{R}$ for each Morse order. Conversely, each injective labeling $\lambda\colon T \to \mathbb{R}$ induces an order $\leq_\lambda$ on the nodes of T by usage of the total order on $\mathbb{R}$.

Definition 2.19 For Morse orders $\leq$ we call the map $\lambda_{\leq}\colon (V(T), \leq) \to \{0, 1, \dots, i(T) + l(T) - 1\}$ the *Morse labeling induced by* $\leq$. We call an arbitrary labeling $\lambda\colon T \to \mathbb{R}$ a *Morse labeling* if it induces a Morse order on T.

We call a merge tree (T, λ) with a Morse labeling $\lambda\colon T \to \mathbb{R}$ a *Morse labeled merge tree (Ml tree)*.

For a Mo tree $(T, \leq)$ we call the Ml tree $(T, \lambda_{\leq})$ the *Ml tree induced by* $(T, \leq)$.

Proposition 2.20 *Let $f: X \to \mathbb{R}$ be a dMf. The labeling which appears in Construction 2.14 induces a Morse order on $M(X, f)$. Hence, $M(X, f)$ canonically carries the structure of a Mo tree as well as an Ml tree.*

Proof It is proved in Johnson and Scoville (2022, Thm 9) that $M(X, f)$ is a merge tree. We only have to prove that the labeling induces a Morse order. By Remark 2.3 the set $\mathrm{Cr}(f)$ of critical values carries a total order induced by f. Since the critical values of f precisely define the labeling in Construction 2.14, the labeling induces a total order on the nodes of $M(X, f)$. It is only left to prove that this order is a Morse order.

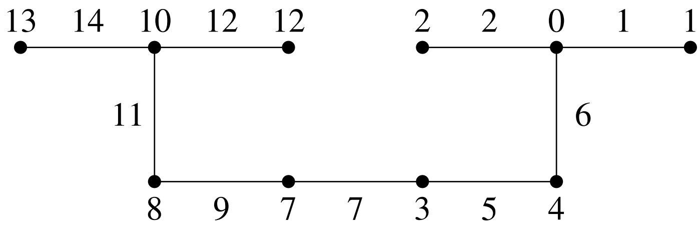
In the construction, each inner node of $M(X, f)$ corresponds to a critical 1-simplex and is labeled with the critical value of said critical 1-simplex. Since parent nodes are created before their child nodes are, and since the critical values are considered from highest to lowest, property (1) of Definition 2.17 is fulfilled. The rule in the construction which decides the chirality of the child nodes is exactly the same as property (2) of Definition 2.17. Hence, it is fulfilled by construction. □

We denote the canonical labeling of $M(X, f)$ by λ_f.

Remark 2.21 In the aforementioned proof, it becomes clear that both conditions of Definition 2.17 are necessary for a total order on $M(X, f)$ to be induced by a dMf f.

Morse orders are useful for the construction of dMfs that induce given merge trees T. We will see in Proposition 3.23 that for a total order on an arbitrary merge tree T condition (1) is sufficient for inducing a dMf in the sense of Definition 3.20 later on. Nonetheless, condition (2) is necessary to ensure that the induced dMf induces the given merge tree T.

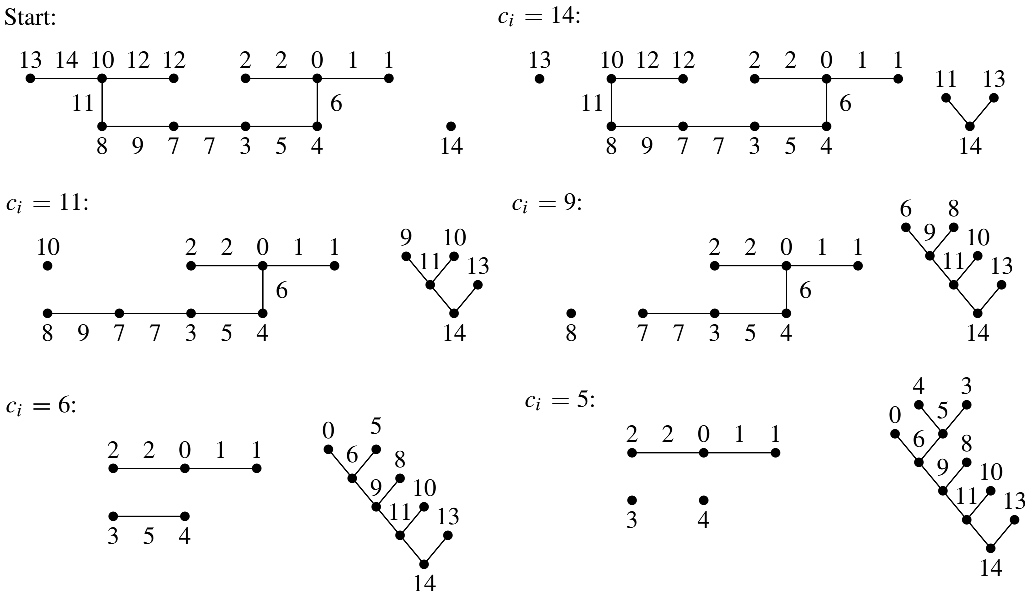
Example 2.22 We consider the following example of a dMf $f: X \to \mathbb{R}$:



The critical values on edges are:

$$5 < 6 < 9 < 11 < 14$$

We now show the construction algorithm of $M(X, f)$ visually by depicting $X_{c_i - \varepsilon}$ on the left and the part of $M(X, f)$ that is created up to the step corresponding to c_i on the right.



There are no more critical edges left, so the construction of $M(X, f)$ is finished.

Since there are different notions of merge trees in the literature and since the merge trees in our setting carry a lot of structure, there are multiple possibilities of how to define equivalences of merge trees. In the remainder of this section, we define and discuss some versions of equivalence of merge trees.

Since merge trees are defined to be chiral rooted binary trees, the obvious notion for isomorphisms of merge trees is isomorphisms of chiral rooted binary trees. In detail, this means bijections between the sets of nodes and the sets of vertices which map the root to the root and are compatible with the chiral child relation. For Mo trees (Definition 2.17) we have more notions of equivalence.

Definition 2.23 Let $(T, \leq)$ and $(T', \leq')$ be Mo trees. An *isomorphism* of Mo trees $(T, \leq) \cong (T', \leq')$ is an order-preserving isomorphism of the underlying merge trees.

Let (T, λ) and (T', λ') be Ml trees. An *isomorphism* of Ml trees is and isomorphism of the underlying merge trees over $\mathbb{R}$, that is, an isomorphism of merge trees $\varphi\colon T \to T'$ such that $\lambda' \circ \varphi = \lambda$.

Definition 2.24 Let (T, λ) and (T', λ') be Ml trees. A *shuffle equivalence* $(\varphi, \psi)\colon (T, \lambda) \to (T', \lambda')$ of Ml trees is a pair of an isomorphism of the underlying merge trees $\varphi\colon T \to T'$ and a bijection $\psi\colon \mathbb{R} \to \mathbb{R}$ such that

- $\psi \circ \lambda = \lambda' \circ \varphi$ holds,
- The restriction of ψ to values on leaves is order-preserving, and
- The restriction of ψ to values on inner nodes is order-preserving.

In the special case that the restriction $\psi_{|\mathrm{im}(\lambda)}\colon \mathrm{im}(\lambda) \to \mathrm{im}(\lambda')$ is an order preserving bijection, we call (φ, ψ) an *order equivalence*.

A *shuffle equivalence* $(T, \leq) \to (T', \leq')$ between Mo trees is an isomorphism φ of the underlying merge trees such that

- The restriction of φ to leaf nodes is order-preserving, and
- The restriction of φ to inner nodes is order-preserving.

Remark 2.25 The name of shuffle equivalences hints at the fact that two given total orders, one total order on the leaves, and another total order on the inner nodes, might be combined to produce a total order on all nodes in different ways, a bit like shuffling cards. Shuffle equivalence checks if two Morse orders arise from the same underlying orders by different ways of shuffling. But the necessity of ranking ancestors higher than descendants and being compatible with the chirality prevents arbitrary ways of shuffling two given orders on the leaves and inner nodes from producing Morse orders.

Shuffle equivalences induce by definition isomorphisms on the underlying merge trees. We now make the relationship between Mo trees and Ml trees precise.

Proposition 2.26 *The Morse labeling induced by a Morse order and the order induced by a Morse labeling define inverse bijections* $\mathrm{iMl}\colon MoT/_{\cong} \overset{\cong}{\leftrightarrow} MlT/_{\sim}\colon \mathrm{iMo}$ *where* $\sim$ *denotes order equivalence.*

Proof Let $(T, \leq)$ be a Mo tree. It is immediate that $(T, \lambda_{\leq})$ has the property that $\lambda_{\leq}$ induces $\leq$ as its induced order on T. Thus, the composition $\mathrm{iMo} \circ \mathrm{iMl}$ is the identity on MoT.

Let (T, λ) be an Ml tree. By definition, the labeling λ induces a Morse order $\leq_{\lambda}$ on T which makes $(T, \leq_{\lambda})$ a Mo tree. Since the induced Morse labeling $\lambda_{\leq_{\lambda}}$ by construction induces $\leq_{\lambda}$ as its induced order, it follows that (T, λ) and $(T, \lambda_{\leq_{\lambda}})$ are order equivalent. □

Corollary 2.27 *The induced Morse order* iMo *and the induced Morse labeling* iMl *induce a bijection between shuffle equivalences of Mo trees and shuffle equivalences of Ml trees.*

Proof The assignment iMl maps shuffle equivalences of Mo trees to shuffle equivalences of Ml trees by Definition 2.24.

Let $\lambda\colon T \to \mathbb{R}$ and $\lambda'\colon T' \to \mathbb{R}$ be shuffle-equivalent Morse labelings and let (φ, ψ) be the corresponding shuffle equivalence. Then by Definition 2.19 the map $\mathrm{iMo}(\varphi, \psi) = \varphi\colon (T, \leq_\lambda) \to (T', \leq_{\lambda'})$ has the property that the restriction of φ to leaf nodes is order-preserving, and the restriction of φ to inner nodes is order-preserving. Hence, $\varphi = \mathrm{iMo}(\varphi, \psi)$ is a shuffle equivalence of Mo trees. □

Remark 2.28 In particular, the aforementioned proposition and corollary mean that two Mo trees are isomorphic (respectively shuffle equivalent) if and only if the corresponding Ml trees are order equivalent (respectively shuffle equivalent) and vice versa.

2.2 Generic properties and equivalences of DMFs

In this subsection, we will take a closer look at generic properties that dMfs can be assumed to have. Furthermore we consider some notions of equivalences between dMfs.

A first example of a generic property is the notion of index-ordered dMfs as defined above. It is inspired by the eponymous notion from the smooth case. We use index-ordered dMfs in order to distinguish critical simplices by their dimension because merge trees have the same property: Critical 0-simplices appear as leaves whereas critical 1-simplices appear as inner nodes of the induced merge tree (see Construction 2.14). Thus, index-ordered dMfs seem to be especially suitable for working with merge trees.

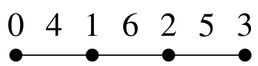
Nonetheless, index-ordered dMfs are not compatible with the structure of rooted subtrees. In detail, consider the following:

Remark 2.29 Let $f\colon X \to \mathbb{R}$ be an index-ordered dMf on a tree such that the following holds: The tree X has a critical 1-simplex σ such that the corresponding inner node p in $M(X, f)$ has two inner nodes c and c' as children. Then the image of f on at least one of the connected components corresponding to c or c' is not an interval in $f(\mathrm{Cr}(f))$.

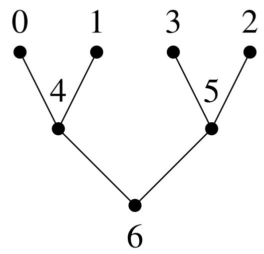
Example 2.30

The following is a small example:

(X, f)



$M(X, f)$



Neither the subtree with root labeled 4, nor the subtree with root labeled 5 is labeled with an interval in $f(\mathrm{Cr}(f)) \subseteq \mathbb{R}$.

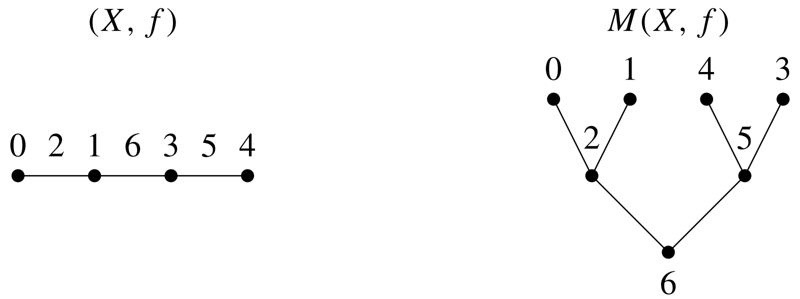
Still, we can assume compatibility with the structure of rooted subtrees as a property:

Springer

Definition 2.31 Let $f: X \to \mathbb{R}$ be a dMf. The function f is called *sublevel-connected* if for all critical 1-simplices v the set $f(X_{f(v)}[v])$ is an interval in $f(\mathrm{Cr}(f))$.

Remark 2.32 Since both 'index-ordered' and 'sublevel-connected' are properties that only rely on the values of f on critical simplices, they can easily be arranged without changing the partial matching if the tree X is finite. It is also possible to do this without changing the induced merge tree. One way to obtain a sublevel-connected dMf would be to choose a collapsing order for the induced matching such that the respective connected components of sublevel sets correspond to intervals in said collapsing order. But as seen in Remark 2.29, the two properties are in most cases mutually exclusive.

Example 2.33 Here is a possibility how to modify the dMf from the previous example in order to make it sublevel-connected without changing the induced merge tree:



Since dMfs can be modified to fulfill either of the two properties, one can always choose the one which is more convenient for the task at hand. Thus, we give two different constructions in this work, one for each property.

We recall that by Remark 2.3 each dMf $f: X \to \mathbb{R}$ induces an order on the 0-simplices of X and on the 1-simplices of X, respectively. This yields the following definition, which induces a merge-tree-invariant notion of equivalence between dMfs.

Definition 2.34 Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be dMfs on a tree X. We call f and g *shuffle-equivalent* if they have the same critical simplices and if they induce the same order on the critical 0-simplices as well as the same order on the critical 1-simplices.

Let (X, f) and (X', f') be two dMfs on trees. A *shuffle equivalence* $(\varphi, \psi): f \to f'$ between f and f' consists of a simplicial map $\varphi: X \to X'$ and a bijection $\psi: \mathbb{R} \to \mathbb{R}$ such that

- $\psi \circ f = f' \circ \varphi$,
- $\varphi_{|\mathrm{Cr}(f)}: \mathrm{Cr}(f) \to \mathrm{Cr}(f')$ is a bijection,
- The restriction of ψ to values on critical 0-simplices is order preserving, and
- The restriction of ψ to values on critical 1-simplices is order preserving.

In the special case that the restriction $\psi_{|\mathrm{Cr}(f)}: \mathrm{Cr}(f) \to \mathrm{Cr}(f')$ is an order preserving bijection, we call (φ, ψ) an *order equivalence*.

Remark 2.35 It is immediate that shuffle equivalence of dMfs is an equivalence relation. The name is inspired analogously as for the eponymous notion for Mo trees and Ml trees. Shuffle equivalence checks if two dMfs arise from the same underlying orders by different ways of shuffling. Nonetheless not all ways of shuffling given orders on the critical 0-simplices and critical 1-simplices produce a dMf because dMfs have to be weakly increasing.

Furthermore, shuffle equivalence, and in particular order equivalence, also considers dMfs to be equivalent if they only differ by scaling because a different scaling does not change the induced orders on simplices.

We split the definition of shuffle equivalences in two steps to simplify the proofs of the following propositions. It is immediate that for two dMfs on the same tree X there is a shuffle equivalence between them if and only if they are shuffle equivalent.

Proposition 2.36 *Let X be a tree and let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be two shuffle-equivalent dMfs. Then $M(X, f)$ and $M(X, g)$ are isomorphic as merge trees.*

The following two lemmas will be helpful for the proof of the proposition:

Lemma 2.37 *Let X be a tree and let $f: X \to \mathbb{R}$ be a dMf such that X has at least one critical 1-simplex. Then the function $f_{|\mathrm{Cr}(f)}$ attains its maximum on a critical 1-simplex.*

Sketch of Proof Backtracking gradient paths, see Forman (1998, Def 8.4), as long as possible leads to a local maximum which turns out to be a critical 1-simplex. □

Lemma 2.38 *Let X be a tree and let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be two shuffle-equivalent dMfs. We denote the critical 1-simplices of f and g by $c_0 < c_1 < \cdots < c_n$ where $<$ denotes the ordering induced by f or g, respectively. Then the connected components $X^f_{f(c_i)}[c_j]$ of sublevel complexes contain the same critical simplices as $X^g_{g(c_i)}[c_j]$ for all $j < i$.*

Sketch of Proof First we observe that restrictions of dMfs on trees to connected components of sublevel complexes are again dMfs on trees. With help of Lemma 2.37, it can be proved inductively that in the construction of $M(X, f)$ and $M(X, g)$ the same critical 1-simplices are considered in the same order. The statement then follows inductively. □

Proof of Proposition 2.36 We consider the construction of the induced merge tree (see Construction 2.14) and prove inductively the slightly stronger result that both functions yield isomorphic merge trees at every step of the construction. This implies that f and g induce isomorphic merge trees.

Since f and g impose the same order on the set of critical 1-simplices, the construction algorithm considers the same critical 1-simplices during the same steps for both functions. This already proves the base case. In particular, this means that the created root node corresponds to the same critical 1-simplex for both functions. Although the label of the root node might be different for the two dMfs, it does not affect the isomorphism type of the induced merge tree because the labeling is not part of the data of merge trees.

For the inductive step, we observe that in every step of the construction we consider a connected component of sublevel complexes $X^f_{f(c_i)}[c_j]/X^g_{g(c_i)}[c_j]$ that contains at least one critical 1-simplex, namely the one with the highest remaining critical value

c_j. Thus, by Lemma 2.38 in each step of the construction, the same critical simplices occur.

Assume we are at the step that considers the critical 1-simplex c_i. For the two new nodes which are created in the inductive step, two pieces of information are important for the isomorphism type of the induced merge tree, namely the chirality of the new nodes and which critical simplices the new nodes correspond to. The chirality of the new nodes affects the isomorphism type of the induced merge tree directly. The critical simplex corresponding to a child node c decides which connected component of the respective sublevel complex is used to build the subtree with root c and at which point said connected component will be subdivided next.

Both pieces of information are defined by the two connected components that belong to the boundary 0-simplices of c_i. The two child nodes correspond to the critical simplices with the highest critical values.

There are three cases:

(1) Both connected components contain at least one critical 1-simplex c_j.
(2) One connected component contains at least one critical 1-simplex c_j whereas the other one only contains one critical 0-simplex c.
(3) Each of the two connected components contains only one critical 0-simplex c.

It follows by Lemma 2.37 that in case (1) the corresponding 1-simplices with the highest critical values are critical 1-simplices c_j. In case (2) the same is true for the connected component that contains at least one critical 1-simplex. For the connected components in case (2) and (3) that only contain one critical 0-simplex, respectively, it is true that the critical 0-simplex is the only critical simplex in its corresponding connected component. Thus, the new nodes correspond to the same critical simplices for f and for g because both functions induce the same order on 1-simplices and because connected components only correspond to critical 0-simplices if they are the only critical simplex left in the corresponding connected component. Furthermore, connected components $X^f_{f(c_i)}[c]/X^g_{g(c_i)}[c]$ that only contain one critical 0-simplex do not have any influence on the induced merge tree $M(X, f)/M(X, g)$ because their corresponding nodes have already been created during a step that considered a critical 1-simplex with a higher critical value and they are not considered in later steps of the construction.

The chirality of the new nodes depends on the minimal values on critical simplices of the two respective connected components $X^f_{f(c_i)}[c_j]/X^g_{g(c_i)}[c_j]$ or $X^f_{f(c_i)}[c]/X^g_{g(c_i)}[c]$. By Lemma 2.7, these minima belong to critical 0-simplices. By assumption, f and g induce the same order on the critical 0-simplices, so the same 0-simplex is minimal with respect to both functions. Thus, f and g assign the same chirality to the new nodes. □

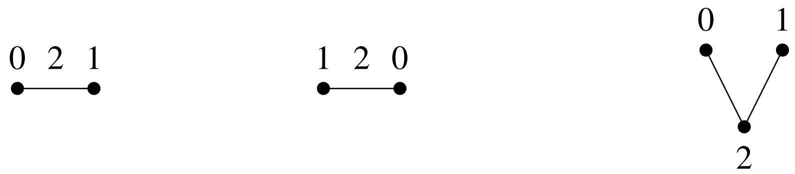
Remark 2.39 The functions f and g might induce different order relations between 0-simplices and 1-simplices. Therefore, in sublevel complexes that appear during the same step of the construction there might be different connected components that contain only one critical 0-simplex each with respect to the two dMfs. However, those connected components that contain only one critical 0-simplex and no critical 1-simplices do not affect the isomorphism type of the induced merge tree. This is because those connected components correspond to leaves of the merge tree which have already been created during the step that considered the critical 1-simplex between said connected components and other connected components. Furthermore, those connected components do not appear in later steps of the construction of the induced merge tree because they do not contain any critical 1-simplices.

Proposition 2.40 *Shuffle equivalences of dMfs induce shuffle equivalences of the induced Ml trees. Moreover, order equivalences of dMfs induce order equivalences of the induced Ml trees.*

Sketch of Proof Since φ is bijective on critical simplices and simplicial, it follows that φ induces a bijection between connected components of sublevel complexes. With this, it follows analogously to the proof of Proposition 2.36 that $M(X, f) \cong M(X', f')$ holds. The proof that the induced Morse labelings are shuffle equivalent is straightforward and only uses that the restrictions of ψ to 0-simplices and to 1-simplices are order preserving, and the compatibility between φ, ψ, f and f'. □

Remark 2.41 The criterion from Proposition 2.36 for merge tree equivalence is sufficient but not necessary, as the following example shows:

(X, f) (X, g) $M(X, f) \cong M(X, g)$



The two dMfs f and g induce inverse orders on the two 0-simplices. Nonetheless, f and g induce the same unlabeled merge tree.

This remark leads us to yet another kind of equivalence relation between dMfs that arises from symmetries of sublevel complexes. In order to make this notion of symmetry precise, we need some preparations.

Definition 2.42 Let $f : X \to \mathbb{R}$ be a dMf on a tree. For each non-empty connected component $X_c^f[v]$ of a sublevel complex X_c^f we denote by $\mathrm{Aut}(X_c^f[v])$ the group of simplicial automorphisms of $X_c^f[v]$. For each $a \in \mathrm{Aut}(X_c^f[v])$ there is an extension to a self-bijection $X \to X$ by the identity. The group $\widetilde{\mathrm{Aut}}(X_c^f[v])$ is defined to be the group of said extensions of elements of $\mathrm{Aut}(X_c^f[v])$ by the identity. We consider $\widetilde{\mathrm{Aut}}(X_c^f[v])$ as a subgroup of the group of all self-bijections of X. The total order on $\mathrm{Cr}(f)$ induced by f induces chains $\widetilde{\mathrm{Aut}}(X_{c_0}^f[v]) \subset \widetilde{\mathrm{Aut}}(X_{c_1}^f[v]) \subset \ldots$ of inclusions of subgroups. Moreover, we have inclusions $\widetilde{\mathrm{Aut}}(X_{c_i}^f[v]) \subset \widetilde{\mathrm{Aut}}(X_{c_j}^f[v]) = \widetilde{\mathrm{Aut}}(X_{c_j}^f[w]) \supset \widetilde{\mathrm{Aut}}(X_{c_i}^f[w])$ if v and w are in different connected components of some sublevel complex $X_{c_i}^f$ that merge together in some other sublevel complex $X_{c_j}^f$ for $j > i$. We define the *sublevel automorphism group* of (X, f), denoted by $\mathrm{Aut}_{sl}(X, f)$, to be the subgroup generated by $\bigcup_{c \in \mathrm{Cr}(f), v \in X} \widetilde{\mathrm{Aut}}(X_c^f[v])$. We call the elements of $\mathrm{Aut}_{sl}(X, f)$ *sublevel automorphisms*.

Remark 2.43 Even though sublevel automorphisms are built out of simplicial automorphisms of connected components $X_c^f[v]$ of sublevel complexes, they are in general not simplicial maps $X \to X$. To be precise, if a simplicial automorphism of $X_c^f[v]$ was used to construct a sublevel automorphism $a \in \mathrm{Aut}_{sl}(X, f)$, then a will fail to be simplicial at the boundary of $X_c^f[v] \subset X$.

Proposition 2.44 *Let $f: X \to \mathbb{R}$ be a dMf on a tree and let $a \in \mathrm{Aut}_{sl}(X, f)$ be a sublevel automorphism. Then $f * a$ defined by $f * a(\sigma) := f(a(\sigma))$ is a dMf on X. Moreover, this defines a right group action of $\mathrm{Aut}_{sl}(X, f)$ on the set of dMfs on X.*

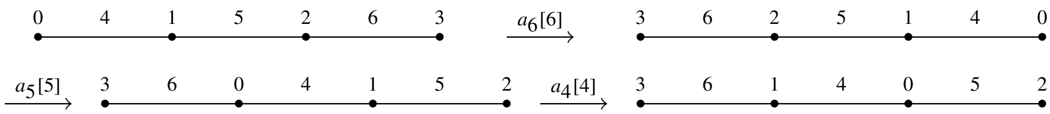
Sketch of Proof The proof that $f * a$ is a dMf is straightforward and the compatibility of the group action follows directly by associativity of the composition of maps. □

Remark 2.45 Since automorphisms of simplicial complexes preserve the dimension of simplices, the action of $\mathrm{Aut}_{sl}(X, f)$ on the set of dMfs on a tree X preserves the properties of being index-ordered or sublevel-connected.

Definition 2.46 Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be dMfs on a tree X. We call f and g *sublevel-equivalent* if $\mathrm{Cr}(f) = \mathrm{Cr}(g)$ and $X_c^f \cong X_c^g$ for all $c \in \mathrm{Cr}(f) = \mathrm{Cr}(g)$. If additionally $g = f * a$ holds for a sublevel automorphism $a \in \mathrm{Aut}_{sl}(X, f) = \mathrm{Aut}_{sl}(X, g)$, then we call f and g *symmetry-equivalent*. We call the map a a *symmetry equivalence* from f to g.

We call two dMfs $f: X \to \mathbb{R}$ and $g: Y \to \mathbb{R}$ *symmetry-equivalent* if there is a simplicial isomorphism $\varphi: X \to Y$ such that f and $g \circ \varphi$ are symmetry-equivalent.

Example 2.47 We give a list of some symmetry-equivalent dMfs on a path with four vertices. Here we denote the sublevel equivalence induced by the reflection of the connected component of the sublevel complex of level k by $a_k[k]$:



Proposition 2.48 *Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be symmetry-equivalent dMfs on a tree X. Then $M(X, f)$ and $M(X, g)$ are isomorphic as Ml trees.*

Sketch of Proof The proposition can be proved by induction over the level c of the sublevel automorphisms that the given symmetry equivalence consists of. We check that in each step the single sublevel automorphism a of the connected component $X_c[\sigma]$ with the 1-simplex σ labeled c in X_c only affects steps of the construction of the induced Ml tree that consider simplices of $X_c[\sigma]$. Moreover, it is straightforward to prove that in these steps, the created nodes and their induced Morse labels are the same as without the application of a. □

Remark 2.49 Sublevel automorphisms of dMfs on paths only consist of reflections of the corresponding connected component of a subcomplex. When such a connected component of a sublevel complex of a critical level c is considered during the construction of the induced Ml tree, the reflection of the connected component only causes the two new parts that are obtained by considering a slightly lower level $c-\varepsilon$ to appear as their mirror images. In particular, the given dMf attains the same values on the two new parts as before. Hence, the two parts that appear are simplicially isomorphic to the ones that appear without application of the reflection.

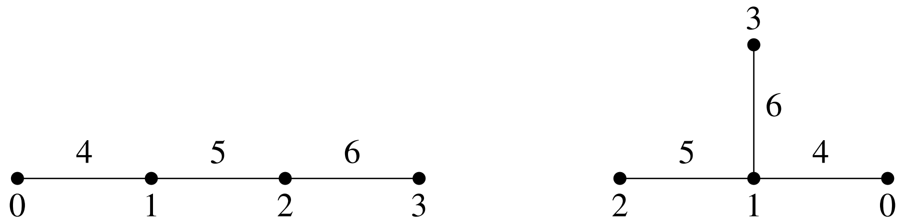
Definition 2.50 Let (X, f) and (X', f') be dMfs on trees. A *component-merge equivalence (cm equivalence) of level a* is a bijection $\varphi\colon X \to X'$ such that one of the following, not necessarily exclusive, cases holds:

(1) φ is a symmetry equivalence.
(2) φ fulfills the following:

- $f' \circ \varphi = f$,
- φ induces a bijection between the sets of connected components of sublevel complexes such that each restriction $\varphi_{|X_{a-\varepsilon}[v]}\colon X_{a-\varepsilon}[v] \to X'_{a-\varepsilon}[\varphi(v)]$ is a cm equivalence of some level $b < a$, and
- The edge $\sigma \in X$ with $f(\sigma) = a$ merges the two connected components $X_{a-\varepsilon}[v_1]$ and $X_{a-\varepsilon}[v_2]$ in $X_a[v_1] = X_a[v_2]$ if and only if the edge $\varphi(\sigma)$ merges the two connected components $X'_{a-\varepsilon}[\varphi(v_1)]$ and $X'_{a-\varepsilon}[\varphi(v_2)]$ in $X'_a[\varphi(v_1)] = X'_a[\varphi(v_2)]$.

If φ fulfills property (2) but not property (1), we call φ *non-trivial*.

Example 2.51 We give an example of two cm-equivalent dMfs on trees. The non-trivial cm equivalence from the left-hand-side to the right-hand-side consists of a symmetry equivalence of level 5 and the attachment of the edge labeled 6 between the vertices labeled 1 and 3 rather than 2 and 3. That is, it is a cm-equivalence of level 6.



Proposition 2.52 *Cm equivalent dMfs on trees induce isomorphic Ml trees.*

Proof Let $\varphi\colon (X, f) \to (X', f')$ be a cm equivalence. By property (ii) of Definition 2.1, at most one non-trivial cm equivalence of level a can occur for any level a because there is at most one edge labeled a in (X, f), (X', f'), respectively. Thus, we can decompose any cm equivalence into a sequence $(\varphi_a)_a$ of non-trivial cm equivalences of decreasing levels such that each φ_a only changes the attachment of the single edge σ with $f(\sigma) = a$ and acts as a symmetry equivalence on the rest of path and dMf. It suffices to consider a single level a because the statement then follows by induction from highest to lowest over all levels a.

For such a non-trivial cm equivalence φ_a we consider the step of the construction of the induced Ml trees that considers the edge σ with $f(\sigma) = a$ and the edge $\varphi(\sigma)$. We inductively assume that φ induces an isomorphism of induced Ml trees everywhere outside the subtrees corresponding to the two connected components of $X^f_{a-\varepsilon}$ that are merged by the edge σ with $f(\sigma) = a$. That is, on the rest of $M(X, f)$ the map

$M(\varphi)$ is a bijection compatible with the chiral child relation onto $M(X', f')$ except for the subtrees of $M(X', f')$ which correspond to the connected components of $X'^{f'}_{a-\varepsilon}$ which are merged by the edge $\varphi(\sigma)$. Since the map φ is compatible with the dMfs and because it restricts to a cm equivalence $X^f_{a-\varepsilon} \to X'^{f'}_{a-\varepsilon}$, the dMf f attains the same minima and maxima on the two relevant connected components of $X^f_{a-\varepsilon}$ as f' does on their counterparts of $X'^{f'}_{a-\varepsilon}$ via φ. Since Construction 2.14 only considers which two connected components are merged by the considered edge, it makes no difference for the isomorphism type of the induced Ml trees that σ in general merges the two connected components of $X^f_{a-\varepsilon}$ at vertices that do not correspond via φ to the ones adjacent to $\varphi(\sigma)$ in $X'^{f'}_{a-\varepsilon}$. Thus, the construction of the induced Ml tree produces nodes with the same chirality and label for both induced Ml trees in the steps that consider σ, $\varphi(\sigma)$, respectively. By assumption, the restriction $\varphi_{X^f_{a-\varepsilon}} : X^f_{a-\varepsilon} \to X'^{f'}_{a-\varepsilon}$ is a symmetry equivalence, so the isomorphism of Ml trees extends to the subtrees that correspond to the respective connected components. □

Proposition 2.53 *Let (X, f) be a dMf on a tree. There is a dMf on a path (P, f') such that (X, f) is cm-equivalent to (P, f').*

Sketch of Proof A suitable cm equivalence can be constructed inductively by reattaching 1-simplices of level a that would become the third 1-simplex incident to some 0-simplex in X_a. □

Remark 2.54 The way we defined cm equivalence makes it a generalization of symmetry equivalence. In fact, cm equivalences are the same as symmetry equivalences, i.e. they are always trivial, if we restrict ourselves to dMfs on paths: Without loss of generality, cm equivalences of some level a of a dMf on a path (P, f) describe all different possibilities of how two glue two paths together with a new edge in order to obtain a path again. This means that the edge labeled a can only be adjacent to the vertices that are adjacent to less than two edges, respectively, of the two old paths in $P_{a-\varepsilon}$. Thus, there are at most four possibilities for the two vertices which may be adjacent to the edge labeled a. All of these possibilities result in dMfs which are related to each other by reflections of the original two paths in $P_{a-\varepsilon}$. Hence, they are all symmetry-equivalent to each other.

3 Construction of the induced index-ordered DMF

We address the inverse question: For any given merge tree T, is there a discrete Morse function f on a path P such that $M(P, f) \cong T$? We answer this question affirmatively by presenting an explicit construction of P and two possible choices for f. The basic idea for the construction is to reverse-engineer the construction of the induced merge tree from Construction 2.14.

To start with the index-ordered case, we define two different orders on T. First we define a Morse order on T, which we call the *index Morse order*. Afterwards, we define the simplex order on the nodes of T, which we use to turn the Morse labeling induced by the index Morse order into a dMf on P. In Sect. 4 we will present an alternative dMf which represents T, namely the sublevel-connected dMf.

We will discuss in Sect. 5 to what extent the constructed dMf is a unique representative for T.

3.1 The index Morse order

To define the index Morse order, we first observe that every node a of T is uniquely determined by the shortest path from the root to a. We recall that the depth of T is the maximal length of any path in T that appears as the shortest path from the root to a leaf. Because T is chiral, we can identify such shortest paths with certain words:

Definition 3.1 Let T be a merge tree of depth n and let a be a node of T. The *path word* corresponding to a is a word $a_0a_1 \dots a_n \in \{L, R, \_\}^{n+1}$ where _ denotes the empty letter. If a is of depth k, the letters $a_0 \dots a_k$ are given by the chirality of the nodes belonging to the shortest path from the root to a. The letters $a_{k+1} \dots a_n$ are then empty.

Remark 3.2 Let a, b be nodes of a merge tree T and let $a_0a_1 \dots a_n$ be the path word corresponding to a and $b_0b_1 \dots b_n$ be the path word corresponding to b. Then the equation $a_0 = b_0 = L$ always holds because we consider paths that begin at the root. Because of $a_0 = b_0 = L$ and because we consider finite trees, there is always a maximal $k \in \mathbb{N}$ such that $a_i = b_i$ holds for all $i \leq k$. Furthermore, the last non-empty letter of a path word is always the chirality of the considered node.

We now define the index Morse order, which will produce an index-ordered dMf on P afterwards.

Definition 3.3 Let T be a merge tree. We define the *index Morse order* $\leq_{io}$ on the nodes of T as follows:

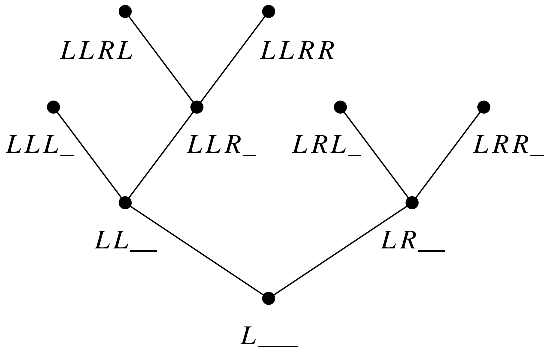
Let a and b be arbitrary nodes of T. If a is a leaf node and b is an inner node, then we define $a \leq_{io} b$. If either both a and b are leaf nodes or both a and b are inner nodes, we consider the following:

Let $a_0a_1 \dots a_n$ be the path word corresponding to a and $b_0b_1 \dots b_n$ the path word corresponding to b. Furthermore, let $k \in \mathbb{N}$ be maximal such that $a_i = b_i$ for all $i \leq k$. If $a_k = b_k = L/R$ we define $a \leq_{io} b$ if and only if one of the following cases hold:

(a) $a_{k+1} = L$ and $b_{k+1} = R$/$a_{k+1} = R$ and $b_{k+1} = L$
(b) $b_{k+1} = \_$
(c) $a = b$ ($\Leftrightarrow k = n$)

The index Morse order is tailor-made to induce an index-ordered dMf later on. Nonetheless, we will see in Example 5.2 that it is in general not the only Morse order which induces an index-ordered dMf. In Sect. 4 we will introduce a different, perhaps more natural, Morse order that is more closely related to the sublevel filtration of the induced dMf. But for now we consider an example of the index Morse order and prove that $\leq_{io}$ is actually is a Morse order.

Example 3.4 We consider the following merge tree T:



The path words are written underneath their corresponding nodes. The index Morse order produces the following chain of inequalities where we denote the nodes by their corresponding path words:

$$LLL\_ \lneq LLRR \lneq LLRL \lneq LRR\_ \lneq LRL\_ \lneq LLR\_ \lneq LL\_\_ \lneq LR\_\_ \lneq L\_\_\_$$

The inequalities from $LLL\_$ to $LRL\_$ arise from the path words of the leaf nodes. The inequality $LRL\_ \lneq LLR\_$ holds because the node corresponding to $LRL\_$ is a leaf node and the node corresponding to $LLR\_$ is an inner node. The inequalities from $LLR\_$ to $L\_\_\_$ arise from the path words of the inner nodes.

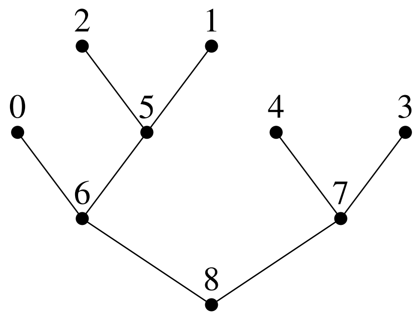
Remark 3.5 By definition, the root node is always the maximal element of $(V(T), \leq_{io})$. Furthermore, the leftmost leaf node of T is always the minimal element of $(V(T), \leq_{io})$.

Proposition 3.6 *The index Morse order is a Morse order on T.*

Sketch of Proof The proof is a straightforward application of the definitions and involves case distinctions corresponding to the cases a), b), and c) from Definition 3.3. □

Definition 3.7 We call the Morse labeling $\lambda_{io}\colon (V(T), \leq_{io}) \to \{0, 1, \ldots, i(T) + l(T) - 1\}$ induced by the index Morse order, see Definition 2.19, the *index Morse labeling* on T. That is, a node c of T is labeled with $\lambda_{io}(c)$.

Example 3.8 The index Morse order from Example 3.4 induces the following index Morse labeling:



3.2 The simplex order

We now define the simplex order on the nodes of T. The simplex order will tell us which nodes of T correspond to which simplices of P.

Remark 3.9 Let T be a merge tree and let a, b be nodes of T. Because T is in particular a rooted binary tree, there is a unique node p which is a common ancestor of a and b and has no descendants which are common ancestors of a and b.

Definition 3.10 We call the node p from Remark 3.9 the *youngest common ancestor of a and b*.

Definition 3.11 Let T be a merge tree. We define the *simplex order* $\preceq$ on $V(T)$ as follows: For two nodes a and b of T we define $a \preceq b$ if and only if one of following mutually exclusive cases holds, where p denotes the youngest common ancestor of a and b:

(1) a is a node of the subtree with root p_l and b is a node of the subtree with root p_r.
(2) a is a node of the subtree with root b_l (in particular $b = p$).
(3) b is a node of the subtree with root a_r (in particular $a = p$).
(4) $a = b$.

Proposition 3.12 *The simplex order is a total order on the nodes of T.*

Sketch of Proof of Proposition 3.12 The proof is a bit tedious and consists of many careful case distinctions corresponding to the different cases from Definition 3.11. Otherwise, the proof is a straightforward application of the definitions, paired with a contradiction argument here and there. □

We use the following definition to make the intuition of leaves being adjacent precise. This allows us to analyze the simplex order further.

Definition 3.13 Let T be a merge tree and let a and b be leaves of T. We call a and b *adjacent* if one of the following holds:

(1) $a \preceq b$ and there is no leaf node c of T such that $a \preceq c \preceq b$ holds.
(2) $b \preceq a$ and there is no leaf node c of T such that $b \preceq c \preceq a$ holds.

Lemma 3.14 *Subtrees of T form chains of cover relations in $(V(T), \preceq)$. In detail, this means the following:*

Let p be an inner node of T and let a/b be the leftmost/rightmost leaf of the subtree with root p. Then the nodes of the subtree with root p in T form the chain of cover relations $a \prec \cdots \prec p \prec \cdots \prec b$ in $(V(T), \preceq)$. Moreover, any two adjacent leaves of T have the property that the left one of the two adjacent leaves is covered by the youngest common ancestor of the two, whereas the right one covers the youngest common ancestor.

Proof We prove the lemma inductively. Let p be a node of T such that both child nodes of p are leaves. It follows directly by Definition 3.11 that p_r covers p and p

covers p_l. That is, the subtree with root p forms the chain of cover relations $a = p_l \prec p \prec p_r = b$.

If p is an arbitrary inner node, then by the inductive hypothesis the subtree with root p_l/p_r forms the chain of cover relations $a_1 \prec \cdots \prec p_l \prec \cdots \prec b_1/a_2 \prec \cdots \prec p_r \prec \cdots \prec b_2$ where a_1/a_2 is the leftmost and b_1/b_2 the rightmost leaf of the subtree with root p_l/p_r. Since b_1/a_2 is a node of the subtree with root p_l/p_r, it follows by case (2)/(3) of Definition 3.11 that $b_1 \prec p/p \prec a_2$ holds. For all nodes c of T which are not nodes of the subtree with root p, the same case from Definition 3.11 holds for c and p as for c and b_1/a_2. Thus, and because b_1/a_2 is *maximal* / *minimal* in the subtree with root p_l/p_r by the inductive assumption, there is no node c such that $b_1 \prec c \prec p/p \prec c \prec a_2$ holds. In conclusion, p covers b_1/a_2 *covers* p. □

As mentioned before, we will use the simplex order to relate the nodes of T to the simplices of a path P. In order to do that, we now define a corresponding simplex order on the simplices of P.

Definition 3.15 Let P be a path. There are two 0-simplices p_0 and p_1 in P which belong only to one respective 1-simplex. For each simplex σ of P there is a unique shortest path γ_σ from p_0 to σ. We denote the length, that is, the number of simplices, of such a path γ_σ by $L(\gamma_\sigma)$. The *simplex order* on P is defined as follows: For two simplices σ and τ of P we define $\sigma \preceq \tau$ if and only if $L(\gamma_\sigma) \leq L(\gamma_\tau)$.

Lemma 3.16 *The simplex order on P is a total order on the simplices of P.*

Sketch of Proof The proof is straightforward and only uses that P is a path and that the integers are linearly ordered. □

Remark 3.17 Connected subcomplexes of P correspond to chains of cover relations with respect to the simplex order. Furthermore, any 1-simplex covers its left boundary 0-simplex and is covered by its right boundary 0-simplex. If one visualizes P as being horizontally embedded in a plane such that p_0 is the leftmost 0-simplex of P and p_1 is the rightmost 0-simplex of P, then for simplices $s, s' \in P$ the relation $s \prec s'$ holds if and only if s is left of s'. This reminds us of the fact that the simplex order is only defined up to a choice of orientation.

Remark 3.18 Let T be a merge tree and let P be a path with $i(T)$ 1-simplices. Then we have a unique isomorphism $\phi\colon (P, \preceq) \xrightarrow{\cong} (V(T), \preceq)$ of totally ordered sets. The isomorphism ϕ only depends on the choice of p_0 and p_1, that is, on a choice of orientation on P. Choosing p_0 and p_1 the other way around would reverse the simplex order on P.

Before we continue with the definition of the index-ordered dMf, we consider how the simplex order can be used to classify the connected components of sublevel complexes of dMfs on paths (P, f):

Proposition 3.19 *Let $f\colon P \to \mathbb{N}_0$ be a dMf on a path P. Then the connected components $P_c[v]$ of sublevel complexes P_c of P are precisely maximal sequences $\sigma := (s_0, \ldots, v, \ldots, s_k)$ of simplices of P such that $s_i \in P_c$ for all $i = 0, \ldots, k$ and $s_0 \prec \cdots \prec v \prec \cdots \prec s_k$ is a chain of cover relations in $(P, \preceq)$.*

Proof The proof is straightforward and uses Remark 3.17 □

3.3 The induced index-ordered DMF

Now we explain how the simplex order can be used to construct dMfs on P from Morse orders on T:

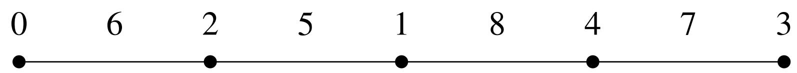
Definition 3.20 Let T be a merge tree and P be a path such that the number of 1-simplices is $i(T)$ and let $\phi\colon (P, \preceq) \to (V(T), \preceq)$ be the isomorphism from Remark 3.18.

For a Morse order $\leq$ and its induced Morse labeling λ we define a map $f_\lambda\colon P \to \mathbb{N}_0$ by $f_\lambda := \lambda \circ \phi$. The map f_λ is then called the *dMf induced by the Morse order* $\leq$ or the *dMf induced by the Morse labeling* λ.

In particular, the map $f_{io} := \lambda_{io} \circ \phi$ induced by the index Morse order is called the *induced index-ordered dMf*.

Remark 3.21 Although ϕ and λ are order-preserving maps with respect to the previously defined total orders, the map f_{io} does not respect the simplex order in general. Since f_{io} is supposed to be an index-ordered dMf, it does not need to respect the simplex order. Because the map f_{io} is supposed to be index-ordered, it rather needs to be compatible with face relation on P, which we will see to be true later on.

Example 3.22 The index Morse order from Example 3.4, respectively the index Morse labeling from Example 3.8, produces the following pair (P, f_{io}):



Proposition 3.23 *For any given Morse order $\leq$ on any merge tree T the dMf induced by $\leq$ is a dMf that has only critical cells.*

Sketch of Proof The proof is straightforward and uses Lemma 3.14, Remark 3.17, and property (1) of Definition 2.17. □

Remark 3.24 The previous proposition proves that Morse orders $\leq$ on merge trees T always induce dMfs f_λ. It is a priori unclear though whether the induced dMf f_λ induces the given merge tree T as its induced merge tree $M(P, f_\lambda)$. We prove this to be true in Theorem 5.5.

Furthermore, condition (1) from Definition 2.17 is necessary for f_λ to be a dMf, because a violation of (1) between an inner node and a leaf would result in a violation of f_λ being weakly increasing on the corresponding simplices.

Before we continue with the sublevel-connected dMf, we consider how the simplex order can be used to improve our understanding of sublevel complexes of dMfs on paths (P, f_λ) and how they are related to subtrees of T. The condition for this approach to be applicable is that the dMf f_λ is induced by a Morse order $\leq$ on T as in Definition 3.20, which we will see to be the general case later on. We will apply this approach to the sublevel-connected case in Sect. 4 where it will be of more importance.

Proposition 3.25 *Let $f_\lambda\colon P \to \mathbb{N}_0$ be a dMf on a path P that is induced by a Morse order $\leq$ on T. Then the connected components $P_c[v]$ of sublevel complexes P_c of (P, f) induce subtrees of T via ϕ.*

Sketch of Proof It follows by Proposition 3.19 and the definition of ϕ in Remark 3.18 that connected components of sublevel complexes $P_c[v]$ induce maximal chains of cover relations such that the corresponding simplices are of at most level c in $(V(T), \preceq)$.

The next step is to prove that such chains are equal to the subtree with the chain's maximum as root. The proof that the chain is contained in the subtree is straightforward. The other inclusion can be proved by contradiction, using that a node outside the subtree would contradict the property of being a chain of cover relations. □

4 The sublevel-connected DMF

As remarked in Sect. 2.2 it might sometimes be more convenient to work with sublevel-connected dMfs rather than with index-ordered dMfs. In this section we introduce a slightly different version of the Morse order from Definition 3.3 to construct a sublevel-connected dMf which is shuffle-equivalent to the induced index-ordered dMf and, hence, induces the same given merge tree.

Definition 4.1 Let T be a merge tree. We define the *sublevel-connected Morse order* $\leq_{sc}$ on the nodes of T as follows:

Let a, b be arbitrary nodes of T. Let $a_0 a_1 \dots a_n$ be the path word corresponding to a and $b_0 b_1 \dots b_n$ the path word corresponding to b (see Definition 3.1). Furthermore, let $k \in \mathbb{N}$ be maximal such that $a_i = b_i$ for all $i \leq k$. If $a_k = b_k = L/R$ we define $a \leq_{sc} b$ if and only if one of the following cases hold:

(a) $a_{k+1} = L$ and $b_{k+1} = R/a_{k+1} = R$ and $b_{k+1} = L$
(b) $b_{k+1} = \_$
(c) $a = b$

Remark 4.2 The only difference between the definition of the index Morse order and the definition of the sublevel-connected Morse order is that we do not treat leaves and inner nodes differently anymore. Thus, both orders induce the same order on inner nodes and the same order on leaves, which makes the index Morse order and the sublevel-connected Morse order shuffle-equivalent. However, the order relation between a leaf and an inner node is in general different then in the index Morse order.

Remark 4.3 The fact that the sublevel-connected Morse order is a Morse order can be proved the same way as the corresponding statement Proposition 3.6 for the index Morse order was proved. There are just fewer case distinctions to be made for the sublevel-connected Morse order.

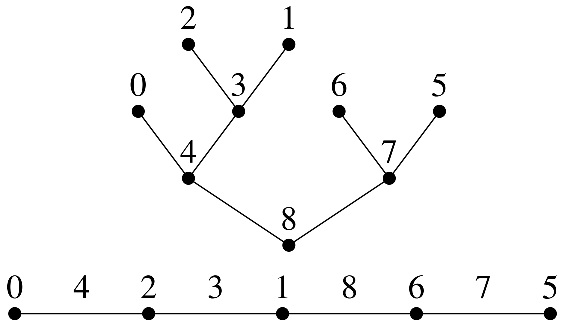
Lemma 4.4 *Subtrees of T form intervals in $(V(T), \leq_{sc})$.*

Sketch of Proof The proof is straightforward and uses the fact that all path words of a subtree T' start with the same couple of letters corresponding to the root of T'. □

The induced sublevel-connected labeling λ_{sc} and the induced sublevel-connected dMf f_{sc} are defined analogously to the index-ordered case:

Definition 4.5 Let T be a merge tree and P be a path such that the number of 1-simplices is $i(T)$. We call the Morse labeling $\lambda_{sc} \colon (V(T), \leq_{sc}) \to \{0, 1, \ldots, i(T) + l(T) - 1\}$ induced by the sublevel-connected Morse order the *sublevel-connected Morse labeling* on T. The dMf $f_{sc} = \lambda_{sc} \circ \phi$ induced by λ_{sc} is called the *induced sublevel-connected dMf*.

Example 4.6 Let T be the merge tree from Example 3.4. The sublevel-connected Morse labeling on T and the induced sublevel-connected dMf are given below.



There are now two things left to prove: that the map f_{sc} is indeed a sublevel-connected dMf and that it induces the given merge tree T.

Proposition 4.7 *The induced sublevel-connected dMf f_{sc} is a sublevel-connected dMf that has only critical cells.*

Proof As a dMf induced by a Morse order, the map f_{sc} is by Proposition 3.23 a dMf that has only critical cells. It is left to prove that f_{sc} is sublevel-connected:

By Proposition 3.25, the connected components of sublevel complexes of (P, f) induce subtrees of T via ϕ. By Lemma 4.4, subtrees of T form intervals in $(T, \leq_{sc})$. The sublevel-connected Morse labeling λ_{sc} by definition maps intervals of $(T, \leq_{sc})$ to intervals of $\mathbb{N}_0$. By concatenation of these arguments, it follows that $f_{sc} = \lambda_{sc} \circ \phi$ maps connected components of sublevel complexes to intervals of $\mathbb{N}_0$, i.e. the map f_{sc} is sublevel-connected. □

Theorem 4.8 *Let T be a merge tree and let P be a path such that the number of 1-simplices in P is $i(T)$. Then f_{io} and f_{sc} are shuffle-equivalent where f_{sc} is the induced sublevel-connected dMf and f_{io} is the induced index-ordered dMf. Thus, $M(P, f_{sc}) \cong M(P, f_{io})$ holds as merge trees.*

Proof By Remark 4.2, the index Morse order and the sublevel-connected Morse order are shuffle equivalent. It follows by Corollary 2.27 and Definition 3.20 the the index-ordered dMf and the sublevel-connected dMf are shuffle equivalent. Thus, $M(P, f_{sc}) \cong M(P, f_{io})$ follows by Proposition 2.36, where f_{io} is the induced index-ordered dMf defined in Definition 3.20. □

5 Relationships between merge trees and DMFs on paths

In this section we want to take a look at the bigger picture again and consider how we can relate dMfs on paths and trees to merge trees. In order to do this in a structured way, we consider the different sets of dMfs and merge trees and relate them to each other by bijections that are compatible with the various notions of equivalence we introduced earlier. Afterwards, we use said bijections in order to prove that the aforementioned dMfs f_{io}, Definition 3.20, and f_{sc}, Proposition 4.7, both represent the given merge tree. We also relate the sets of dMfs and merge trees mentioned here to the setting of Curry (2019).

Remark 5.1 In order to obtain bijections that are compatible with the various notions of equivalence, we restrict ourselves to the case of dMfs for which all simplices are critical. Since the induced merge tree does not take matched cells into account, we would otherwise need to define a notion of equivalence similar to Forman equivalence (see Johnson and Scoville 2022, Def 4.1) of dMfs that in particular takes simple homotopy equivalences as well as combinatorial aspects of dMfs into account. This could be done by additional pre- and post-composition of the equivalences as we defined them with simple homotopy equivalences that are compatible with the given dMfs. For simplicity, we chose to leave this aspect out of this work.

We denote the set of merge trees up to isomorphism by Mer and the set of dMfs on paths with only critical simplices up to symmetry equivalence by DMF_P^{crit}. It follows by Proposition 2.48 that the assignment $M(\_, \_)$ is well-defined on DMF_P^{crit}. Furthermore, the construction of the induced dMf from Definition 3.20 extends to a map which we denote by Φ. Since ϕ is well defined up to a choice of orientation, the map Φ is in particular well defined up to symmetry equivalence. This leaves us with the diagram in Fig. 1.

The arrows induced by the index Morse order and the sublevel-connected Morse order are not left-inverse to the forget arrow as the following example shows. We have marked them in red because they are the only arrows that prevent the diagram from commuting completely. Nonetheless, it is obvious that the arrows induced by the two Morse orders are right-inverses of the forgetful arrow.

Example 5.2 Consider the following Ml trees and their corresponding Mo trees:

It is immediate the (T, λ) and (T', λ') are neither isomorphic nor shuffle equiva-

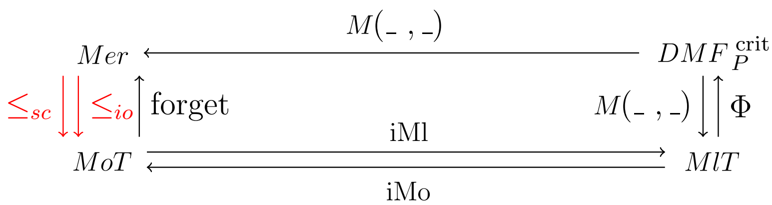


Fig. 1 Relationships Between Merge Trees and DMFs on Paths

lent. Thus, $(T', \lambda') \ncong (T, \lambda) = (\text{forget}(T', \lambda'), \lambda_{io})$ holds. Nonetheless, the Ml tree (T', λ') induces an index-ordered dMf.

But we clearly have the following.

Remark 5.3 Let T, T' be isomorphic merge trees. Then $(T, \leq_{io}) \cong (T', \leq_{io})$ and $(T, \leq_{sc}) \cong (T', \leq_{sc})$ holds as Mo trees. Furthermore, we have $\text{forget}(T, \leq_{io}) \cong T \cong \text{forget}(T, \leq_{sc})$ as merge trees.

We have already seen in Proposition 2.26 that the maps iMo and iMl are inverse to each other in the sense that they are bijections compatible with isomorphisms, order equivalences and shuffle-equivalences. We now show that $M(\_, \_)$ and Φ are also inverse to each other up to the respective notions of equivalence.

Theorem 5.4 *The induced labeled merge tree $M(\_, \_)$ and the induced dMf Φ define maps $M(\_, \_)\colon DMF_P^{crit} \leftrightarrow MlT\colon \Phi$ that are inverse to each other in the following sense:*

(1) *For any dMf (P, f) with only critical cells, the dMf $\Phi(M(P, f), \lambda_f)$ is symmetry-equivalent to (P, f), and*
(2) *For any Ml tree (T, λ), the Ml tree $M(\Phi T, f_\lambda)$ is isomorphic to (T, λ).*

Proof (1) Let (P, f) be a dMf on a path. We construct a symmetry equivalence between f and f_{λ_f}. It is given as follows: For any simplex σ of $\Phi M(P, f)$ there is exactly one simplex $\tilde{\sigma}$ of P such that $f(\tilde{\sigma}) = f_{\lambda_f}(\sigma)$. This induces a bijection $\varphi\colon P \to \Phi M(P, f)$ which is compatible with f, f_{λ_f} and $\text{id}_{\mathbb{R}}$ by definition. But in general, the map φ is not simplicial. This is because the simplex order on $M(P, f)$ might be different than the left/right relation on the corresponding simplices of P. In other words, the map $M(\_)\colon P \to M(P, f)$ is in general not compatible with the two different simplex orders. Nonetheless, the Morse labeling λ_f induced by f orders the nodes of $M(P, f)$ in the same order as their corresponding simplices of P. Thus, connected components of sublevel complexes of (P, f) still correspond to subtrees of $M(P, f)$. Since the induced merge tree assigns the chirality of child nodes according to which connected component carries the minimal value of f, at each inner node of $M(P, f)$ the chirality of the two child nodes is either assigned in accordance with the left/right relation on the corresponding sublevel complex, or it is the opposite. If it is the opposite, this can be corrected by application of the reflection of the corresponding sublevel complex, that is, by application of a sublevel equivalence. In consequence, the difference between the right/left relation of the corresponding simplices in P and the simplex order only lies in symmetry equivalences of P. Thus, φ can be decomposed into a symmetry equivalence of P and a simplicial isomorphism $\tilde{\varphi}$. Hence, (φ, ψ) is a symmetry equivalence of dMfs on paths.

(2) Let (T, λ) be an Ml tree. Let $c_0 < c_1 < \cdots < c_n$ be the critical values of f_λ and let $\sigma_i \in \Phi T$ such that $f_\lambda(\sigma_i) = c_i$. We recall that the induced merge tree M defines in particular a bijection between the critical simplices of ΦT and the nodes of $M(\Phi T, f_\lambda)$. For any simplex $\sigma \in \Phi T$, we denote the node of $M(\Phi T, f_\lambda)$ that corresponds to σ by $M(\sigma)$. An isomorphism $\varphi\colon (T, \lambda) \to M(\Phi T, f_\lambda)$ is given

by $\varphi := M \circ \phi^{-1}$. It is immediate that φ is a bijection because M and ϕ are. Furthermore, φ is by construction compatible with the respective Morse labelings. It is only left to show that φ is compatible with the chiral child relation and the respective roots.

Consider $\sigma_n \in \Phi T$. For both trees, the simplex σ_n corresponds to the root of the respective tree. In $M(\Phi T, f_\lambda)$ this is the case because σ_n carries the maximal value of f_λ. In (T, λ) this holds because $\phi(\sigma_n)$ holds the maximal Morse label $\lambda(\phi(\sigma_n)) = c_n$. Thus, the map φ maps the root of (T, λ) to the root of $M(\Phi T, f_\lambda)$.

Let σ_i be a simplex of ΦT. We now prove that φ is compatible with the chiral child relation, that is, that $\varphi(\phi(\sigma_i)_l) = M(\sigma_i)_l/\varphi(\phi(\sigma_i)_r) = M(\sigma_i)_r$ holds. If σ_i is a 0-simplex then there is nothing to show because then both $M(\sigma_i)$ and $\phi(\sigma_i)$ are leaves. Let σ_i be a critical edge with chirality L. The case for chirality R works symmetrically to the case with chirality L.

By Construction 2.14, the node $M(\sigma_i)_l/M(\sigma_i)_r$ corresponds to a critical simplex of the connected component of $\Phi T_{c_i-\varepsilon}$ that carries/does not carry the minimal value of f_λ on these two connected components. Furthermore, the node $M(\sigma_i)_l/M(\sigma_i)_r$ corresponds to the critical simplex that carries the maximal value of f_λ of the respective connected component of $\Phi T_{c_i-\varepsilon}$.

By application of Proposition 3.25, we see that the connected components $\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_l)]/\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_r)]$ induce subtrees of (T, λ) via ϕ. It follows that the node $\phi(\sigma_i)_l/\phi(\sigma_i)_r$ is contained in the subtree that corresponds to $\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_l)]/\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_r)]$ via ϕ because by (2) of Definition 2.17 the subtree with root $\phi(\sigma_i)_l$ does/$\phi(\sigma_i)_r$ does not carry the minimal Morse label of the subtree with root $\phi(\sigma_i)$ in (T, λ). Furthermore, the node $\phi(\sigma_i)_l/\phi(\sigma_i)_r$ corresponds to the simplex that carries the maximal value of f_λ of $\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_l)]/\Phi T_{c_i-\varepsilon}[M^{-1}(M(\sigma_i)_r)]$ because by (1) of Definition 2.17 it carries the maximal Morse label on said subtree and because ϕ is order-preserving. Thus, $\varphi(\phi(\sigma_i)_l) = M(\sigma_i)_l/\varphi(\phi(\sigma_i)_r) = M(\sigma_i)_r$ holds.

□

By considering Fig. 1 we see that the difference between taking the induced merge tree of a dMf on a path is the same as taking its induced Ml tree and forgetting the Morse labeling. Thus, constructing a dMf that represents a given merge tree T is up to symmetry equivalence the same as choosing a Morse order on T. This leads us to:

Theorem 5.5 *Let T be a merge tree and let P be a path such that the number of 1-simplices in P is $i(T)$. Then $T \cong M(P, f_{io}) \cong M(P, f_{sc})$ holds as merge trees where f_{io} denotes the induced index-ordered dMf* (Definition 3.20) *and f_{sc} denotes the sublevel-connected dMf* (Definition 4.5).

Proof The statement follows by Theorem 5.4, Proposition 2.26, and the fact that by Definition 2.23, isomorphisms of Mo trees are in particular isomorphisms of the underlying merge trees. Furthermore, $M(P, f_{io}) \cong M(P, f_{sc})$ holds by Theorem 4.8. □

Using the notion of component-merge equivalence, Definition 2.50, we can extend Theorem 5.4 to a bijection between the set of dMfs with only critical simplices on trees up to cm equivalence DMF_X^{crit} and the set of Ml trees up to isomorphism MlT:

Theorem 5.6 *The induced labeled merge tree $M(\_,\_)$ and the induced dMf Φ define maps $M(\_,\_)\colon DMF_X^{crit} \leftrightarrow MlT : \Phi$ that are inverse to each other in the sense that:*

(1) *For any dMf (X, f) with only critical cells, the dMf $\Phi(M(X, f), \lambda_f)$ is cm-equivalent to (X, f), and*
(2) *For any Ml tree (T, λ), the Ml tree $M(\Phi T, f_\lambda)$ is isomorphic to (T, λ).*

Proof The proof for statement (2) works exactly as in the proof for Theorem 5.4 because symmetry equivalences are in particular cm equivalences. For (1) we apply Proposition 2.53 to consider a representative of the cm equivalence class of (X, f) which is a dMf on a path (P, f'). By Proposition 2.52, the isomorphism type of the induced Ml tree does not depend on this choice. Thus, Theorem 5.4 implies that $\Phi(M(P, f'))$ is symmetry-equivalent to (P, f'). Since (P, f') is cm-equivalent to (X, f), so is $\Phi(M(P, f'))$. □

Corollary 5.7 *Let T be a merge tree. By Theorems 5.5 and 5.6 it follows that there are dMfs on trees (X, f) such that $M(X, f) \cong T$ as merge trees.*

Corollary 5.8 *Applying Theorem 5.6 together with Propositions 2.26 and 2.40 yields the result that there is a bijection $DMF_X^{crit}/_\sim \cong MoT$ where $\sim$ denotes order equivalence.*

We conclude this section by discussing our results and comparing them to the results of Curry (2019). Using the bijections appearing in Fig. 1 and Theorem 5.6 we replaced the question of finding dMfs on paths or arbitrary trees that represent a given merge tree T by finding Morse orders on T instead. This argument can be used to replace the question of classifying merge equivalence classes of dMfs on trees by classifying Morse orders. Example 5.2 tells us that, if a merge tree T has at least three leaves, there might be different Morse orders on T which are not shuffle-equivalent. That is, there are dMfs on trees (X, f) contained in the image of $\Phi \circ iMl$ which induce T as their merge tree but are neither isomorphic, nor symmetry-equivalent, nor cm-equivalent, nor shuffle-equivalent, nor a combination of the four to each other. We found a non-empty shuffle equivalence class of Morse orders on any merge tree, defined by either the index Morse order or the sublevel-connected Morse order, since the two have been shown to be shuffle-equivalent and, hence, merge-equivalent in Theorem 4.8. This allowed us to answer the aforementioned question of Johnson–Scoville affirmatively. That is, any merge tree is indeed represented by a dMf on a path, which is induced by a Morse order. Using a non-trivial cm equivalence, one can also find a dMf on a tree other than a path as a representative.

Furthermore, our results from Sects. 2.1, 2.2 and 5 allow us to structure the study of the set of merge equivalence classes of dMfs on trees using the following four notions of merge-invariant equivalences between dMfs on trees: Forman equivalence, symmetry equivalence, cm equivalence, and shuffle equivalence.

We discussed in Remark 5.1 how Forman equivalences could be considered together with the other notions of equivalence and why we left Forman equivalences out of this work. In Theorem 5.4 it becomes quite clear that passing to the induced Ml tree identifies symmetry-equivalent dMfs with each other up to isomorphism. Passing from the induced Ml tree to the induced Mo tree then identifies dMfs up to order equivalence with each other up to isomorphism of the underlying Mo tree. Last of all, passing from the induced Mo tree to the induced merge tree in particular identifies shuffle-equivalent dMfs with each other. This allows us to study the different equivalence classes separately for dMfs on paths. The more liberal notion of cm equivalence, Definition 2.50, allowed us to generalize Theorem 5.4 to dMfs on arbitrary trees as seen in Theorem 5.6.

In Curry (2019), the author establishes a bijection between graph-equivalence (Curry 2019, Def 6.1) classes of Morse-like (Curry 2019, Def 6.9) continuous functions on the interval that attain minima at the boundary on the one hand and isomorphism classes of chiral merge trees (Curry 2019, Def 5.3) on the other hand.

At first glance it might seem likely that Ml trees in the sense of Definition 2.19 and chiral merge trees in the sense of Curry (2019, Def 5.3) are directly related by geometric realization and considering the corresponding abstract simplicial complex but, as mentioned in the introduction, there is a subtle difference in the construction of the induced merge tree. To be precise, the two constructions only differ in the induced chirality. In Curry (2019, Sec 5) the chirality is given by which of the two merging components is the left or right one with respect to the chosen orientation on the interval. In Johnson and Scoville (2022) the chirality is given such that drawing the induced merge tree is compatible with the elder rule: The component with the minimal value gets the same chirality as the merged component. This means that following the same chirality leads to the oldest component.

This convention leads to the necessity to assume property (2) of Definition 2.17: a certain compatibility between the Morse order and the chirality. The compatibility between Morse orders and the chirality implies, as seen in Proposition 2.48, that the induced Ml tree does not distinguish between symmetry-equivalent dMfs. If one defines induced Ml trees analogously to Curry (2019), that is by inducing the chirality by a chosen orientation of the path, this notion of induced Ml trees would distinguish symmetry-equivalent dMfs on paths. Moreover, Ml trees induced by symmetry-equivalent dMfs would be related by sequences of reflections of subtrees. The definition could be as follows:

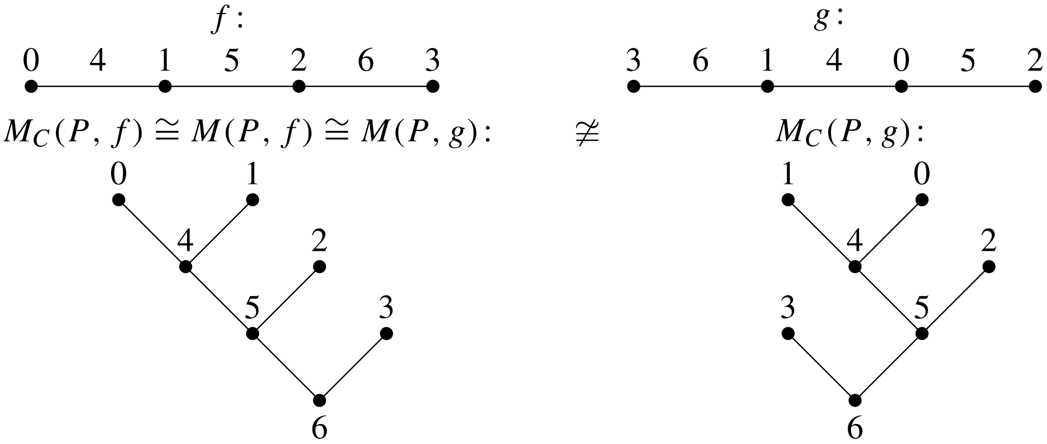
Definition 5.9 Let T be a merge tree. A *Curry Morse order* is a total order $\leq$ on the nodes of T such that the maximal node of any subtree is the root of said subtree. A *Curry Morse labeling* on a merge tree T is a labeling λ on the nodes of T that induces a Curry Morse order on T.

A pair $(T, \leq)$ of a merge tree with a Curry Morse order on it is called a *Curry Morse ordered merge tree (CMo tree)*. A pair (T, λ) of a merge tree with a Curry Morse labeling on it is called a *Curry Morse labeled merge tree (CMl tree)*.

Let $f : P \to \mathbb{R}$ be a dMf where P is an oriented path. Then the CMl tree induced by f is constructed as is Construction 2.14 with the difference that the chirality is assigned according to the position of the corresponding connected components with respect to the orientation instead of according to critical values.

Remark 5.10 CMl trees as defined above are basically the same concept as generic merge trees as defined in Curry et al. (2021, Def 2.2). We stick to the name CMl trees instead of generic merge trees because we already use the term merge tree in a different way.

Example 5.11 We consider two of the symmetry-equivalent dMfs from Example 2.47 and see how the induced CMl tree distinguishes them whereas the induced Ml trees identifies them as one:



The notion of CMl trees is related to the notion of chiral merge trees in the sense of Curry (2019) by the interplay between abstract and geometrical simplicial complexes. In detail, the bijection is given as follows:

Construction 5.12 Let (T, λ) be a CMl tree. We define a chiral merge tree $|(T, \lambda)|$ associated to (T, λ) as follows:

The compact rooted tree is given by the geometric realization $|T|$. We attach a distinguished edge e_∞ to the vertex which corresponds to the root of T in order to obtain a cell complex which we will by abuse of notation also refer to as $|T|$. The map $\pi : |T| \to \mathbb{R}$ is given by λ on vertices, and by a linear extension of λ on edges.

For the other way around let $\pi : T \to \mathbb{R}$ be a chiral merge tree in the sense of Curry (2019, Def 5.3). We define an Ml tree $\mathrm{abs}(T)$ associated to T as follows:

We take the 0-skeleton T_0 as the vertex set and the 1-skeleton $(T \setminus \{e_\infty\})_1$ as the set of edges. We define the node which corresponds to v_∞ to be the root of $\mathrm{abs}(T)$. The labeling λ is given by π.

The proof that the two constructions are inverse to each other is straightforward.

Furthermore, there is a similar bijection between the two notions of Morse functions:

Construction 5.13 Let (P, f) be a dMf with only critical cells on an oriented path. We define a Morse-like function $\tilde{f}$ on the interval which attains minima at the boundary as follows:

Let $k + 1$ be the number of 0-simplices in P. We denote the 0-simplices of P by $n_0, n_1, \ldots, n_k$ from left to right with respect to the given orientation. Then we define $\tilde{f}(\frac{i}{k}) := f(n_i)$ for $i \in \{0, 1, \ldots, k\}$. We denote the 1-simplices of P by $e_1, \ldots, e_k$, again according to the given orientation. Then we define $\tilde{f}(\frac{2i-1}{2k}) := f(e_i)$ for $i \in \{0, 1, \ldots, k\}$. We define $\tilde{f}$ on the rest of the interval as the linear extension. This makes $\tilde{f}$ a distinct-valued PL function that attains minima at the boundary. Hence, $\tilde{f}$ is Morse-like.

For the other way around let $f: I \to \mathbb{R}$ be a Morse-like function that attains minima at the boundary. We define a path P and a dMF f' on P as follows:

Let $n_0, n_1, \ldots, n_k$ be the local minima of f ordered by the orientation on I, so in particular $n_0 = 0$ and $n_k = 1$. We define P to be a path with $k + 1$ simplices of dimension 0. We choose one of the endpoints to be denoted by s_0 and the one by s_k. We denote the other 0-simplices such that their indices are in accordance with their position in the simplex order (Definition 3.15), making s_0 the minimal simplex with respect to the simplex order. The dMf f' is defined by $f'(s_i) := f(n_i)$ on 0-simplices. Let $c_1, \ldots, c_k$ be the local maxima of f, ordered in accordance to the orientation on I, and let $e_1, \ldots, e_k$ be the 1-simplices of P, ordered in accordance to the aforementioned simplex order on P. Then f' is defined by $f'(e_i) := f(c_i)$ on 1-simplices.

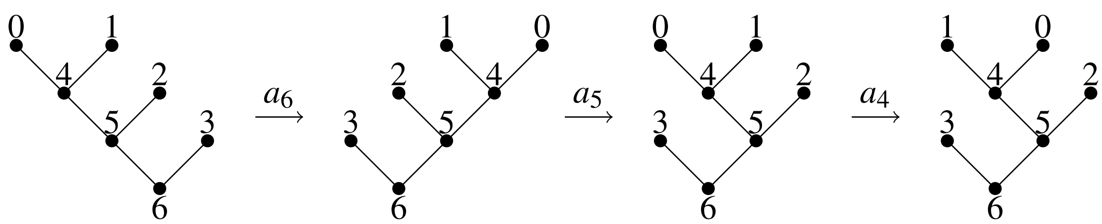
It is easy to check that the two given constructions are inverse to each other.

A theorem which, analogously to Theorem 5.4, defines a pair of inverse bijections $M_C(\_,\_): {}^+DMF_P^{\text{crit}} \leftrightarrow CMIT: \Phi$ can be proved similarly to the proof of Theorem 5.4. The difference is that the induced CMl tree keeps track of symmetry equivalences in the sense that symmetry equivalences of dMfs induce reflections at roots of subtrees on the induced CMl trees. The map Φ can be defined the same way as before. Alternatively, one could define the bijection Φ as the composition $\text{abs} \circ \Psi^{-1} \circ |\_|$ in the diagram in Fig. 2:

Here Ψ denotes the bijection between the set of Morse-like functions on the interval $\mathcal{M}$ and the set of chiral merge trees $\mathcal{X}$ from Curry (2019, Cor 6.11). The map i is the inclusion induced by considering Ml trees as CMl trees. Ml trees can be considered as a special kind of CMl trees because they only differ in the additional property (2) of Definition 2.17 which does not need to hold for CMl trees. The map $/_{symm}$ is the quotient map that identifies symmetry equivalent dMFs. The map o_{JS} is defined as the composition $\Phi \circ i \circ M(\_,\_)$. This definition coincides with choosing the representative of a dMf (P, f) with respect to symmetry equivalence such that the simplex order on P is compatible with f in the following way: At each critical edge e with $f(e) = c$ the connected component of $P_{c-\varepsilon}$ that corresponds to $M(e)_l/M(e)_r$ is left/right is left/right of the edge e with respect to the orientation induced by the simplex order. Here we recall that $M(e)$ denotes the inner node of $M(P, f)$ that corresponds to the critical edge e and that $M(e)_l/M(e)_r$ denotes the left/right denotes the left/right child node of $M(e)$. In Example 5.11 we have $o_{JS}(P, g) = (P, f)$ where (P, f) is oriented from left to right.

The map JS is defined as $JS := M(\_,\_) \circ /_{symm} \circ \Phi$. This definition coincides with division by the equivalence relation generated by reflections of subtrees.

Example 5.14 We consider how a sequence of reflections of subtrees maps the CMl trees from Example 5.11 to each other. Here, we denote by a_λ the reflection of the subtree with root labeled λ.



The proof that the two possible definitions for Φ coincide and that the diagram from Fig. 2 commutes are straightforward.

As a result, the map Φ can be seen as a discrete version of the map Ψ from Curry (2019, Cor 6.11). Moreover, moving from the setting of Curry (2019) to the setting of Johnson and Scoville (2022) basically means to divide by symmetry equivalences which allows the authors of Johnson and Scoville (2022) to generalize the construction of the induced merge tree to dMfs on trees. In order to extend our generalization Theorem 5.6 to the oriented case, one would need to find a notion of orientation-preserving cm equivalences which distinguishes symmetry-equivalent dMfs.

6 Further directions and possible applications

In this section we want to take a look at possible applications of our results. The structured overview of the different notions of equivalence of discrete Morse functions and their connections to each other might be useful to explore the space of discrete Morse functions on a given simplicial complex. Even though the construction of the induced merge tree from Johnson and Scoville (2022) does not easily extend to arbitrary simplicial complexes, the notions of equivalence from this article do. Hence, one could try to use e.g. the notion of symmetry equivalence to structure the space of discrete Morse functions in a nice way, i.e. into orbits of a groupoid action. In a second step, one could then try to define the space of merge trees as a quotient of the space of discrete Morse functions on some "large enough" simplicial complex. With such a construction, one could assemble the results from this article and the results from Curry et al. (2021) into an analysis of a larger instance of the persistence map.

Aside from this, one could try to generalize and enhance the construction of the induced merge tree from Johnson and Scoville (2022) to arbitrary simplicial complexes. Then, in a second step, one could try to use the induced merge tree to find possible cancellations of pairs of critical simplices. This way, the induced merge tree might be helpful to optimize discrete Morse functions.

Furthermore, one could study the possible Morse orders on a given merge tree. With a classification of all Morse orders on a given merge tree T, one could classify all discrete Morse functions on any given tree that induce T as the induced merge tree with the help of Definition 3.20 and cm equivalences.

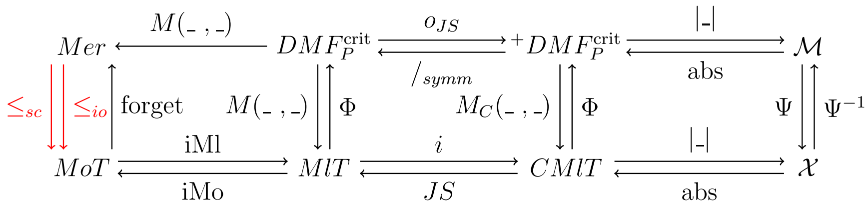


Fig. 2 Relationship to the continuous case

Ultimately, the proofs of Theorem 5.6 and all the lemmas that lead to it describe exactly which information is lost when considering the induced MI tree instead of the original dMf. This knowledge might be useful for applications in TDA because it tells the user which kind of features will not be seen by the induced merge tree, and hence, by the persistent zeroth homology and the barcode. Moreover, the knowledge about the exact lost information can be used to enhance the induced merge tree with extra structure such that it no longer disregards certain desired information.

Acknowledgements The author would like to thank Benjamin Johnson and Nicholas A. Scoville for the suggested questions in their article (Johnson and Scoville 2022) which led to this project. Furthermore, the author would like to thank the mathematical faculty of Ruhr-University of Bochum, especially the chair for topology, for the great scientific environment in which this endeavor was started. In addition to that the author thanks Max Planck Institute for Mathematics for the great scientific environment in which this project was finished. Most notably the author thanks his advisor, Viktoriya Ozornova, for her advice and the many helpful discussions which increased the quality of this article. Last but not least, the author thanks the anonymous referees for the helpful and detailed feedback.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The author states that there is no conflict of interest.

References

Baryshnikov, Y.: Time Series, Persistent Homology and Chirality. (2019). arXiv:1909.09846

Curry, J, DeSha, J., Garin, A., Hess, K., Kanari, L., Mallery, B.: From Trees to Barcodes and Back Again II: Combinatorial and Probabilistic Aspects of a Topological Inverse Problem. 07 (2021). arXiv:2107.11212v2

Carr, H., Snoeyink, J., Axen, U.: Computing contour trees in all dimensions. Comput. Geom.: Theory Appl. **24**(2), 75–94 (2003)

Curry, J.: The fiber of the persistence map for functions on the interval. J. Appl. Comput. Topol. **2**(3), 301–321 (2019)

Forman, R.: Morse theory for cell complexes. Adv. Math. **134**, 90–145 (1998)

Forman, Robin: A user's guide to discrete morse theory. Sem. Lothar. Combin. **48**, 12 (2001)

Heine, C., Leitte, H., Hlawitschka, M., Iuricich, F., De Floriani, L., Scheuermann, G., Hagen, H., Garth, C.: A survey of topology-based methods in visualization. Comput. Graph. Forum **35**, 643–667 (2016)

Johnson, B., Scoville, N.A.: Merge trees in discrete Morse theory. Res. Math. Sci. **9**(3), 1–7 (2022)

Kweon, I.S., Kanade, T.: Extracting topographic terrain features from elevation maps. CVGIP: Image Underst. **59**, 171–182 (1994)

Liu, S., Maljovec, D., Wang, B., Bremer, P.-T., Pascucci, V.: Visualizing high-dimensional data: advances in the past decade. IEEE Trans Vis. Comput. Graph. **23**(3), 1249–1268 (2016)

Nanda, V., Tamaki, D., Tanaka, K.: Discrete Morse theory and classifying spaces. Adv. Math. **340**, 723–790 (2018)

Oesterling, P., Heine, C., Weber, G.H., Morozov, D., Scheuermann, G.: Computing and Visualizing Time-Varying Merge Trees for High-Dimensional Data. Topological Methods in Data Analysis and Visualization, pp. 87–101 (2017)

Oesterling, P., Heine, C., Weber, G.H., Scheuermann, G.: Visualizing ND point clouds as topological landscape profiles to guide local data analysis. IEEE Trans. Vis. Comput. Graph. **19**, 514–526 (2013)

Shinagawa, Y., Kunii, T.L., Kergosien, Y.L.: Surface coding based on morse theory. IEEE Comput. Graph. Appl. **11**(05), 66–78 (1991)

Tarasov, S.P., Vyalyi, M.N.: Construction of contour trees in 3D in O(n log n) steps. In: SCG 98: Proceedings of the Fourteenth Annual Symposium on Computational Geometry, pp. 68–75 (1998)

van Kreveld, M., van Oostrum, R., Bajaj, C., Pascucci, V., Schikore, D.: Contour trees and small seed sets for isosurface traversal. In: SCG 97: Proceedings of the Thirteenth Annual Symposium on Computational Geometry, pp. 212–220 (1997)

Weber, G.H., Bremer, P.T., Pascucci, V.: Topological landscapes: a terrain metaphor for scientific data. IEEE Trans. Vis. Comput. Graph. **13**, 1416–1423 (2007)

Yan, L., Wang, Y., Munch, E., Gasparovic, E., Wang, B.: A structural average of labeled merge trees for uncertainty visualization. IEEE Trans. Vis. Comput. Graph. **26**, 832–842 (2019)

4. Erratum

In Definition 2.42, although the total order on $\mathrm{Cr}(f)$ induces chains of inclusions $X^f_{c0}[v] \subset X^f_{c1}[v] \subset \ldots$, these inclusions only induce chains of group homomorphisms $\widetilde{Aut}(X^f_{c0}[v]) \to \widetilde{Aut}(X^f_{c0}[v]) \to \ldots$, which are not necessarily inclusions of subgroups. Moreover, it turns out that the definition of the sublevel automorphism group is not quite desirable for further investigations. For a better definition, we refer to Definition 2.9.

CHAPTER III

On Cycles and Merge Trees

Julian Brüggemann and Nicholas A. Scoville

This chapter is available as a preprint under Brüggemann and Scoville 2023. The project "On cycles and merge trees" is collaborative work of the author of this thesis (referred to as the first author) and Nicholas A. Scoville (referred to as the second author). Both authors discussed all the contents of "On cycles and merge trees", and reviewed and edited all its written parts. The mathematical ideas, details, and initial formulations for the different sections have been developed separately. In particular, the first author is mainly responsible for sections 2, 3, 5 and 6. The second author is mainly responsible for section 4. The introduction was written jointly by both authors.

Abstract. In this paper, we extend the notion of a merge tree to that of a generalized merge tree, a merge tree that includes 1-dimensional cycle birth information. Given a discrete Morse function on a 1-dimensional CW complex, i.e. a multigraph, we construct the induced generalized merge tree. We give several notions of equivalence of discrete Morse functions based on the induced generalized merge tree and how these notions relate to one another. As a consequence, we obtain a complete solution to the inverse problem between discrete Morse functions on 1-dimensional CW complexes and generalized merge trees. After characterizing which generalized merge trees can be induced by a discrete Morse function on a simple graph, we give an algorithm based on the induced generalized merge tree of a discrete Morse function $f\colon X \to \mathbb{R}$ that cancels the critical cells of f and replaces it with an optimal discrete Morse function.

1. Introduction

Let X be a simplicial complex along with a sequence of subcomplexes $\emptyset = X_0 \subseteq X_1 \subseteq \cdots \subseteq X_n = X$ known as a filtration. In the burgeoning field of topological data analysis, a filtration is often given by a sampling of points based on some increasing parameter. Geometrical and topological features of X are then estimated by studying the persistence of certain topological features Polterovich et al. 2020. When the topological feature in question is the number of connected components, the persistence over the lifetime of the filtration is given by birth and death information and is summarized in a barcode or persistence diagram Oudot 2015; Carlsson and Vejdemo-Johansson 2022. If one wishes to not only determine birth and death information from the filtration but also how the components are evolving, that is, which components are merging with which, one associates a merge tree tree to the filtration. Because the merge tree carries with it this extra information, merge trees are a rich topic of study in both the theoretical and computational settings Curry et al. 2022; Curry 2019; Morozov, Beketayev, and Weber 2013; Gasparovic et al. 2022; Cardona et al. 2022. Merge trees have originally been introduced as an approximation to contour trees, a special case of Reeb graphs, in the context of visualization Carr, Snoeyink, and Axen 2003.

One way to induce a filtration on X is with a discrete Morse function (dMf) Forman 1998; Forman 2002. Such a function f induces a filtration by considering subcomplexes associated to each critical value of f. The induced merge tree of a dMf on a tree, or 1-dimensional acyclic complex, was introduced in Johnson and Scoville 2022. There the authors showed that a certain class of merge trees could be realized as the induced merge tree of a star graph. The authors went on to conjecture that any merge tree could be the induced merge tree of a certain dMf on a path. This conjecture was recently proved in Brüggemann 2022.

The goal of this paper is to extend the theory of merge trees and discrete Morse theory to include cycles. More specifically, given any 1-dimensional CW complex (i.e. a graph with or without multiedges) equipped with a dMf, we define a generalized induced Morse labeled merge tree (Definition 2.7) associated to this dMf. The generalized induced Morse labeled merge tree keeps track of not only component birth, death, and merge information but also cycle birth information via a node with a single child. After defining some basic properties, we introduce an equivalence relation on connected graphs called component-merge equivalence (CM equivalence, Definition 2.11) and show that there is a one-to-one correspondence between the set of CM equivalence classes of dMfs with only critical cells and the set of isomorphism classes of generalized Morse labeled merge tree in Theorem 3.1. In addition, we determine when a given generalized merge tree can be realized by an induced Morse function on a graph without multiedges. Unlike the case of merge trees, not all generalized merge trees can be realized. Theorem 4.1 gives a simple counting condition for when a generalized merge tree can be realized by a dMf on a simple graph. The proof is constructive and builds off of the merge tree construction in Brüggemann 2022, Theorem 5.9. Finally in Section 5, we give an algorithm on merge trees induced by a dMf in order to cancel critical cells of the dMf. The algorithm allows for some options depending on whether one wishes to preserve homeomorphism type of the graph or find an optimal matching. We briefly compare the algorithm to some similar algorithms from the literature Lewiner, Lopes, and Tavares 2003a; Rand and Scoville 2020.

In the last section, we consider possible future directions and applications.

2. Preliminaries on DMfs and Merge Trees

We recall and introduce the necessary notions for this work. In this article, we use the term graph for finite abstract multigraphs, possibly with self-loops. That is, graphs in this work may have multiple edges between two given vertices, and they can have self-loops, i.e. edges of the form (x, x). This notion of graph can be geometrically interpreted as 1-dimensional CW complexes.

On the other hand, we will use the term regular[1] if X does not contain a self-loop, and we call X a simple graph if X is regular and there is at most one edge between two given vertices. Simple graphs correspond to 1-dimensional simplicial complexes. Since we consider graphs as geometric objects, we also use geometric terms like cells, simplices, and faces to describe them. For any graph X, we use $v(X), e(X)$, and $b_1(X)$ to denote the number of vertices, edges, and cycles of X, respectively. If an edge $e = uv$ for vertices u and v, we say that u and v are the endpoints of e.

One key feature of this work is that, as usual in works related to topological data analysis, the involved filtrations are considered as part of the data of the space under investigation, rather than just a tool to analyze a space. In this work, the filtrations are given by one of the most central notions of the article, namely that of a discrete Morse function.

Definition 2.1 (Johnson and Scoville 2022, Definition 2.2/Benedetti 2016, Section 2.1). Let X be a graph, not necessarily connected. A function $f\colon X \to \mathbb{R}$ is a *discrete Morse function (dMf)* if it fulfills:

Monotonicity: For cells $\sigma \subset \tau$ we have $f(\sigma) \leq f(\tau)$.
Semi-injectivity: $|f^{-1}(\{z\})| \leq 2$ for all $z \in \mathbb{R}$.

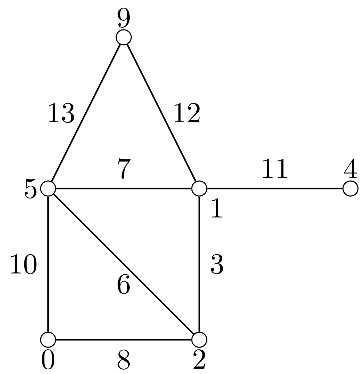
[1]We use the term regular in this fashion because we consider the graphs to be combinatorial models for topological spaces, i.e. 1-dimensional regular CW complexes. This should not be confused with the use of the term regular in graph theory.

Generacity: For cells $\sigma, \tau \in X$, if $f(\sigma) = f(\tau)$, then either $\sigma \subseteq \tau$ or $\tau \subseteq \sigma$ holds.

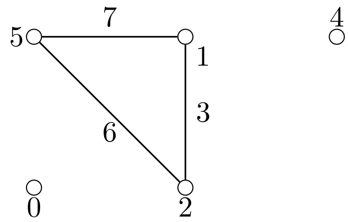
A cell σ of X is *critical* if σ is the unique preimage of $f(\sigma)$. Otherwise, σ is called *matched*. If X is not regular, we additionally require that all self-loops are critical with respect to f.

For any $a \in \mathbb{R}$, the *sublevel subcomplex* of X at a is $X_a = \{\sigma \in X : f(\sigma) \leq a\}$. The connected component of $\sigma \in X$ is denoted $X[\sigma]$. We use the notation $X_{a-\varepsilon}$ to denote the sublevel subcomplex of X immediately preceding a, i.e., $X_{a-\varepsilon} := \{\sigma : f(\sigma) < a\}$.

Example 2.1. Define $f\colon X \to \mathbb{R}$ by



Then f is a dMf with each value critical. The sublevel complex X_7 is given by



where $X_7[f^{-1}(1)] = X_7[f^{-1}(6)]$ is the 3-cycle while $X_7[f^{-1}(4)]$ is simply the isolated vertex labeled 4.

Remark 2.1. This definition of dMfs, due to B. Benedetti, is not equivalent to the more general definition originally given by Forman Forman 1998. Nonetheless, the given definition is generic in the sense that any dMf in the sense of Forman can be modified to fulfill the definition above without changing the induced acyclic matching. Nonetheless, adjusting a dMf in Forman's sense to become a dMf in the sense of Definition 2.1 will in general change the associated filtration of the complex at hand. The definition stated above has the advantage that critical cells are distinguished by their critical values and at each level, at most either one critical cell or one pair of matched cells is added to the sublevel complex. We chose to still use the term "dMf" in order to follow Benedetti's notation. This should not lead to confusion because all dMfs in this work satisfy Definition 2.1.

The condition that self-loops must be critical is a standard approach to dMfs on CW complexes that fail to be regular. In detail, one requires non-regular faces to remain unmatched. In the 1-dimensional case, self-loops are the only way to break regularity.

While discrete Morse theory provides a well-developed framework for filtered spaces, we apply this framework to one specific kind of topological information, namely the development of connected components throughout the filtration. The development of connected components is summarized in the merge tree.

Definition 2.2 (Johnson and Scoville 2022, Definition 3.1). A rooted tree is called *binary* if it is a rooted tree where each vertex has at most two children. A binary tree is called *full* if each node has either 0 or 2 children. A binary tree is called *chiral*[2] if each vertex is equipped with a label of either L or R with the children of a vertex having one label L and the other R. A *merge tree* is a chiral full binary tree.

For a node p of a rooted tree, we denote by $T(p)$ the *rooted subtree with root* p, that is, the subtree that consists of p and all of p's descendants.

Remark 2.2. Merge trees have been originally introduced in Carr, Snoeyink, and Axen 2003[3] in order to compute contour trees efficiently. Although the original introduction of merge trees was of a combinatorial nature, it is also common in the literature to perceive merge trees of filtration maps $f\colon X \to \mathbb{R}$ in a more geometric way: as the quotient space $X/\sim$, where $\sim$ is defined by $x \sim y$ if and only if $f(x) = f(y)$ and x and y are in the same connected component of the sublevel set $f^{-1}(-\infty, f(x)]$. If one assumes the filtration function f to be generic[4] in the sense that at each time at most two connected components merge, the induced merge tree will be binary. From that point of view one obtains a merge tree in the sense of Definition 2.2 in the following way: consider new connected components in the filtration as leaves and consider points where connected components merge as inner nodes. This way the merge tree in the geometric sense becomes a full binary tree. The chirality in Definition 2.2 is an additional structure that is somewhat motivated by the Elder rule: connected components that were created earlier should persist longer. This is reflected in the way that any (generalized) merge tree constructed according to Definition 2.7 has the property that bars in the induced 0-barcode correspond to maximal paths in the induced merge tree that only go through nodes of the same chirality.

In that sense, it would be more adequate to refer to merge trees in the sense of Definition 2.2 as chiral generic merge trees but we decide against that in the interest or brevity because all merge trees in this work are chiral and generic.

Our main object of study is given in Definition 2.3, that of a generalized merge tree. It generalizes the notion of a merge tree in the following sense: while a merge tree keeps track of component information, the generalized merge tree will also keep track of 1-dimensional cycle information. A cycle is represented by a child vertex with no sibling. See Example 2.3.

Definition 2.3. A *generalized merge tree* T is a chiral binary tree T such that each leaf has a sibling, and inner nodes without a sibling have the same chirality as their parent node. By convention, we say that the root always has chirality L. Furthermore, the root is never regarded as a leaf, even if it only has one child node.

For nodes c of generalized merge trees we use the notation c_l/c_r for the $left/right$ child node of c.

Remark 2.3. Generalized merge trees may have nodes without siblings. Due to that, generalized merge trees generalize merge trees in the sense of Definition 2.2 in the way that the generalized merge trees allows to break the condition of being full in the way that some nodes might have only one child node. This breaks the property of being full in a different way than the usual notion of merge tree does: non-generic[5] filtration functions induce merge trees where each node might have two or more children but single children will never occur. We impose the condition that an only child has the same chirality as its parent node for technical

[2]For full binary trees, the notion of chirality is equivalent to the notion of ordered trees, i.e. trees that for each node have a specified total order on the set childrens of that node. We define chirality in a more general fashion here in order to use the same term for the definition of general merge trees.

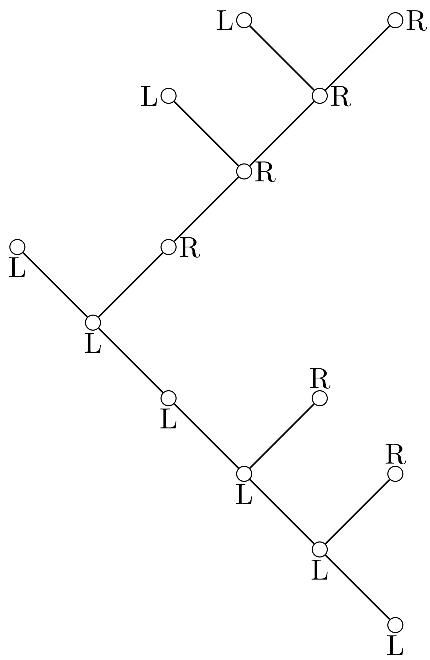
[3]In that work, merge trees have been referred to as join trees.

[4]In this work, this sense of generacity is ensured by the generacity and semi-injectivity properties in Definition 2.1.

[5]In the sense as in Remark 2.2.

reasons. We need this convention so the constructions in Definition 2.8 and Definition 2.7 make part 2 of Theorem 3.1 work.

Example 2.2. The tree T below is a generalized merge tree:



Note that each vertex has at most one child and that inner nodes without a sibling have the same chirality (label of L or R) as their parent. This generalized merge tree is said to have 6 leaves, as the root node at the bottom (labeled L) is not considered a leaf by convention.

Definition 2.4 (Brüggemann 2022, Definition 2.17). Let T be a generalized merge tree. We call a total order $\leq$ on the nodes of T a *Morse order* if it fulfills the following two properties for all generalized merge subtrees T' of T:

(1) The restriction $\leq_{|T'}$ attains its maximum on the root p of T'.
(2) The minimum of $\leq_{|T'}$ has the same chirality as p.

We call a generalized merge tree together with a Morse order $(T, \leq)$ a *generalized Morse ordered merge tree (gMo tree)*.

Remark 2.4. Assuming property 2 of Definition 2.4 for every subtree T' with root p of T is equivalent to either of the following:

- For any subtree T' with root p of T, the restriction $\leq_{|T'}$ attains its minimum on the subtree with root p_l/p_r if L/R is the chirality of the root p of T'.
- For any subtree T' with root p of T, all nodes on the shortest path between p and the minimum of $\leq_{|T'}$ have the same chirality as p.

The equivalence can be proved by an inductive argument over all nodes of the shortest path between p and the minimum.

Definition 2.5 (Brüggemann 2022, Definition 2.19). We call a generalized merge tree (T, λ) with an injective map $\lambda\colon T \to \mathbb{R}$ such that λ induces a Morse order on T a *generalized Morse labeled merge tree (gML tree)*. Any such map λ is called a *Morse labeling* on T.

Let (T, λ) and (T', λ') be gMl trees. An *order equivalence* $(\varphi, \psi)\colon (T, \lambda) \to (T', \lambda')$ of gMl trees is a pair of maps consisting of an isomorphism of the underlying generalized merge trees $\varphi\colon T \to T'$ and a bijection $\psi\colon \mathbb{R} \to \mathbb{R}$ such that the restriction $\psi_{|\mathrm{im}(\lambda)}\colon \mathrm{im}(\lambda) \to \mathrm{im}(\lambda')$ is order preserving.

Proposition 2.1. *Let gMoT be the set of generalized Morse ordered merge trees and let gMlT be the set of generalized Morse labeled merge trees. Then (i) taking the Morse order induced by a Morse labeling and (ii) using a Morse order with labels* $\{0, \cdots, |V(T)|-1\}$ *to induce a Morse labeling, define inverse bijections*

$$iMl\colon gMoT/_{\cong} \longleftrightarrow gMlT/_{\sim}\colon iMo$$

where $\sim$ *denotes order equivalence.*

Proof. The proof is analogous to Brüggemann 2022, Proposition 3.26. □

Definition 2.6. Let (X, f) be a dMf on a graph. We call a critical edge $\sigma \in X$ a *closing edge* if there is a subdivision Δ of S^1 contained as a subcomplex $\Delta \subseteq X$ which contains σ such that $f(\sigma)$ is the maximum of f on Δ.

We define $C(X, f) \coloneqq \{c \in X | c \text{ is closing}\}$ to be the *set of closing edges* of (X, f) and $(\bar{X}, \bar{f}) \coloneqq (X \setminus C(X, f), f_{|X\setminus C(X,f)})$ to be the *spanning tree (or spanning forest, if* X *is not connected) induced by* f of X.

Remark 2.5. In the previous definition, the subdivison of S^1 that any closing edge σ must be part of does not need to be unique. Nonetheless, the removal of σ would lead to the reduction of the first Betti number by one. Moreover, the notion of closing edges is well-defined because the edge σ being closing implies that it is the unique maximal edge of all subdivisions of S^1 in $X_{f(\sigma)}[\sigma]$ that contain σ.

Furthermore, it is immediate that $(\bar{X}, \bar{f}) \coloneqq (X \setminus C(X, f), f_{|X\setminus C(X,f)})$ is a dMf on a tree. It is also immediate that self-loops are always closing edges.

Definition 2.7 (Induced gML Merge Tree). [6] Let $f\colon X \to \mathbb{R}$ be a dMf on a connected graph X. If f has a single critical vertex v with $f(v) = r$ and no critical edges, then the induced generalized Morse labeled merge tree is a single vertex c_{λ_v} with label $\lambda_f(c_{\lambda_v}) = r$ and chirality L. Otherwise, let $\sigma_n > \sigma_{n-1} > \cdots > \sigma_1 > \sigma_0$ be the critical edges of (X, f) ordered by their values under f. The gML merge tree induced by (X, f), denoted $M(X, f)$ with labeling λ_f, is constructed inductively, inducing over the decreasing order of the critical edges.

Start by constructing a root node called $M(\sigma_n)$, labeled $\lambda_f(M(\sigma_n)) \coloneqq f(\sigma_n)$, and left chirality.

Now begin the induction over the decreasing order of the critical edges starting from $n, \ldots, 0$. For σ_i a critical edge with endpoints u and v, the addition of σ_i at level subcomplex $X_{f(\sigma_i)}$ either creates a cycle or connects two components. Formally,

(1) The critical edge σ_i is closing.[7]
(2) The critical edge σ_i is not closing.[8]

If σ_i satisfies (1), we construct a child node c of $M(\sigma_i)$ with label $\lambda_f \coloneqq \max\{f(\sigma) | \sigma \in X_{f(\sigma_i)-\varepsilon},\ \sigma \text{ is critical}\}$ and the same chirality as $M(\sigma_i)$. The node c then corresponds to the edge of X labeled λ.

If σ_i satisfies (2), we construct two child nodes c_{λ_v} and c_{λ_w} of $M(\sigma_i)$. Define $\lambda_v \coloneqq \max\{f(\sigma) | \sigma \in X_{f(\sigma_i)-\varepsilon}[v], \sigma \text{ critical}\}$ and $\lambda_w \coloneqq \max\{f(\sigma) | \sigma \in X_{f(\sigma_i)-\varepsilon}[w], \sigma \text{ is critical.}\}$. Then label the new nodes $\lambda_f(c_{\lambda_v}) \coloneqq \lambda_v$ and $\lambda_f(c_{\lambda_w}) \coloneqq \lambda_w$. If $\min\{f(\sigma) | \sigma \in X_{f(\sigma_i)-\varepsilon}[v]\} <$

[6] This construction generalizes Johnson and Scoville 2022, Theorem 3.5.

[7] This condition is equivalent to $b_1(X_{f(\sigma_i)} \setminus \sigma_i) = b_1(X_{f(\sigma_i)}) - 1$, which in turn is equivalent to $b_0(X_{f(\sigma_i)} \setminus \sigma_i) = b_0(X_{f(\sigma_i)})$.

[8] This condition is equivalent to $b_1(X_{f(\sigma_i)} \setminus \sigma_i) = b_1(X_{f(\sigma_i)})$, which in turn is equivalent to $b_0(X_{f(\sigma_i)} \setminus \sigma_i) = b_0(X_{f(\sigma_i)}) + 1$.

$\min\{f(\sigma)|\sigma \in X_{f(\sigma_i)-\varepsilon}[w]\}$, we assign c_{λ_v} the same chirality (L or R) as c_{σ_i} and give c_{λ_w} the opposite chirality.

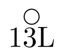
Continue the induction over all of the critical edges of X to obtain the Morse labeled merge tree $M(X, f)$ induced by f along with labeling λ_f.

Remark 2.6. It is important to note that in the inductive step after creating the child(ren) of $M(\sigma_i)$, if critical edge σ_{i-1} exists, then $M(\sigma_{i-1})$ is a vertex on the current constructed generalized merge tree (it may be a child of a vertex other than $M(\sigma_i)$).

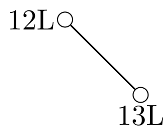
Example 2.3. We will construct the induced Morse labeled merge tree of Example 2.1. Since the dMf in this example is injective, we will name each vertex or edge by σ_i where $f(\sigma) = i$. The first step of Definition 2.7 is to list the critical edges in increasing order:

$$\sigma_{13} > \sigma_{12} > \sigma_{11} > \sigma_{10} > \sigma_8 > \sigma_7 < \sigma_6.$$

Here we index by the value under the dMf as opposed to the integers $6, \ldots, 0$ but it does not matter. For the base case, we create a node called $M(\sigma_{13})$ with value $\lambda_f(M(\sigma_{13})) = f(\sigma_{13}) = 13$ with chirality L (by definition); that is, we begin with



Moving on from the base case, σ_{13} creates a cycle, i.e., it is a closing edge so that it has a single child $M(\sigma_{12})$ with label $\lambda_f = \max\{f(\sigma) : \sigma \in X_{13-\epsilon}, \sigma \text{ critical }\} = 12$ and chirality that of $M(\sigma_{13})$ which is L. This yields



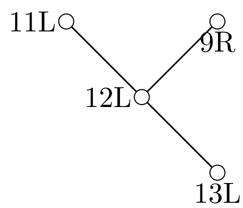
Now since $M(\sigma_{12})$ is not a closing edge, we construct two child does $c_{\lambda_{\sigma_9}}$ and $c_{\lambda_{\sigma_1}}$ of $M(\sigma_{12})$. There values are then computed as

$$\lambda_{\sigma_9} = \max\{f(\sigma) : \sigma \in X_{12-\epsilon}[\sigma_9], \sigma \text{ critical }\} = 9$$

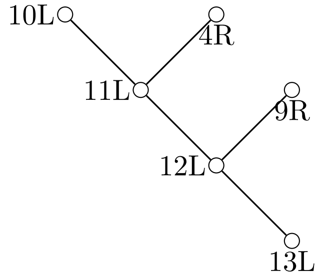
and

$$\lambda_{\sigma_1} = \max\{f(\sigma) : \sigma \in X_{12-\epsilon}[\sigma_1], \sigma \text{ critical }\} = 11.$$

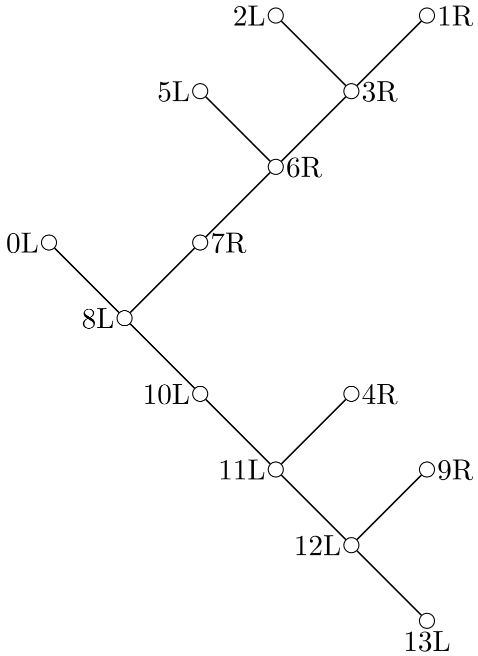
This amounts to determining the largest critical value in the connected component of the vertex in question. hence the two children of $M(\sigma_{12})$ are labeled 9 and 11. Finally, $0 = \min\{f(\sigma) : \sigma \in X_{12-\epsilon}[\sigma_1], \sigma \text{ critical }\} < \min\{f(\sigma) : \sigma \in X_{12-\epsilon}[\sigma_9], \sigma \text{ critical }\} = 9$ so that $c_{\lambda_{\sigma_1}}$ shares the same chirality as its parent while $c_{\lambda_{\sigma_9}}$ has the opposite chirality. In sum, we have so far



The induction again continues at σ_{11} which is not a closing edge. The two child nodes of $M(\sigma_{11})$ have values 4 and 10 with the nodes given value 4 sharing the same chirality as $M(\sigma_{11})$ so that we have



Continuing in this manner we arrive at the induced Morse labeled merge tree given by



which is the same merge tree as in Example 2.2.

REMARK 2.7. The construction of the induced gMl tree comes with a bijection $M\colon X \to V(M(X,f))$ that restricts to bijections between the critical vertices of X and leaves of $M(X,f)$, between the non-closing critical edges of X and parents with two children of $M(X,f)$, and between the closing edges (cycles) of X and parents with one child in $M(X,f)$.

Furthermore, the proof that the construction indeed produces a gMl tree is completely analogous to Brüggemann 2022, Proposition 2.20, respectively Johnson and Scoville 2022, Theorem 9.

It is also possible to apply the construction to dMfs on non-connected graphs. In that case the algorithm produces a merge forest and one can deal with each connected component separately.

LEMMA 2.1. *Let (X,f) be a dMf on a graph and let $M(X,f)$ be the induced gML tree. For any critical cell $s \in X$, the rooted subtree $T(M(s))$ of $M(X,f)$ is induced by the connected component $X_{f(s)}[s]$ of s in the sublevel complex of level $f(s)$. Moreover, the rooted subtree $T(M(s))$ is isomorphic to $M(X_{f(s)}[s], f_{|X_{f(s)}[s]})$ as merge trees if and only if $M(s)$ has chirality L. If $M(s)$ has chirality R, then $T(M(s))$ is isomorphic to $M(X_{f(s)}[s], f_{|X_{f(s)}[s]})$ as rooted binary trees but the chiralities of all nodes are opposite to the ones of their respective nodes in the other tree.*

Proof. We observe that by Definition 2.7 the label of $M(s)$ is $f(s)$ and the chirality of $M(s)$ is decided by the minimum of $f_{|X_{f(s)}[s]}$ in comparison to the minimum of the connected component that $X_{f(s)}[s]$ got divided from at level $f(s)$. It follows inductively by construction that all nodes of the subtree $T(M(s))$ are induced by critical cells of $X_{f(s)}[s]$ because they are constructed by removing critical edges of $X_{f(s)}[s]$.
The isomorphism as rooted binary trees is constructed by the same inductive argument. Since the chirality depends on the chirality of the respective parent node, said isomorphism is compatible with the chirality if and only if the root of the rooted subtree $T(M(s))$, namely $M(s)$, has chirality L. This is true because the root of $M(X_{f(s)}[s], f_{|X_{f(s)}[s]})$ by convention always has chirality L. □

Definition 2.8. Let (T, λ) be a gML tree. Let $C(T) \subset V(T)$ be the set of nodes that have exactly one child node. We refer to the elements of $C(T)$ as *cycle nodes*. We denote by $(\bar{T}, \lambda)$ the Morse labeled merge tree that is obtained from (T, λ) by removing the cycle nodes by connecting their parent nodes directly to their child nodes. We call $(\bar{T}, \lambda)$ the underlying Morse labeled merge tree of (T, λ).

We obtain a dMf on a graph $f_\lambda \colon X \to \mathbb{R}$ from (T, λ) in two steps as follows: In a first step, we construct the induced dMf on a path (P, f_λ) as in Brüggemann 2022, Definition 3.21. For the second step, for each node c of $C(T)$ we add an edge parallel to the edge corresponding to c's oldest descendant which has two children to P. We denote the graph obtained this way by X and extend the function $f_\lambda \colon P \to \mathbb{R}$ to X using the values of λ on the corresponding nodes. We denote the pair (X, f_λ) by $\Phi(T, \lambda)$ and consider that we also obtained a bijection $\phi \colon V(T) \to \Phi(T, \lambda)$.

Lemma 2.2. *We have $M(\bar{X}, f_{|\bar{X}}) \cong \bar{M}(X, f)$ as Morse labeled merge trees.*

Proof. The construction of the induced generalized merge tree induces a bijection $M \colon X \to V(M(X, f))$. It follows immediately by construction that M bijectively maps closing edges to nodes of $C(M(X, f))$. Hence removing the closing edges from (X, f), that is, passing on to $(\bar{X}, f)$, precisely removes the nodes of $C(M(X, f))$, which corresponds to passing on to $\bar{M}(X, f)$. Hence, the statement holds because the values of f on non-closing edges are not changed. □

Definition 2.9 (Brüggemann 2022, Definition 2.42). Let $f \colon X \to \mathbb{R}$ be a dMf on a graph. For each non-empty connected component $X_c[v]$ of a sublevel complex X_c we denote by $\mathrm{Aut}(X_c[v])$ the group of simplicial automorphisms of $X_c[v]$. Each $\xi \in \mathrm{Aut}(X_c[v])$ can be extended by the identity to a set function that is a self-bijection $X \to X$. The group $\widetilde{\mathrm{Aut}}(X_c[v])$ is defined to be the group of said extensions of elements of $\mathrm{Aut}(X_c[v])$ by the identity[9]. The group operation on $\widetilde{\mathrm{Aut}}(X_c[v])$ is the composition of self-bijections of X. We call the elements of $\widetilde{\mathrm{Aut}}(X_c[v])$ *elementary sublevel automorphisms*. We define the *sublevel automorphism group* of (X, f), denoted by $\mathrm{Aut}_{sl}(X, f)$, as

$$\mathrm{Aut}_{sl}(X, f) := \mathop{\ast}_{c \in \mathrm{Cr}(f), v \in X} \widetilde{\mathrm{Aut}}(X_c[v]) / \sim,$$

where $*$ denotes the free product of groups and $\sim$ is defined by

$$\xi \xi' b \sim \begin{cases} \xi \circ \xi' & \text{if } a, b \in \widetilde{\mathrm{Aut}}(X_c[v]) \text{ for the same } X_c[v] \\ \xi' \xi & \text{if } \xi, \xi' \text{ belong to different connected components of sublevel complexes} \end{cases}$$

We call the elements of $\mathrm{Aut}_{sl}(X, f)$ *sublevel automorphisms*.

[9]That is, elements of $\mathrm{Aut}(X_c[v])$ are self-bijections of X that restrict to a simlicial automorphism on $X_c[v]$) and to the identity on $X \setminus X_c[v]$).

2. PRELIMINARIES ON DMFS AND MERGE TREES

Note that an element $\xi \in \mathrm{Aut}_{sl}(X, f)$ is not necessarily an automorphism of X, but only a collection of self-bijections of X that restrict to automorphisms on certain connected components of some sublevel complex $X_c[v]$, and the identity outside of $X_c[v]$. Furthermore, notice that $\mathrm{Aut}_{sl}(X, f)$ is by definition isomorphic to the product of the $\widetilde{\mathrm{Aut}}(X_c[v])$. We chose to phrase $\mathrm{Aut}_{sl}(X, f)$ as a quotient of a free product to provide more clarity in Lemma 2.3 andProposition 2.2. Moreover, $\mathrm{Aut}_{sl}(X, f)$ is not a subgroup of the group of self-bijections of X.

Example 2.4. We consider three instructive examples of sublevel-automorphism groups:

(1) The path with n vertices P,
(2) the star graph with $n+1$ vertices S, and
(3) the cycle graph with n vertices C.

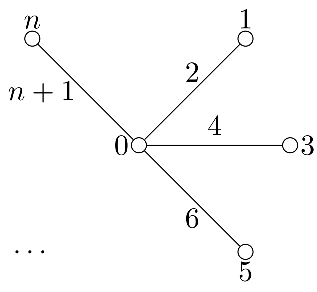
(1)

Consider the path P with n vertices with a critical dMf defined by counting from left to right:

We observe that each connected component of a sublevel set only has either the trivial group or a group generated by exactly one reflection as its automorphism group. Hence, we have $\mathrm{Aut}(P^f_{f(e)}[e]) \cong \Sigma_2$ for all edges e and $\mathrm{Aut}_{sl}(P, f) \cong \prod_{k=1}^{n-1} \Sigma_2$, where Σ_k denotes the symmetric group on k elements. This way, we realized the Young subgroup $\prod_{k=1}^{n-1} \Sigma_2 \subset \Sigma_{2(n-1)}$ as a group of sublevel automorphisms of a filtered space with the associated constant sequence of automorphism groups of sublevel complexes $\mathrm{Aut}(P^f_{f(e)}[e]) \cong \Sigma_2$.

(2)

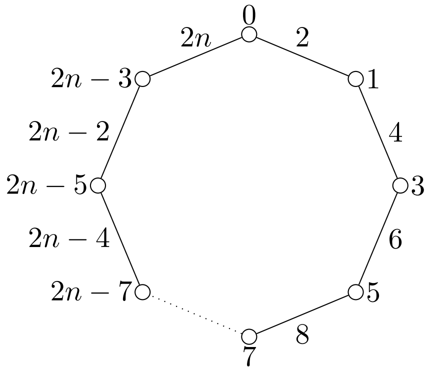
We consider the star graph with $n+1$ vertices S together with a critical dMf that attains its minimum at the center and otherwise assigns values pairwise to the outer vertices and their adjacent edges:



Let v_k be the vertex with label k. It is immediate that $S^f_{2k+2}[v_{2k+1}]$ is the star graph with $k+2$ vertices. Hence, we have $Aut(S^f_2[v_1]) \cong \Sigma_2$ and $\mathrm{Aut}(S^f_{2k+2}[v_{2k+1}]) \cong \Sigma_{k+1}$ for $k \geq 1$, where Σ_k denotes the symmetric group on k elements. We have $\mathrm{Aut}_{sl}(S, f) \cong \Sigma_2 \times \prod_{k=2}^{n} \Sigma_k$. So we realized $\Sigma_2 \times \prod_{k=2}^{n} \Sigma_k$ as a group of sublevel automorphisms of a filtered space with the associated sequence of groups of automorphisms of connected components of sublevel complexes $\mathrm{Aut}(S^f_2[v_1]) \cong \Sigma_2$ and $\mathrm{Aut}(S^f_{2k+2}[v_{2k+1}]) \cong \Sigma_{k+1}$ for $k \geq 1$.

(3)

Consider the cycle graph with n vertices C together with the following critical dMf:



Let e be any edge except for the one labeled $2n$. Then we have $\mathrm{Aut}(P^f_{f(e)}[e]) \cong \Sigma_2$ as in the first example. Let $\tilde{e}$ be the edge labeled $2n$. Then we have $\mathrm{Aut}(P^f_{f(\tilde{e})}[\tilde{e}]) \cong D_n$ where D_n denotes the dihedral group of order $2n$, i.e. the symmetries of the regular n-gon. This leads us to $\mathrm{Aut}_{sl}(S, f) \cong D_n \times \prod_{k=1}^{n-1} \Sigma_2$. This way, we realized $D_n \times \prod_{k=1}^{n-1} \Sigma_2$ as a group of sublevel automorphisms of a filtered space with the associated constant sequence of automorphism groups of sublevel complexes $\mathrm{Aut}(C^f_{f(e)}[e]) \cong \Sigma_2$ for $f(e) \leq 2n-2$ and $\mathrm{Aut}(C^f_{2n}[\tilde{e}]) \cong D_n$.

Definition 2.10. Let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ be dMfs on a graph X. We call f and g *sublevel-equivalent* if they have the same critical values and isomorphic sublevel complexes. If additionally $g = f \circ \xi$ holds for a sublevel automorphism $\xi \in \mathrm{Aut}_{sl}(X, f)$, then we call f and g *symmetry-equivalent.* We call the map ξ a *symmetry equivalence* from f to g.

We call two dMfs $f\colon X \to \mathbb{R}$ and $g\colon Y \to \mathbb{R}$ *symmetry-equivalent* if there is a simplicial isomorphism $\varphi\colon X \to Y$ such that f and $g \circ \varphi$ are symmetry-equivalent.

Having these definitions established, we are able to consider the action of $\mathrm{Aut}_{sl}(X, f)$ on the symmetry equivalence class of (X, f).

Remark 2.8. If two dMfs on graphs $f\colon X \to \mathbb{R}$ and $g\colon Y \to \mathbb{R}$ are symmetry-equivalent, then their sublevel automorphism groups are isomorphic because the two dMfs induce isomorphic filtrations.

We want to remark at this point that even though the elements of sublevel automorphism groups $\mathrm{Aut}_{sl}(X, f)$ are collections of self-bijections of the respective graph X, the group structure is different from the group of self-bijections of X. In particular, the group structure of $\mathrm{Aut}_{sl}(X, f)$ is constructed such that symmetry equivalences of different levels always commute with each other although the corresponding self-bijections of X do not necessarily commute. This is necessary for the desired action on the set of dMfs on X: symmetry equivalences exist due to the existence of different filtrations with isomorphic sublevel complexes that are embedded differently into X. That is, if φ and ψ are symmetry equivalences that belong to two connected components of sublevel complexes such that one is contained in the other, wlog φ is of a higher level than ψ, then φ changes the location of the connected component that belongs to ψ precisely such that the actions of φ and ψ on the set of dMfs commute.

Lemma 2.3. *Let $f\colon X \to \mathbb{R}$ be a dMf on a graph X and let $\xi \in \widetilde{\mathrm{Aut}}(X_c[v]) \subset \mathrm{Aut}_{sl}(X, f)$ be an elementary sublevel automorphism. Then $f \circ \xi$ is a dMf on X, which is symmetry equivalent to f.*

Moreover, a induces an isomorphism $\mathrm{Aut}_{sl}(X, f) \cong \mathrm{Aut}_{sl}(X, g)$ by precomposition.

Proof. In order to prove that function $f \circ \xi$ is a dMf, we note that ξ is a simplicial automorphism on $X_c[v]$ and the identity outside of $X_c[v]$. Since ξ is in particular a self-bijection of X, $f \circ \xi$ is still at most 2-1. Due to $X_c[v]$ being contained in a sublevel complex, all values of f outside of $X_c[v]$ are strictly larger than the ones inside $X_c[v]$. In particular, f is strictly monotone on all face relations at the boundary of $X_c[v]$, i.e. between simplices of $X_c[v]$ and simplices outside $X_c[v]$. Thus, the action of ξ outside of $X_c[v]$ does not affect monotonicity and generacy. Inside $X_c[v]$, ξ acts as a simplicial automorphism, which is why $f \circ \xi$ also satisfies monotonicity and generacy. Furthermore, $f \circ \xi$ is symmetry-equivalent to f by Definition 2.10.

For the second statement, note that ξ induces an isomorphism between the filtrations induced by $f, f \circ \xi$, respectively. That is, ξ bijectively maps connected components of sublevel complexes of f to connected components of sublevel complexes of g in an inclusion and filtration preserving way. Hence, it follows directly from the presentation of Aut_{sl} given in Definition 2.9 that ξ is an isomorphism. □

PROPOSITION 2.2. *Let X be a graph, let $f\colon X \to \mathbb{R}$ be a discrete Morse function, and let $\mathrm{Aut}_{sl}(X, f)$ be the group of sublevel automorphisms of (X, f). Then $\mathrm{Aut}_{sl}(X, f)$ acts on the symmetry equivalence class of f as follows: for any dMf g that is symmetry-equivalent to f, and any elementary sublevel equivalence ξ, we define $g * \xi := g \circ \tilde{\xi}$, where $\tilde{\xi} \in \mathrm{Aut}_{sl}(X, g)$ is the elementary sublevel equivalence that corresponds to ξ under the automorphism $\mathrm{Aut}_{sl}(X, f) \cong \mathrm{Aut}_{sl}(X, g)$ from Lemma 2.3. For arbitrary elements of $\mathrm{Aut}_{sl}(X, f)$, the group action is defined by the successive action of elementary sublevel automorphisms.*

PROOF. It follows from successive application of Lemma 2.3 that $g * \xi$ is well defined for any dMf g that is symmetry-equivalent to f, and any elementary equivalence ξ. The compatibility, i.e. that for any $\xi, \xi' \in \mathrm{Aut}_{sl}(X, f)$, we have $g * \xi * \xi' = g * (\xi \cdot \xi')$, where $\cdot$ denotes the multiplication in $\mathrm{Aut}_{sl}(X, f)$ follows by construction of the group action and $\mathrm{Aut}_{sl}(X, f)$, and by Remark 2.8. □

Next, we introduce the more general notion of component-merge equivalence:

DEFINITION 2.11 (Brüggemann 2022, Definition 2.50). Let (X, f) and (X', f') be critical dMfs on connected graphs. A *component-merge equivalence (CM equivalence)* of level a[10] is a bijection $\varphi\colon X \to X'$ such that at least one of the following two cases holds:

(1) φ is a symmetry equivalence that involves sublevel automorphisms of at most level a.
(2) φ fulfills the following:
  - $f' \circ \varphi = f$,
  - φ induces a bijection between the sets of connected components of sublevel complexes such that the restriction $\varphi_{|X_{a-\varepsilon}[v]}\colon X_{a-\varepsilon}[v] \to X'_{a-\varepsilon}[\varphi(v)]$ to each connected component is a CM equivalence of a level $\leq a$, and
  - the edge $\sigma \in X$ with $f(\sigma) = a$ merges two connected components $X_{a-\varepsilon}[v_1]$ and $X_{a-\varepsilon}[v_2]$ in $X_a[v_1] = X_a[v_2]$ if and only if the edge $\varphi(\sigma)$ merges the corresponding two connected components $X'_{a-\varepsilon}[\varphi(v_1)]$ and $X'_{a-\varepsilon}[\varphi(v_2)]$ in $X'_a[\varphi(v_1)] = X'_a[\varphi(v_2)]$. Otherwise, if the edge $\sigma \in X$ with $f(\sigma) = a$ does not merge two connected components but rather closes a circle within a connected component $X_{a-\varepsilon}[v]$, then and only then $\varphi(\sigma)$ closes a circle within $X'_{a-\varepsilon}[\varphi(v)]$.

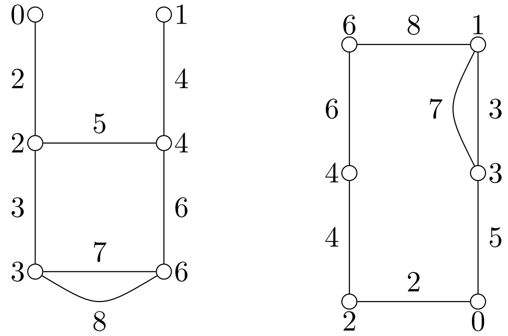
If φ re-attaches the critical edge labeled a, we call φ *non-trivial*. Moreover, if φ re-attaches the critical edge of level a and acts as a symmetry equivalence everywhere else, we say that φ is *elementary of level a*. If φ does not re-attach any critical edge, i.e., if φ is a symmetry equivalence, we call φ a *trivial CM equivalence*.

REMARK 2.9. Extending the notion of CM equivalences to dMfs with matched cells is a bit tedious. We would like to suggest getting rid of matched cells by identifying arbitrary dMfs

[10]We emphasize the notion of the level of CM equivalences in order to highlight the recursive nature of this definition. In situations when the specific level of a CM equivalence is not of importance, we sometimes drop the level in the notation.

on graphs with critical dMfs on the corresponding graph that arises by collapsing matched cells beforehand even though self-loops might arise in this process. Nonetheless, the newly created self-loops are critical by construction and the definition above works in this context.

Example 2.5. Let $f\colon X \to \mathbb{R}$ be the complex with dMf on the left and $f'\colon X' \to \mathbb{R}$ be the complex with dMf on the right.



Then a CM equivalence $\varphi\colon X \to X'$ of critical levels $a = 7$ is given by $\varphi(v) = v'$ whenever $f(v) = f'(v')$ on vertices and $\varphi(e) = e'$ whenever $f(e) = f'(e')$ on edges. We remark that according to Remark 2.9 matched simplices can be arbitrarily added and removed from connected components by CM equivalences. After performing the prescribed collapsed, it becomes clear that φ only re-attaches the edge labeled 7 from an edge between the critical vertices labeled 0 and 1 to a critical self-loop at the critical vertex labeled 1.

Remark 2.10. It is clear from the case distinction made in Definition 2.11 that any CM equivalence $\varphi\colon (X, f) \to (X', f')$ restricts to a bijection $\varphi_{|C(X,f)}\colon C(X, f) \to C(X', f')$.

Lemma 2.4. *Let $f\colon X \to \mathbb{R}$ and $f'\colon X' \to \mathbb{R}$ be CM-equivalent dMfs on multigraphs. Then $M(X, f) \cong M(X', f')$ holds as gML trees.*

Proof. Let φ be a CM equivalence $\varphi\colon (X, f) \to (X', f')$. Since we work with a generic version of dMfs which are at most 2-1, at most one non-trivial elementary CM equivalence of level a can occur for any level a because there is at most one critical edge labeled a in $(X, f), (X', f')$, respectively. Thus, we can decompose any CM equivalence into a sequence $(\varphi_a)_a$ of non-trivial elementary CM equivalences of decreasing levels such that each φ_a only changes the attachment of the single edge σ with $f(\sigma) = a$ and acts as a symmetry equivalence on the rest of graph and dMf. It suffices to consider a single level a because the statement then follows by induction from highest to lowest over all levels a.

For such a non-trivial elementary CM equivalence φ_a we consider the step of the construction of the induced Ml trees that considers the critical edge σ with $f(\sigma) = a$ and the critical edge $\varphi(\sigma)$. If σ is not closing, neither is $\varphi(\sigma)$ by Remark 2.10 and the inductive step follows by Brüggemann 2022, Proposition 2.52. In the case that σ is closing, so is $\varphi(\sigma)$ and we inductively assume that φ induces an isomorphism of induced generalized Ml trees everywhere outside the subtree corresponding to the connected component of $X^f_{a-\varepsilon}$ that the edge σ with $f(\sigma) = a$ is attached to. That is, on the rest of $M(X, f)$ the map $M(\varphi)$ is a bijection compatible with the chiral child relation onto $M(X', f')$ except possibly for the subtree of $M(X', f')$ which corresponds to the connected component of $X'^{f'}_{a-\varepsilon}$ that the edge $\varphi(\sigma)$ is attached to.

Since the map φ is compatible with the dMfs and because it restricts to a CM equivalence $X^f_{a-\varepsilon} \to X'^{f'}_{a-\varepsilon}$, the dMf f attains the same minima and maxima on the two relevant connected component of $X^f_{a-\varepsilon}$ as f' does on its counterpart of $X'^{f'}_{a-\varepsilon}$ via φ. Since Definition 2.7 only considers which connected component the considered edge is attached to, it makes no difference for the isomorphism type of the induced Ml trees that in general σ is attached to said connected component of $X^f_{a-\varepsilon}$ at vertices that do not correspond via φ to the ones adjacent to $\varphi(\sigma)$ in $X'^{f'}_{a-\varepsilon}$. Thus, the construction of the induced generalized Ml tree produces nodes with the same chirality and label for both induced Ml trees in the steps that consider $\sigma, \varphi(\sigma)$, respectively. By assumption, the restriction $\varphi_{X^f_{a-\varepsilon}} : X^f_{a-\varepsilon} \to X'^{f'}_{a-\varepsilon}$ is a symmetry equivalence, so the isomorphism of Ml trees extends to the subtrees that correspond to the respective connected components. □

At the end of this section, we want to provide a different point of view on CM equivalences. As opposed to the case of symmetry equivalences, we cannot describe the action of the group of CM equivalences on the CM equivalence class of some dMf on a graph in terms of some group action on a space because CM equivalences change the space at hand.

Instead we propose to consider this operation as a digraph which has the CM equivalence class of a dMf on a graph (X, f) as vertices and elementary CM equivalences, i.e. ones that are either a symmetry equivalence of only one connected component or a CM equivalence of some level a, as edges. In the previous proof, we already used the fact that CM equivalences can be decomposed into a sequence of CM equivalences of separate levels a. Such CM equivalences of level a are determined by which edge e they reattach, that is which vertices the boundary vertices of e are swapped with. This allows us to order the outgoing edges at each vertex linearly:

We identify edges e with the ordered pair of the Morse labels of their boundary vertices (c_1, c_2) with the convention that the smaller label always comes first, that is $c_1 < c_2$. Let φ be a CM equivalence of level $f(e)$ that maps $(c_1, c_2) \mapsto (c'_1, c'_2)$. Then we label φ with the ordered tuple (c_1, c_2, c'_1, c'_2) and order the outgoing CM equivalences by level and among the same level by the lexicographical ordering on these labels. We define trivial CM equivalences, i.e., symmetry equivalences, to be less then non-trivial ones and order them by level and inside their level by minimal label on the connected component involved. For the same connected component at the same level, we define the order given by a lexicographic order on words which describe the symmetries similar to the case of CM equivalences. Thus, we can explore the CM equivalence class of any dMf on a graph with any standard exploration algorithm for any edge labeled digraph.

Remark 2.11. The mentioned point of view on CM equivalence classes can be phrased in a more categorical language: the described digraph encodes the data of a groupoid that describes the action of CM equivalences on their corresponding CM equivalence class.

3. Inverse Problem for Multigraphs

In this section we want to describe the relationship between dMfs on graphs, generalized Ml trees, generalized Mo trees, and generalized merge trees. The results are summarized in Figure 1.

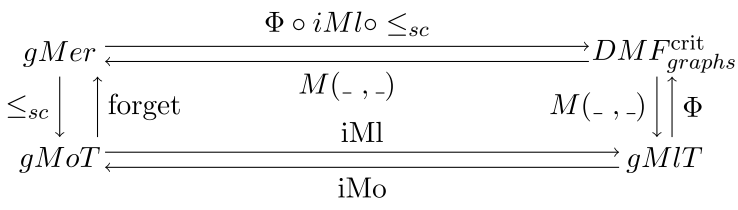


Figure 1. Relationships between dMfs and merge trees

THEOREM 3.1. *Let DMF^{crit}_{graphs} denote the set of CM equivalence classes of dMfs with only critical cells on multigraphs. Let $gMlT$ denote the set of isomorphism classes of gML trees. Then the induced dMf Φ, Definition 2.8, and the induced Morse labeled merge tree $M(\_\,,\_)$, Definition 2.7, define maps $M(\_\,,\_) \colon DMF^{crit}_{graphs} \leftrightarrow gMlT \colon \Phi$ that are inverse of each other in the sense that:*

(1) for any dMf (X, f) with only critical cells, the dMf $\Phi(M(X, f), \lambda_f)$ is CM-equivalent to (X, f), and

(2) for any gML tree (T, λ), we have $M(\Phi T, f_\lambda) \cong (T, \lambda)$.

PROOF. (1) Let (X, f) be a dMf with only critical cells on a graph X. We construct a CM equivalence $\varphi(X, f) \to \Phi(M(X, f))$ as follows: First we consider the spanning trees induced by (X, f) and $(\Phi(M(X, f)), f_{\lambda_f})$ and show that they are CM equivalent. Then we define φ on the closing edges and prove that φ is a CM equivalence.

By application of Brüggemann 2022, Theorem 5.6 we have a CM equivalence $\tilde{\varphi} \colon (\bar{X}, \bar{f}) \to (\Phi(M(\bar{X}, \bar{f})), \bar{f}_{\lambda_{\bar{f}}})$. We extend $\tilde{\varphi}$ to a CM equivalence $\varphi \colon (X, f) \to \Phi(M(X, f))$ by mapping each closing edge $\sigma \in X$ such that $f(\sigma) = a$ to the unique edge $\sigma' \in \Phi(M(X, f))$ with $f_{\lambda_f}(\sigma') = a$. The edge $\sigma' \in \Phi(M(X, f))$ is closing because a does not appear as a label on $(M(\bar{X}, \bar{f})), \lambda_{\bar{f}}) \cong (\bar{M}(X, f), \bar{\lambda}_f)$ since a is the value of the closing critical edge $\sigma \in X$. Furthermore, the connected component of $X_{a-\varepsilon}$ that σ is attached to corresponds to the subtree of $M(X, f)$ that consists of all descendants of $M(\sigma)$. By Definition 2.8, the edge σ' is attached to the connected component of $\Phi(M(X, f)_{a-\varepsilon}$ that corresponds to said subtree. It follows that φ is a CM equivalence.

(2) Let (T, λ) be a gML tree. Let $c_0 < c_1 < \cdots < c_n$ be the critical values of f_λ and let $\sigma_i \in \Phi T$ such that $f_\lambda(\sigma_i) = c_i$. We recall that the induced merge tree M defines in particular a bijection between the critical cells of ΦT and the nodes of $M(\Phi T, f_\lambda)$. For any cell $\sigma \in \Phi T$, we recall that we denote the node of $M(\Phi T, f_\lambda)$ that corresponds to σ by $M(\sigma)$. We also recall that Φ, as constructed in Brüggemann 2022, Definition 3.21, comes with a bijcetion that we extended to cycle nodes in Definition 2.8 $\phi \colon V(T) \to \Phi T$. An isomorphism $(\varphi, \mathrm{id}_{\mathbb{R}}) \colon (T, \lambda) \to M(\Phi T, f_\lambda)$ is given by $\varphi := M \circ \phi^{-1}$. It is immediate that φ is a bijection because M and ϕ are. Furthermore, φ is by construction compatible with the respective Morse labelings. It is only left to show that φ is compatible with the chiral child relation and the respective roots.

Consider $\sigma_n \in \Phi T$. For both trees, the cell σ_n corresponds to the root of the respective tree. In $M(\Phi T, f_\lambda)$ this is the case because f_λ attains its maximum on σ_n. In (T, λ) this holds because $\phi(\sigma_n)$ holds the maximal Morse label $\lambda(\phi(\sigma_n)) = c_n$. Thus, the map φ maps the root of (T, λ) to the root of $M(\Phi T, f_\lambda)$.

For each critical edge $\sigma_i \in \Phi T$ we have one of the two cases:

a) σ_i is closing, or
b) σ_i is not closing.

For case b), the proof is identical to the proof of case (2) of Brüggemann 2022, Theorem 5.4. For case a), let σ_i be a closing critical edge. In this case, the compatibility with the chiral child relation follows directly by case 1 of Definition 2.7 and the property that only children of generalized merge trees need to have the same chirality as their parent node.

□

COROLLARY 3.1. Since the bijection from Theorem 3.1 is compatible with the Morse labels, it induces a bijection $M(\_,\ \_) \colon DMF^{crit}_{graphs}/_{\leq} \leftrightarrow MlT/_{\leq} \colon \Phi$ where $/_{\leq}$ denotes dividing by order equivalence.

3. INVERSE PROBLEM FOR MULTIGRAPHS

DEFINITION 3.1. Let $\leq$ and $\leq'$ be two Morse orders on a generalized merge tree T. A merge equivalence $(T, \leq) \to (T, \leq')$ of Mo trees is a self-bijection $\psi\colon V(T) \cong V(T)$ such that

(1) for each inner node a of T, the node a is the maximum of a subtree T' of T with respect to $\leq$ if and only if $\psi(a)$ is the maximum of T' with respect to $\leq'$, and
(2) for each leaf a of T, the node a is the minimum of a subtree T' of T with respect to $\leq$ if and only if $\psi(a)$ is the minimum of T' with respect to $\leq'$.

We call $\leq$ and $\leq'$ *merge equivalent* if there exists a merge equivalence $(T, \leq) \to (T, \leq')$. A merge equivalence $(T, \leq) \to (T', \leq')$ between different gMo trees is a concatenation of an isomorphism $\varphi\colon T \to T'$ of underlying generalized merge trees and a merge equivalence $(T, \leq) \xrightarrow{\psi} (T, \varphi^* \leq') \xrightarrow{\varphi} (T', \leq')$.

PROPOSITION 3.1. *Any two Morse orders $\leq$ and $\leq'$ on a generalized merge tree T are merge equivalent.*

PROOF. The statement is proved inductively. Let a be the minimal leaf of a subtree T' of T with respect to $\leq$. Then a needs to be the minimal leaf of T' with respect to $\leq'$ because otherwise $\leq'$ would fail to be a Morse order due to Remark 2.4. The statement for inner nodes follows similarly. □

COROLLARY 3.2. Two generalized Mo trees have isomorphic underlying generalized merge trees if and only if they are merge equivalent. In particular, two (not generalized) Mo trees have isomorphic underlying (not generalized) merge trees if and only if they are merge equivalent.

For any generalized merge tree T, there are several ways to induce canonical Morse orders on T. We introduce the sublevel-connected Morse order (generalization of Brüggemann 2022, Definition 4.1) on any given generalized merge tree in the following:

To define the sublevel-connected Morse order, we first observe that every node a of T is uniquely determined by the shortest path from the root to a. We recall that the depth of T is the maximal length of any path in T that appears as the shortest path from the root to a leaf. Because T is chiral, we can identify such shortest paths with certain words:

DEFINITION 3.2 (Brüggemann 2022, Definition 3.1). Let T be a generalized merge tree of depth n and let a be a node of T. The *path word* corresponding to a is a word $a_0 a_1 \ldots a_n \in \{L, R, \_\}^{n+1}$ where $\_$ denotes the empty letter. If a is of depth k, the letters $a_0 \ldots a_k$ are given by the chirality of the nodes belonging to the shortest directed path from the root to a. The letters $a_{k+1} \ldots a_n$ are then empty.

REMARK 3.1. Let a, b be nodes of a generalized merge tree T and let $a_0 a_1 \ldots a_n$ be the path word corresponding to a and $b_0 b_1 \ldots b_n$ be the path word corresponding to b. Then the equation $a_0 = b_0 = L$ always holds because we consider paths that begin at the root. Because $a_0 = b_0 = L$ and because we consider finite trees, there is always a maximal $k \in \mathbb{N}$ such that $a_i = b_i$ holds for all $i \leq k$. Furthermore, the last non-empty letter of a path word is always the chirality of the considered node.

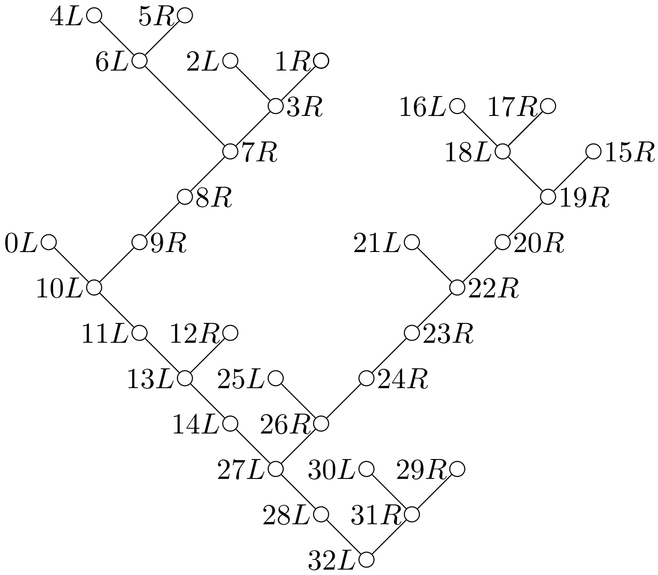
DEFINITION 3.3 (Brüggemann 2022, Definition 4.1). Let T be a generalized merge tree. We define the *sublevel-connected Morse order* $\leq_{sc}$ on the nodes of T as follows:

Let a, b be arbitrary nodes of T. Let $a_0 a_1 \ldots a_n$ be the path word corresponding to a and $b_0 b_1 \ldots b_n$ the path word corresponding to b (see Definition 3.2). Furthermore, let $k \in \mathbb{N}$ be maximal such that $a_i = b_i$ for all $i \leq k$. If $a_k = b_k = L/R$ we define $a \leq_{sc} b$ if and only if one of the following cases hold:

a) $a_{k+1} = L$ and $b_{k+1} = R$/$a_{k+1} = R$ and $b_{k+1} = L$
b) $b_{k+1} = \_$
c) $a = b$

3. INVERSE PROBLEM FOR MULTIGRAPHS

EXAMPLE 3.1. We depict the sublevel-connected Morse order in the following example:



PROPOSITION 3.2. *The construction of the sublevel-connected Morse order and forgetting the Morse order defines a pair of inverse bijections*

$$\leq_{sc} \colon Mer/_{\cong} \longleftrightarrow gMoT/_{\sim} \colon forget$$

where $\sim$ *denotes merge equivalence.*

PROOF. The statement follows directly by Corollary 3.2. □

To summarize our results of this section, we take a look at how Proposition 2.1, Theorem 3.1, and Proposition 3.2 turn the different maps from Figure 1 into bijections by dividing out the needed notion of equivalence. If we do not divide out any equivalence relation, the map Φ is not even well-defined. The maps $M(\_\,,\_)$, iMo, and forget are surjective, but not injective. The maps $\leq_{sc}$ and iMl are injective but not surjective.

Identifying CM-equivalent dMfs makes Φ a well-defined map and, moreover, a bijection which is inverse to $M(\_\,,\_)\colon DMF^{crit}_{graphs} \to gMlT$ by Theorem 3.1. Inverting order equivalences turns iMo and iMl into inverse bijections. Finally, inverting merge equivalences makes $\leq_{sc}$ and forget inverse to each other. As a consequence, we have a complete description of the inverse problem for critical dMfs on multigraphs and their induced merge trees. The characterization for arbitrary dMfs on 1-dim regular CW complexes follows by collapsing matched cells and then applying a version of Theorem 3.1 that incorporates Remark 2.9. However, this procedure secretly makes use of two features which might become problematic if one tries to generalize the result to higher dimensions: on one hand, we use that irregularities of attaching maps can be easily characterized in the 1-dimensional case: here they always produce self-loops. Dealing with irregular faces in higher dimensions would be more difficult

On the other hand, even if we start with regular CW complexes, the complex that arises by performing the simple collapses described by a Morse matching is not arbitrary but subject to being simple homotopy equivalent to a regular CW complex. It is a feature of dimension one that all 1-dimensional CW complexes are simple homotopy equivalent to a 1-dimensional regular CW complex. Hence, defining CM equivalences becomes more difficult in a higher-dimensional setting, in particular, if one wants to work with non-critical dMfs. This would lead to the need to analyze which CW complexes are simple homotopy equivalent to regular CW complexes in order to know for which generality a notion of CM equivalence is needed.

3. INVERSE PROBLEM FOR MULTIGRAPHS

4. Realization Problem with Simple Graphs

Let T be a generalized merge tree. Recall that $C(T) = C$ denotes the set of all cycle nodes of T. For any $c \in C$, let c_u denote the unique child of c. For any $v \in T$, let $T(v)$ denote the subtree of T with root v and let $\ell(v)$ denote the number of leafs of $T(v)$.

THEOREM 4.1. *Let T be a generalized merge tree. Then there exists a simple graph X and dMf $f\colon X \to \mathbb{R}$ such that $M(X, f) = T$ if and only if for every $c \in C(T)$,*

$$|C(T(c_u))| < \frac{(\ell(c_u) - 2)(\ell(c_u) - 1)}{2}.$$

Furthermore, X can be made planar if and only if

$$|C(T(c_u))| < 2 \cdot \ell(c_u) - 5.$$

PROOF. Suppose there exists a simple graph X and dMf $f\colon X \to \mathbb{R}$ such that $M(X, f) = T$, and suppose by contradiction that there is a $c \in C(T)$ with the property that

$$|C(T(c_u))| \geq \frac{(\ell(c_u) - 2)(\ell(c_u) - 1)}{2}.$$

By Lemma 2.1, the rooted subtree $T(c_u)$ is isomorphic as rooted binary trees to the induced Morse labeled merge tree of $X_{f(s)}[s]$ where s is the simplex of X such that $M(s) = c_u$. Letting v be the number of vertices in $X_{f(s)}[s]$, e the number of edges in $X_{f(s)}[s]$, and b_1 the number of cycles in $X_{f(s)}[s]$, we see that

$$\begin{aligned} e &= v - 1 + b_1 \\ &\geq v - 1 + \frac{(v-1)(v-2)}{2} \\ &= v - 1 + \frac{v(v-1)}{2} + 1 - v \\ &= \frac{v(v-1)}{2} \end{aligned}$$

which is the maximum number of edges any connected component can have. Hence it is impossible to add a cycle to this connected component so that

$$|C(T(c_u))| < \frac{(\ell(c_u) - 2)(\ell(c_u) - 1)}{2}.$$

for all $c \in C$. Now suppose further that X is planar, and suppose by contradiction that $|C(T(c_u))| \geq 2 \cdot \ell(c_u) - 5$. Using the same notation as above, we have

$$\begin{aligned} e &= v - 1 + b_1 \\ &\geq v - 1 + 2v - 5 \\ &= 3v - 6. \end{aligned}$$

But it is well known that a simple planar graph satisfies $e \leq 3v - 6$ Bickle 2020, Theorem 5.9. Hence either $X_{f(s)}[s]$ is not planar or maximal planar in the case of equality. In either case, another edge cannot be added to $X_{f(s)}[s]$ without breaking planarity, and thus the result.

For the other direction, given the generalized Merge tree T, construct the sublevel-connected Morse order $\leq_{sc}$ (Definition 3.3) on the nodes of T. Associate to this Morse order a Morse labeling $\lambda\colon T \to \mathbb{R}$ such that $a \leq_{sc} b$ if and only if $\lambda(a) \leq \lambda(b)$. Apply the construction in Definition 2.8 to (T, λ) to obtain the underlying merge tree $(\overline{T}, \overline{\lambda})$. By Brüggemann 2022, Theorem 6.5, there is a path P and dMf $\overline{f}\colon P \to \mathbb{R}$ such that $M(P, \overline{f}) = (\overline{T}, \overline{\lambda})$. We will inductively attach edges to P in one-to-one correspondence with cycle nodes of T. Each edge will be labeled with the same label as its corresponding cycle node.

Induce on the cycle nodes of T with respect to the sublevel-connected Morse order $c_1 \leq_{sc} c_2 \leq_{sc} \cdots$. For the base case $i = 1$, write $P = X^1$. We have by hypothesis that

$$|C(T(c_{1u}))| < \frac{(\ell(c_{1u}) - 2)(\ell(c_{1u} - 1)}{2}.$$

In addition, $M(P, \overline{f}) = (\overline{T}, \overline{\lambda})$ so $c_{1u} = M(s_1)$ for some simplex $s_1 \in P = X^1$. Applying the correspondence noted in Remark 2.7, this inequality means that

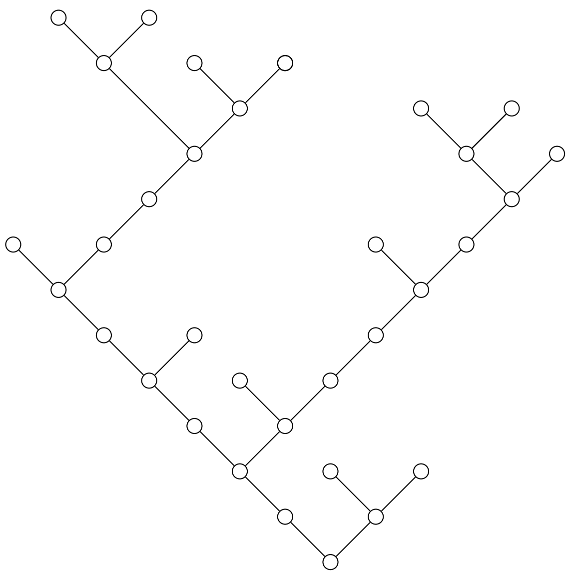
$$b_1(X^1[s_1])| < \frac{(v(X^1[s_1] - 2)(v(X^1[s_1] - 1)}{2}.$$

By the computation in the forward direction, this implies that $e(X^1[s_1]) < \frac{v(X^1[s_1])(v(X^1)-1)}{2}$. Hence there are at least two vertices in $X^1[s_1]$ not connected by an edge. A choice of vertex can be made by defining a lexicographic ordering on a subset of ordered pairs of the vertex set of P where an ordered pair (v, u) satisfies $\overline{f}(v) < \overline{f}(u)$ and $(v, u) < (v', u')$ if $\overline{f}(v) < \overline{f}(v')$ or $\overline{f}(u) < \overline{f}(u')$ when $\overline{f}(v) = \overline{f}(v')$. Since all the vertices of P are given distinct values, $<$ is a total order. Add an edge e_1 incident with the vertices in the minimum pair over all available pairs to create $X^2 = X^1 \cup \{e_1\}$ and extend $\overline{f}$ to $f^1(e_1) := \lambda(c_1)$. Then $M(X^2, f^1) \simeq (T_{\leq \lambda(c_1)}, \lambda|_{T_{\leq \lambda(c_1)}})$. The inductive step is identical to the base case.

Now suppose that $|C(T(c_u))| < 2 \cdot \ell(c_u) - 5$ for all cycle nodes $c \in T$. By the forward direction, this is equivalent to $e < 3v - 6$ in the corresponding sublevel complex of X. The method of construction is analogous to the above construction and utilizes the fact that if a planar simple graph satisfies $e < 3v - 6$, then it is not maximal planar and hence an edge can be added while maintaining planarity Bickle 2020, Corollary 5.11. □

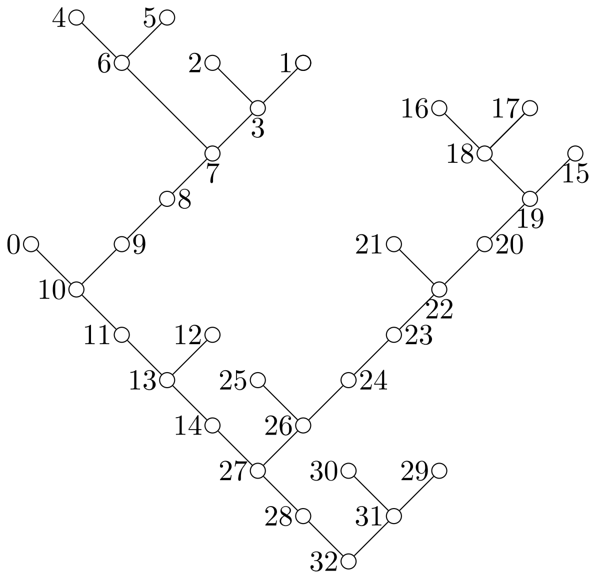
Remark 4.1. While the choices made in the construction of the simple graph X in Theorem 4.1 may be thought of as one canonical choice, the sublevel-connected Morse order is only one possible representative for the Morse order. Another just as natural (and shuffle equivalent[11]) order would be the index Morse order Brüggemann 2022, Definition 3.3. Furthermore, once a Morse order is picked, there are often several possible simple graphs with dMfs all related by CM equivalence that represent the given generalized merge tree.

Example 4.1. To illustrate the construction in the planar case, consider the generalized merge tree T pictured below:

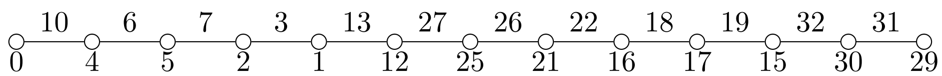


[11]That is, Morse orders that have the same restricted order on leafs as well as the same restricted order on inner nodes. See Brüggemann 2022, Definition 2.24 for details.

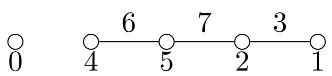
We constructed the sublevel-connected Morse order and induced Morse labeling λ in Example 3.1.



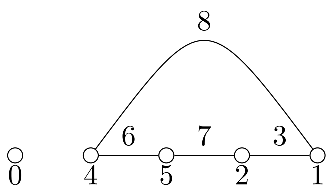
We then pass to the underlying merge tree $\overline{T}$ and restrict λ to $\overline{T}$ in order to apply Brüggemann 2022, Definition 4.5 to obtain the sublevel-connected dMf on the graph below with induced merge tree $\overline{T}$.



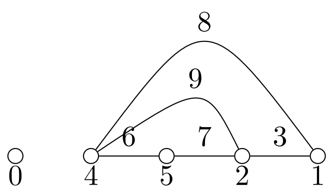
We induce on the cycle nodes ordered by their generalized Morse label. The first cycle to be introduced is cycle node with label 8. This will be a cycle added to the graph



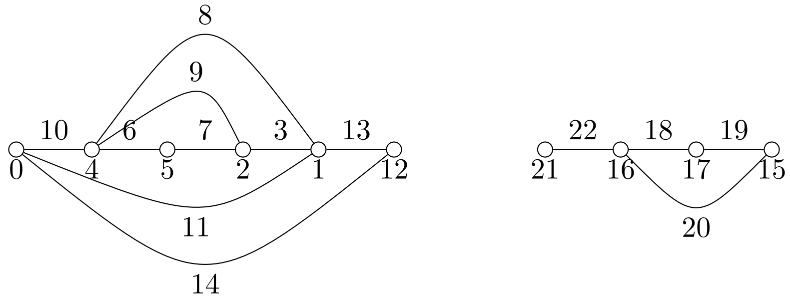
to the component with the edge labeled 7.



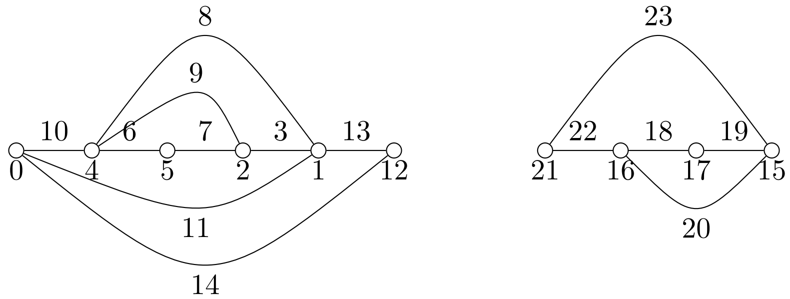
We then add the cycle corresponding to the node labeled 9 to this same graph.



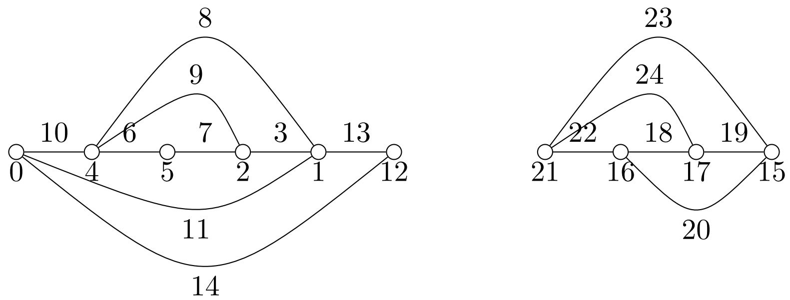
Skipping to the cycle node labeled 23, we see that we need to add a cycle to the component with edge labeled 22:



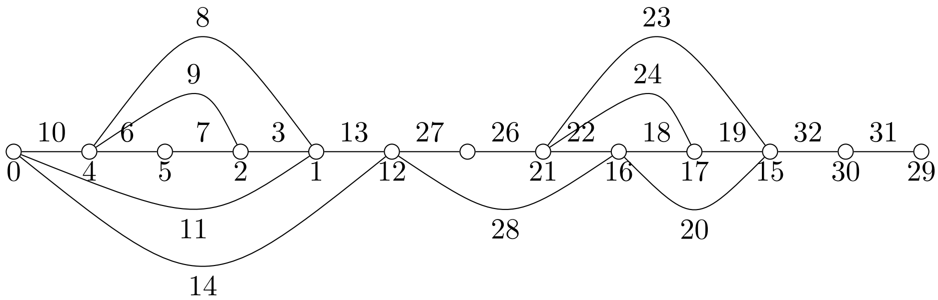
We add this edge



and must add another cycle corresponding to cycle node labeled 24 to this same connected component.



Notice that this component is now a complete graph and that no more cycles can be added. The final graph with dMf that induces the given generalized merge tree is



5. How to Find Cancellations with Merge Trees

In this section, we present a way to find cancellations of critical cells of dMfs with the help of the induced merge tree. The idea is to start with an arbitrary dMf that only has critical cells and to perform cancellations along the merge tree.

5. HOW TO FIND CANCELLATIONS WITH MERGE TREES

Remark 5.1. In order to obtain an arbitrary dMf on a graph X that has only critical cells, one can simply choose any total order on the vertices and any total order on the edges. Then assign the values $0, \dots, |V(X)| - 1$ to the vertices according to the chosen order and the numbers $|V(X)|, \dots, |V(X)| + |E(X)|$ to the edges. This always produces an index-ordered dMf which is not necessary for the following algorithm. Perhaps more sophisticated approaches to finding a critical dMf might be useful, but for now we are satisfied with this simple one.

Given a critical dMf $f\colon X \to \mathbb{R}$, the algorithm proceeds as follows:

(1) Calculate the induced generalized Morse labeled merge tree $M(X, f)$, and let C be the set of leaves of $M(X, f)$

(2) If $C = \emptyset$, end the algorithm. Otherwise, let $c \in C$ be the vertex with maximal label, and let p be the youngest ancestor of c such that p is neither a cycle node nor matched. Then either:
a) The vertex $M^{-1}(c)$ is adjacent to the edge $M^{-1}(p)$
b) The vertex $M^{-1}(c)$ is not adjacent to the edge $M^{-1}(p)$.
If case a), match $M^{-1}(c)$ and $M^{-1}(p)$. This does not produce cycles because we explicitly exclude cycle nodes from the matching. Let $C = C - \{c\}$ and return to (2)

If case b), either:

i) leave $M^{-1}(c)$ critical, let $C = C - \{c\}$ and return to (2)

ii) check for a symmetry equivalence a of (X, f) such that $a(M^{-1}(c))$ is adjacent to $a(M^{-1}(p))$, apply it, and then proceed as in case a). If there is no symmetry equivalence, proceed to i), iii), or iv).

iii) apply a CM equivalence in order to make $M^{-1}(c)$ and $M^{-1}(p)$ adjacent, then proceed as in case a), or

iv) observe that there is a unique gradient flow line from $M^{-1}(c)$ to $M^{-1}(p)$ and cancel the two cells along this flow line. Let $C = C - \{c\}$ and return to (2).

The precise nature of the output depends on the choices the user makes in case b). If case b) never applies, the output will be an optimal discrete Morse function on the exact same graph X. In the case that b) is applied and the user chooses option i), an optimal matching is not guaranteed but we preserve the homeomorphism type of X. If case iii) is consistently chosen, we produce an optimal matching but may change the homeomorphism type of X. If case iv) is consistently chosen, we preserve the homeomorphism type of X and obtain an optimal matching but we change the order of the vertices induced by f on a larger scale. While one could in principle choose different options of case b) at different stages in a single run of the algorithm, this would produce a seemingly undesirable output, as it would suffer all the drawbacks mentioned in each case.

Most of the claims made in the above algorithm are straightforward to prove. For example, the fact that the cases 2a), 2b)i), and 2b)iii) work as described follows immediately from the definition of the used equivalences. However in general it does not appear easy to decide whether case 2b)ii) is applicable. Nonetheless, case 2b)iv) is not so obvious, so we consider it in the following lemma:

Lemma 5.1. *Let X be a graph, $f\colon X \to \mathbb{R}$ a critical dMf, and $M(X, f)$ the induced gML tree. At any point of the cancellation algorithm, there is always a unique gradient flow line from the vertex $M^{-1}(c)$ corresponding to the maximally labeled unmatched leaf c to the edge $M^{-1}(c)$ corresponding to its youngest unmatched ancestor p.*

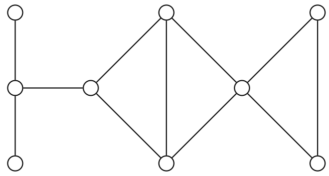
Proof. If $M^{-1}(c)$ and $M^{-1}(p)$ are adjacent, there is nothing to prove. If $M^{-1}(c)$ and $M^{-1}(p)$ are not adjacent then there is no other non-closing critical edge in $X_{f(M^{-1}(p)-\varepsilon)}[M^{-1}(c)]$ because otherwise said other younger critical edge would induce a younger unmatched ancestor of c.

Since $M^{-1}(c)$ is a critical vertex with no adjacent critical edge, all adjacent edges of $M^{-1}(c)$ are matched with their respective other vertex. This means that on all adjacent edges, there is a gradient flow line pointing towards $M^{-1}(c)$. Following these gradient flow lines backwards either leads to matched vertices that are adjacent only to the edge they are matched with, or to the unique critical edge of $X_{f(M^{-1}(p)-\varepsilon)}[M^{-1}(c)]$. One of the flow lines eventually leads to $M^{-1}(p)$ because $X_{f(M^{-1}(p)-\varepsilon)}[M^{-1}(c)]$ is connected.

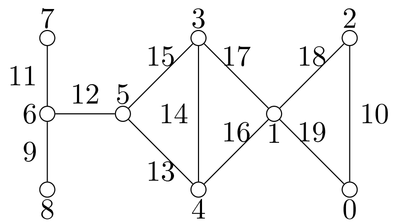
The flow line is unique because closing edges remain critical, that is, because we only match cells along a subtree of X. □

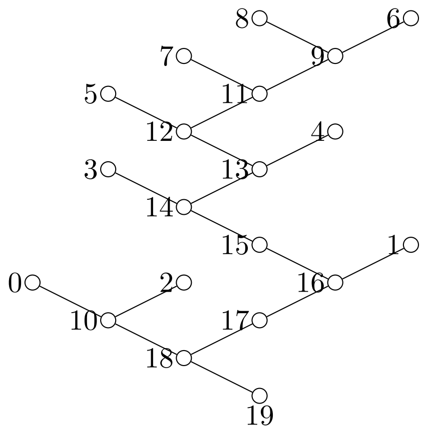
We apply the cancellation algorithm to the following example:

Example 5.1. We consider the graph:

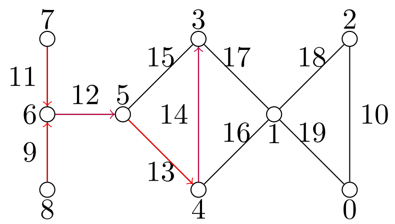


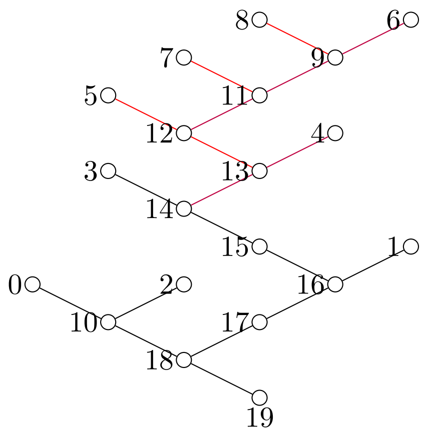
We put some critical dMf on it and calculate the induced generalized merge tree:



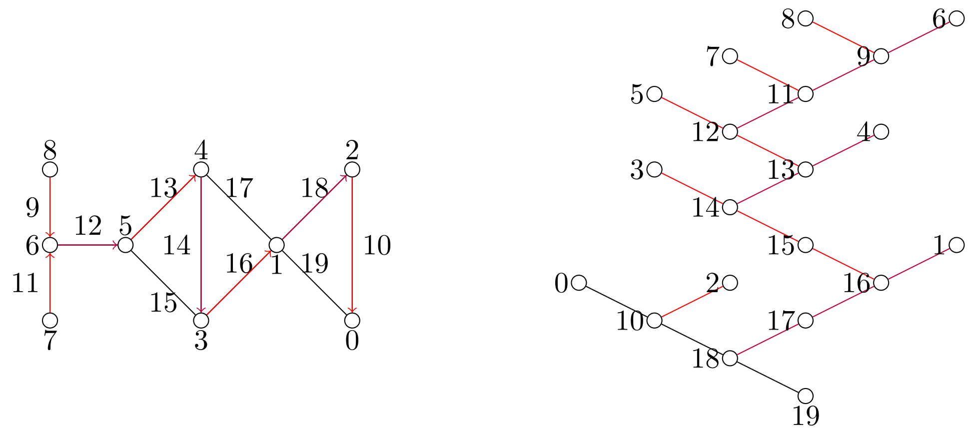


We apply step 2a) as long as possible:





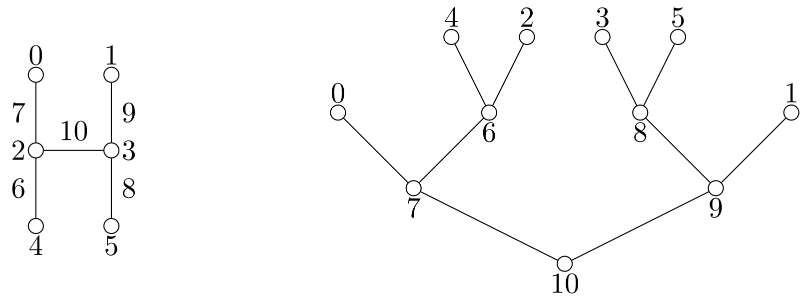
Now is the first time we run into case 2b). We can actually apply case 2b)ii) here:



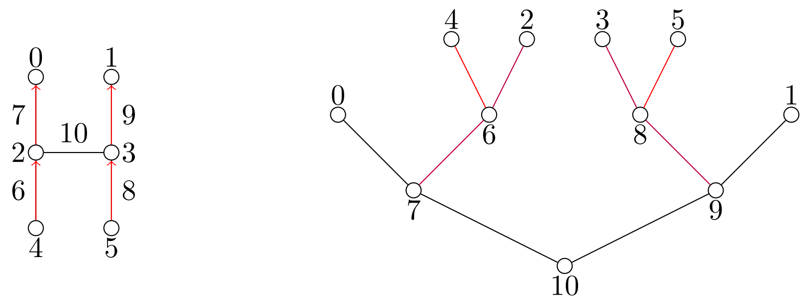
In this example, the cases 2a) and 2b)ii) sufficed.

We consider the following example in order to see how quickly things can fail:

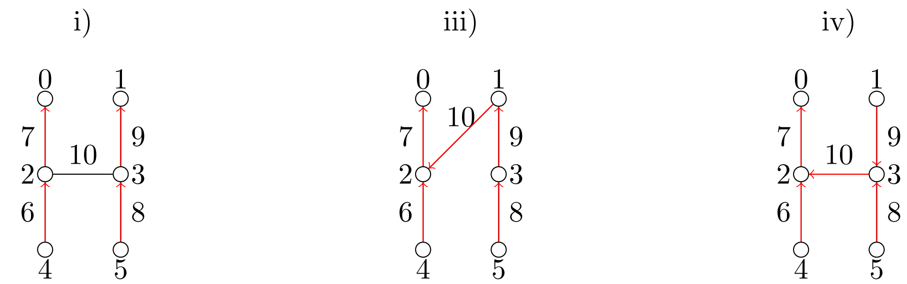
Example 5.2. We consider the following dMf and its induced merge tree:



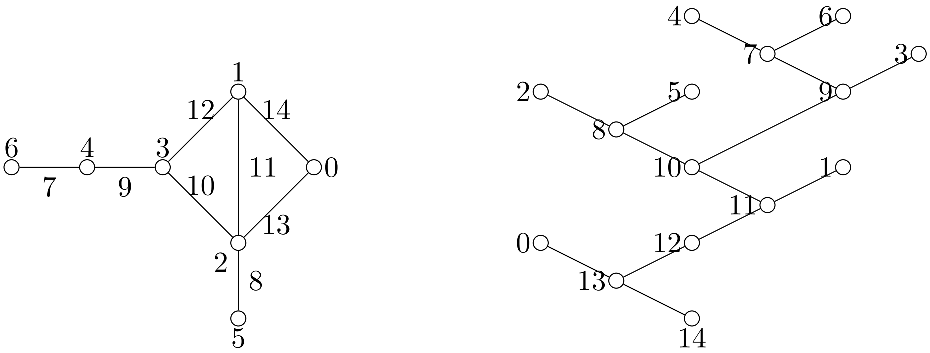
After twofold application of step 2a), we have the following:



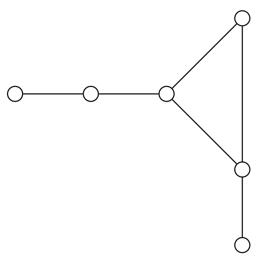
Now we have reached case 2b) and case 2b)ii) is not applicable. We would need to have the vertex labeled 1 adjacent to the edge labeled 10. But this is not possible because all symmetry equivalences leave the vertex labeled 1 adjacent to the edge labeled 9 and no other edge. The three different solutions result in the following:



Example 5.3. A sublevel symmetry of the last sublevel complex before the "merge tree algorithm" fails may not always be sufficient. Consider the graph with dMf given below.

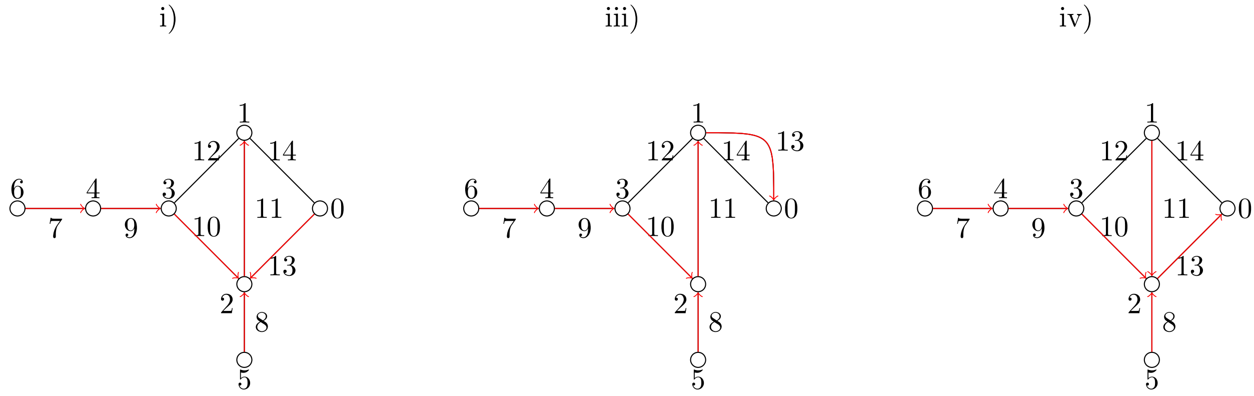


Proceeding as before, we obtain a matching on the graph until the algorithm specifies to match the vertex labeled 1 with the edge labeled 13. Since these cells are not incident, we need to find a sublevel-symmetry of sublevel 12. However, the sublevel subcomplex X_{12} is given by



which is well-known to have no non-trivial automorphisms. There is also no symmetry equivalence of a lower level than 12 that makes the vertex labeled 1 and the edge labeled 13 adjacent. However, the three different workarounds mentioned earlier result in the following:

5. HOW TO FIND CANCELLATIONS WITH MERGE TREES



At the end of this section, we compare our algorithm for finding cancellations of critical cells to similar algorithms from the literature. In Lewiner, Lopes, and Tavares 2003b the authors introduce an algorithm to find optimal dMfs on 2-dimensional manifolds which they generalize to higher dimensions and more general complexes in Lewiner, Lopes, and Tavares 2003a, even though losing the guarantee for optimality in the process. The main similarity to our approach is the use of an auxiliary tree structure, in our case the generalized merge tree, in the case of Lewiner, Lopes, and Tavares 2003a a spanning hyperforest of a hypergraph associated to the Hasse diagram of a dMf.

In Rand and Scoville 2020, the authors provide an algorithm to find optimal dMfs on trees. Said algorithm, combined with any standard algorithm to find spanning trees, can easily be generalized to provide optimal dMfs on graphs with a prescribed critical vertex.

The main feature of our new approach, compared to the pre-existing ones, seems to be that our algorithm allows to preserve certain properties of a given dMf. In certain cases, such a dMf might be given by an application and, therefore, might be worth preserving. We conjecture that, given a suitable version of higher merge trees, our algorithm can be generalized to higher dimensions. Since finding optimal Morse matchings is MAXâ€“SNP hard, such a generalization might either fail to be optimal or be inconvenient to work with in practice. Nonetheless, we hope to find interesting classes of examples in which such a generalized algorithm happens to be performative and informative.

6. Future Directions

In this section, we want to take a look at possible applications and further directions this work might lead to.

Our main results, Theorem 3.1 and Theorem 4.1, give a detailed description of the fiber of the persistence map that takes dMfs on graphs to their persistent connectivity. This approach may be used in applications in which the persistent connectivity turns out to be the most relevant feature, allowing to replace a maybe inconvenient graph with a more convenient one that describes that same persistent connectivity. At the end of Section 2 we sketch how to search through all possible representatives in a structured way. Theorem 4.1 provides an easy-to-check condition for when this replacement can be chosen to be a simple graph.

A similar approach is given by applying the cancellation algorithm from Section 5. The algorithm helps to simplify dMfs on graphs while allowing to preserve either the homeomorphism type or the dynamics induced by the Morse function. One immediate question would be, how much approaches such as these change the original Morse function. Thus, it seems interesting to investigate the diameter of the set of representatives for a given merge class of dMfs with respect to some suitable metric for dMfs. Moreover, we would be interested in finding out how distant the function coming from the cancellation algorithm is from its input function.

In a more pure direction, one could try to set up a version of persistent geometric group theory using the groups of sublevel automorphisms as in Example 2.4. On one hand, it seems interesting in itself to consider actions of sequences of groups on sequences of spaces and which ones can be realized as sublevel automorphisms of a filtered space. On the other hand, the results from such approaches would be useful for applications of the cancellation algorithm mentioned above. It also seems interesting to analyze how the application of CM equivalences affects the sublevel automorphism group.

Furthermore, the notions of symmetry equivalences and CM equivalences might be helpful for the investigation of the space of dMfs on some given graph.

The most straightforward direction would be a generalization of Definition 2.7 to higher dimensions in order to enable the pursuit of all the above mentioned possible future directions in higher dimensions.

7. Conflict of Interest Statement

The authors state that there is no conflict of interest.

8. Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Bibliography

Benedetti, Bruno (2016). "Smoothing discrete Morse theory". In: *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 16, pp. 335–368.

Bickle, Allan (2020). *Fundamentals of graph theory*. Vol. 43. Pure and Applied Undergraduate Texts. American Mathematical Society, Providence, RI, pp. xv+336. ISBN: 978-1-4704-5342-8.

Brüggemann, Julian (Nov. 2022). "On Merge Trees and Discrete Morse Functions on Paths and Trees". In: *J Appl. and Comput. Topology*. DOI: https://doi.org/10.1007/s41468-022-00101-w.

Brüggemann, Julian and Nicholas A. Scoville (2023). *On cycles and merge trees*. arXiv: 2301.01316 [math.AT].

Cardona, Robert et al. (2022). "The universal ℓ^p-metric on merge trees". In: *38th International Symposium on Computational Geometry*. Vol. 224. LIPIcs. Leibniz Int. Proc. Inform. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, Art. No. 24, 20. DOI: 10.4230/lipics.socg.2022.24.

Carlsson, Gunnar and Mikael Vejdemo-Johansson (2022). *Topological data analysis with applications*. Cambridge University Press, Cambridge, pp. xi+220. ISBN: 978-1-108-83865-8. DOI: 10.1017/9781108975704.

Carr, Hamish, Jack Snoeyink, and Ulrike Axen (2003). "Computing contour trees in all dimensions". In: *Computational Geometry: Theory and Applications*, pp. 75–94.

Curry, Justin (Feb. 2019). "The fiber of the persistence map for functions on the interval". In: *J Appl. and Comput. Topology*, pp. 301–321.

Curry, Justin et al. (2022). "Decorated merge trees for persistent topology". In: *J. Appl. Comput. Topol.* 6.3, pp. 371–428. ISSN: 2367-1726. DOI: 10.1007/s41468-022-00089-3.

Forman, Robin (Mar. 1998). "Morse theory for cell complexes". In: *Adv. Math.* 134.1, pp. 90–145. ISSN: 0001-8708. DOI: 10.1006/aima.1997.1650.

— (2002). "A user's guide to discrete Morse theory". In: *Sém. Lothar. Combin.* 48, Art. B48c, 35.

Gasparovic, Ellen et al. (2022). "Intrinsic Interleaving Distance for Merge Trees". In: *trees* 38.37, p. 32.

Johnson, Benjamin and Nicholas A. Scoville (July 2022). "Merge trees in discrete Morse theory". In: *Res. Math. Sci.* 9, Paper No. 49.

Lewiner, Thomas, Hélio Lopes, and Geovan Tavares (2003a). "Optimal discrete Morse functions for 2-manifolds". In: *Computational Geometry* 26.3, pp. 221–233. ISSN: 0925-7721. DOI: https://doi.org/10.1016/S0925-7721(03)00014-2.

— (2003b). "Toward Optimality in Discrete Morse Theory". In: *Experimental Mathematics* 12.3, pp. 271–285. DOI: 10.1080/10586458.2003.10504498.

Morozov, Dmitriy, Kenes Beketayev, and Gunther Weber (2013). "Interleaving distance between merge trees". In: *Discrete and Computational Geometry* 49.22-45, p. 52.

Oudot, Steve Y. (2015). *Persistence theory: from quiver representations to data analysis*. Vol. 209. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, pp. viii+218. ISBN: 978-1-4704-2545-6. DOI: 10.1090/surv/209.

Polterovich, Leonid et al. (2020). *Topological persistence in geometry and analysis*. Vol. 74. University Lecture Series. American Mathematical Society, Providence, RI, pp. xi+128. ISBN: 978-1-4704-5495-1.

Rand, Ian and Nicholas A. Scoville (2020). "Discrete Morse functions, vector fields, and homological sequences on trees". In: *Involve* 13.2, pp. 219–229. DOI: DOI:10.2140/involve.2020.13.219.

CHAPTER IV

On the Parameter Space of Discrete Morse Functions

Julian Brüggemann

This chapter is available as a preprint under Brüggemann 2023.

Abstract. In this work, we give a combinatorial-geometric model for parameter spaces of discrete Morse functions on CW complexes. We investigate properties of parameter spaces of discrete Morse functions and their relation to the spaces of smooth Morse functions on manifolds, as well as spaces of discrete Morse matchings on CW complexes. Moreover, we provide a similar model for the parameter space of merge trees and realise the induced merge tree and the induced barcode as continuous maps between parameter spaces.

1. Introduction

Discrete Morse theory, originally introduced by Forman Forman 1998, is a powerful framework for investigating the simple homotopy type and other properties of CW complexes. It has shown potential to be useful for the study of persistent invariants of filtered CW complexes as they appear in topological data analysis. Other than that, discrete Morse theory has numerous applications in both, pure and applied mathematics.

The idea of using spaces to parametrize mathematical object originates in the concept of moduli spaces as they are used in topology and algebraic geometry. The general idea of parameter spaces is to provide a topological space $\mathcal{M}$, together with a convenient way to parametrize $\mathcal{M}$, such that points of $\mathcal{M}$ correspond to certain mathematical objects one wants to investigate, and geometric features of $\mathcal{M}$ reflect relevant properties of said objects under investigations.

We choose the framework of hyperplane arrangements in vector spaces of discrete functions for an approach to discrete Morse functions on the level of parameter spaces.

We observe that Formans definition of discrete Morse functions Definition 3.2 gives rise to the Definition 3.9 of the Morse arrangement $\mathcal{A}(X)$, a hyperplane arrangement that subdivides the vector space of discrete functions $\mathbb{R}^X$ on a CW complex X into regions, some of which belong to Forman-equivalence classes of discrete Morse functions, and others do not contain discrete Morse functions at all. This way, the introduced parameter space of discrete Morse functions allows a parametrization indexed over the face poset $D(X)$, which helps us to investigate properties of discrete Morse functions using the combinatorial properties of the Morse arrangement $\mathcal{A}$.

In this work, we want to investigate discrete Morse theory and its connection to certain other concepts in mathematics from the viewpoint of parameter spaces.

The concepts we want to relate the parameter spaces of discrete Morse funcitons to are

(1) The parameter spaces of smooth Morse functions on manifolds.
(2) The parameter spaces of discrete Morse matchings on regular CW complexes.
(3) The parameter space of merge trees.
(4) The parameter space of barcodes.

Similar work has been done in Catanzaro et al. 2020, Leygonie and Tillmann 2022, and Cyranka, Mischaikow, and Weibel 2020. Whereas Catanzaro et al. 2020 explores different notions of equicalences of smooth Morse functions on the sphere and the corresponding moduli spaces, both Leygonie and Tillmann 2022 and Cyranka, Mischaikow, and Weibel 2020 investigate fibers of two different instances of persistence maps.

The parameter spaces of Morse functions on manifolds have been introduced in Cerf 1970 and further investigated in Hatcher and Wagoner 1973 as subspaces of the spaces of smooth functions $C^\infty(M)$ on compact smooth manifolds M, governed by a stratification given by a certain notion of regularity for critical points and critical values. The motivation to investigate Morse functions in this way was to give a solid framework for the investigation of the pseudo-isotopoy versus isotopy question and its implications for the h-cobordism theorem. Cerf already found a map that relates certain subspaces of the space of Morse functions to spaces of discrete Morse functions Proposition 4.2 decades before discrete Morse functions had been invented by Forman. We extend this map to path components:

Theorem 4.7.
It follows from Theorem 4.5 that Cerf's map from Proposition 4.2 extends to a map $\eta\colon \mathcal{N} \to \mathcal{M}(\mathcal{M}_{\mathcal{N}}$, *where* $M_{\mathcal{N}}$ *denotes the CW decomposition of* M *induced by any Morse function in the path component* $\mathcal{N}$ *of* M*. Moreover, Cerf's proof of Proposition 4.2 shows that this instance of* η *is a topological submersion and compatible with the respective stratifications, too.*

The parameter space of discrete Morse matchings has been introduced in Chari and Joswig 2005 and further investigated in Capitelli and Minian 2017 and Lin and Scoville 2021. We show that the parameter space of discrete Morse matchings is also canonically associated to the Morse arrangement:

Proposition 5.3.
Let X *be a CW complex and let* $\mathcal{A}$ *be the Morse arrangement on* $\mathbb{R}^X$*. Then there is a canonical embedding of posets* $\mathfrak{M}(X) \subset \mathcal{L}(\mathcal{A})$*.*

For the parameter space of merge trees, we follow a similar approach as for the parameter space of discrete Morse functions. We model merge trees as maps from combinatorial merge trees to the real numbers and topologize using a mixture of the euclidean distance and a kind of edit distance. Definition 5.15. Our euclidean edit distance is similar to a distance in Wetzels and Garth 2022, although their edit moves focus on edges, whereas our approach focuses on nodes.

Proposition 5.16.
The euclidean edit distance d *from Definition 5.15 is a pseudo metric on the parameter space of merge trees* Mer *and a metric on the parameter space of strict merge trees* $Mer_<$ *and the parameter space of well-branched merge trees* Mer_{wb}*.*

Moreover, we show the following:

Theorem 5.18.
Let X *be a regular CW complex. Then the map* $M\colon \mathcal{M}(X) \to Mer$ *from Definition 5.8 that maps a discrete Morse function to its induced merge tree is continuous.*

For the relationship to the parameter space of barcodes, we use a similar approach[1] as in Brück and Garin 2023 for the space of barcodes and a construction for the barcode induced by a merge tree similar to the one in Curry et al. 2024. We obtain the following result:

[1] One can identify the space of barcodes over a given combinatorial barcode in this work with the space $\mathbb{R}^{2n}$ in Brück and Garin 2023, Section 4.1.

Proposition 5.31.
The map $B\colon Mer_{wb} \to Bar$ from Definition 5.30 is continuous.

Acknowledgements. The author would like to thank Max Planck Institute for Mathematics for the great scientific environment in which this project was conducted. Moreover, the author would like to thank Andrea Bianchi, Florian Kranhold and Paul Mücksch for helpful discussions about the project. Most notably the author thanks his advisor, Viktoriya Ozornova, for her advice, the many helpful discussions, and the detailed feedback at multiple occasions.

2. Hyperplane Arrangements and Their Associated Spaces

We review some well established notions from combinatorial geometry and topology, which will prove to be useful for our endeavor. The standard references we refer to are Björner et al. 1999 and Aguiar and Mahajan 2017. While most notions are standard in the literature, we adapt and extend them to the setting of this article.

Definition 2.1 (compare Björner et al. 1999 and Aguiar and Mahajan 2017).
A ***real hyperplane arrangement*** $\mathcal{A}$ is a finite set of hyperplanes embedded in a finite-dimensional real vector space V. A real hyperplane arrangement is called ***central*** if all hyperplanes contain the origin. For a central arrangement $\mathcal{A}$, the ***center*** of $\mathcal{A}$ is the subspace $\mathcal{Z}(\mathcal{A}) := \bigcap_{H\in\mathcal{A}} H \subset V$ given by the intersection of all hyperplanes of $\mathcal{A}$. A central arrangement is called ***essential*** if $\mathcal{Z}(\mathcal{A}) = \{0\}$. For any central arrangement $\mathcal{A} \subset V$ we call the arrangement induced by the projection $V \to \mathcal{Z}(\mathcal{A})^\perp$ to the orthogonal complement of the center the ***essential arrangement associated to*** $\mathcal{A}$.

An ***oriented hyperplane arrangement*** is a real hyperplane arrangement together with a choice of an orientation class of normal vectors for each hyperplane.

We call the path components of the complement $V \setminus \mathcal{A}$ the ***(open) regions*** of the arrangement $\mathcal{A}$. Moreover, we call intersections of half-spaces associated to $\mathcal{A}$, with at least one half-space per hyperplane chosen, ***faces*** of the hyperplane arrangement. For oriented hyperplane arrangements, each of these faces σ is associated with a sign vector $w \in \{+,-,0\}^{\mathcal{A}}$, where $+$ means that σ lies on the positive side of the corresponding hyperplane, $-$ means that σ lies on the negative side of the corresponding hyperplane, and 0 means that σ lies inside corresponding hyperplane. The regions correspond to sign vectors with no entry 0. The ***face poset*** $\Sigma(\mathcal{A})$ is the set of faces of $\mathcal{A}$ ordered by inclusion.

While hyperplane arrangements can be defined in arbitrary vector spaces, it is often useful to use coordinate spaces, in particular partially ordered coordinate spaces. This allows us to use equations in the ordered coordinates in order to specify the arrangements we consider.

Definition 2.2. A ***partially ordered coordinate space*** is a real vector space V together with a choice of a basis which is indexed over some partially ordered set P.

For convenience, we sometimes choose a specific model for partially ordered coordinate spaces:

Definition 2.3. The category of partially ordered coordinate spaces $po\,$Vect consists of

Objects: vector spaces $\mathrm{Map}(P,\mathbb{R})$ for posets P together with the standard basis induced by P, and

Morphisms: maps $\mathrm{Map}(Q,\mathbb{R}) \to \mathrm{Map}(P,\mathbb{R})$ induced by maps $P \to Q$.

Remark 2.4. The category of partially ordered coordinate spaces can be identified with the image of the contravariant functor $\mathrm{Map}(\_,\mathbb{R})\colon \mathrm{Poset} \to \mathrm{Vect}_{\mathbb{R}}$, where Poset denotes the category of posets and monotone maps, and $\mathrm{Vect}_{\mathbb{R}}$ denotes the category of real vector spaces and linear maps. The specific choice of partially ordered basis is then given by the domain poset and preserved by the condition that morphisms need to be induced by poset maps.

In this work, we mostly consider partially ordered coordinated spaces where the bases are indexed over either face posets of CW complexes or posets associated to merge trees or barcodes. We can characterize these three frameworks as certain types of posets, which makes partially ordered coordinate vector spaces a unifying framework for them:

DEFINITION 2.5. A poset P is of ***regular CW type*** if it is isomorphic to the face poset $D(X)$ of a regular CW complex X.

A poset P is of ***merge tree type*** if it

(1) is finite,
(2) has a unique maximum r,
(3) for each $a \in P$ the set $[a, r]$ is a chain, and
(4) each non-minimal element $a \in P$ has at least two elements $b \neq b'$ such that $b \prec a$ and $b' \prec a$, where $\prec$ denotes cover relation.

A poset P is of ***barcode type*** if it is a finite disjoint sum of chains of length one. We call these chains the ***bars*** of the barcode.

Remark 2.6. In the definition above, we deviated from the standard convention in the literature to consider trees oriented away from the root. We do so because merge trees come with a preferred orientation from leaves to the root, induced by increasing filtration levels.

In order to see that Definition 2.5 models merge trees, we remark (2) and (3) model arbitrary rooted trees: the chains in (3) correspond to shortest paths between the root and other nodes of the tree. The feature that inner nodes of merge trees correspond to mergers of path components, i.e. each inner node has at least two children, is reflected in (4).

Since merge trees can, in this fashion, be considered as both, posets and trees in the graph-theoretic sense, we use the notions from both fields interchangeably. For example, we call minimal elements leaves, non-minimal elements inner nodes, and refer to the maximum as the root.

Example 2.7. We consider the following example of a manifold with a height function, the induced merge tree and the corresponding poset of merge tree type.

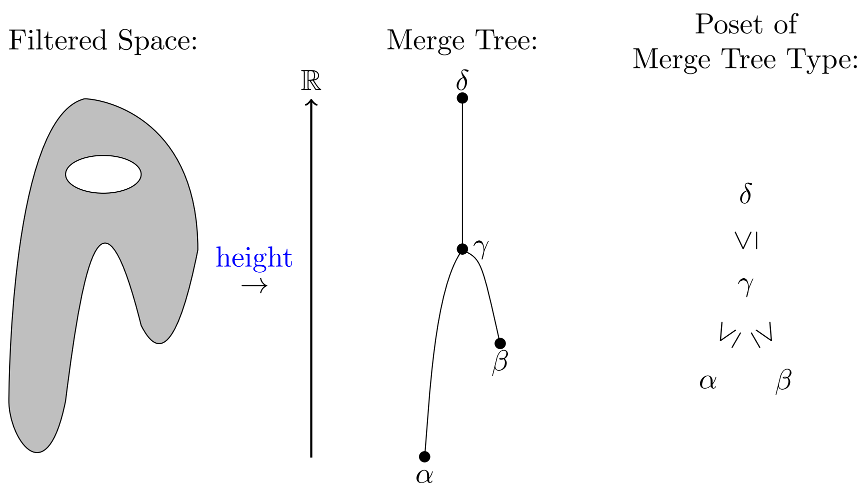


FIGURE 1. An example of a merge tree induced by a height function on a manifold and the corresponding poset of merge tree type.

One classic example of hyperplane arrangements, which is of particular importance for this work, is the braid arrangement.

2. HYPERPLANE ARRANGEMENTS AND THEIR ASSOCIATED SPACES

Example 2.8. We denote by $\mathcal{H}_n$ the hyperplane arrangement in $\mathbb{R}^n$ given by the equations $x_i = x_j$ for all $i \neq j$. We call $\mathcal{H}_n$ the ***braid arrangement*** $\mathcal{H}_n$ (see Björner et al. 1999, Example 2.3.3).

We recall the definitions of the intersection poset and the poset of flats, which are in a sense dual to each other:

Definition 2.9 (Björner et al. 1999, Definition 2.1.3,Aguiar and Mahajan 2017, p. 1.3.1). The ***intersection poset*** $\mathcal{L}(\mathcal{A})$ of a hyperplane arrangement $\mathcal{A}$ is the set of intersections of subfamilies of $\mathcal{A}$, ordered by reverse inclusion. The ***poset of flats*** $\Pi(\mathcal{A})$ is as a set the same as the intersection poset but ordered by inclusion. The elements of the poset of flats, or the intersection poset, respectively, are called the ***flats*** of $\mathcal{A}$.

If $\mathcal{A}$ is central, then $\mathcal{L}(\mathcal{A})$ is a geometric lattice of rank $r(\mathcal{A})$. Since all arrangements in this work are central, we also refer to $\mathcal{L}(\mathcal{A})$ as the ***intersection lattice*** of $\mathcal{A}$. It is straightforward to see that $\Pi(\mathcal{A})$ is never the face poset of a simplicial complex if $\mathcal{A}$ is central and $|\mathcal{A}| > 1$.
On the other hand, $\Sigma(X)$ always has the structure of the face poset of a regular CW complex (see Aguiar and Mahajan 2017, p. 1.1.8).

Definition 2.10 (Aguiar and Mahajan 2017, p. 1.1.9). We call a hyperplane arrangement simplicial if $\Sigma(\mathcal{A}) \setminus \{\mathcal{Z}(\mathcal{A})\}$ has the structure of a face poset of a pure[2] simplicial complex.

In fact, $\Sigma(\mathcal{A}) \setminus \{\mathcal{Z}(\mathcal{A})\}$ always gives a regular CW decomposition of the sphere in $\mathcal{Z}^{\perp}$. Hence, the real condition for a hyperplane arrangement being simplicial is that $\Sigma(\mathcal{A}) \setminus \{\mathcal{Z}(\mathcal{A})\}$ needs to be isomorphic to the face poset of a simplicial complex. Then any such simplicial complex will be pure.

3. The Parameter Space of Discrete Morse Functions

Our goal is to present a framework which allows us to analyze geometric properties of and homotopies between discrete Morse functions on (regular) CW complexes. It turns out that manipulations of discrete Morse functions, like cancellations of critical cells and reordering critical cells by swapping their critical values can be geometrically realized by homotopies between discrete Morse functions, i.e. paths in the space of discrete Morse functions.
In order to define the space of discrete Morse functions on a (regular) CW complex, we consider the space of all discrete functions, which the space of discrete Morse functions will turn out to be a subspace of. As common when working with discrete Morse theory, we adopt a slight abuse of notation by not distinguishing between a CW complex and its set of cells.

Definition 3.1. Let X be a CW complex. We call the space $\mathbb{R}^X := \mathbb{R}^{F(X)} := \prod_{\sigma \in X} \mathbb{R}_\sigma$ the ***space of discrete functions on*** X.

Here, we notice that the coordinates of the space of discrete functions are partially ordered rather than linearly. This partial order will help us to investigate which of the discrete functions are discrete Morse functions.
For that, we recall the definition of discrete Morse functions as given by Forman:

Definition 3.2 (Forman 1998, Definition 2.1). Let X be a finite CW complex and let $F(X)$ be the face poset of X. A ***discrete Morse function*** on X is a function $f\colon F(X) \to \mathbb{R}$ such that for every p-dimensional cell $\alpha^{(p)} \in X$

(1) $\#\{\beta^{(p+1)} \supset \alpha | f(\beta) \leq f(\alpha)\} \leq 1$, and
(2) $\#\{\gamma^{(p-1)} \subset \alpha | f(\gamma) \geq f(\alpha)\} \leq 1$.
(3) $f(\alpha) < f(\beta)$ whenever $\alpha \subset \beta$ is not regular.

[2]A simplicial complex is called ***pure*** if all its inclusion-maximal simplices have the same dimension d.

Cells for which the inequalities in (1) and (2) both are strict are called ***critical***.

Remark 3.3. We include the general version for non-regular CW complexes mostly for the comparison to the smooth case in Section 4. For obvious reasons, we can drop condition (3) if X is regular. It is highly inconvenient to work with discrete Morse theory on non-regular CW complexes because in that case the face poset does not contain all information of the homotopy type at hand. Moreover, we make the definition for the non-regular case a bit stricter than Forman in order to simplify the definition of the induced gradient field, i.e. Morse functions are automatically strictly monotone along face inclusions of codimension higher than one.

We recall that by Forman 1998, Lemma 2.5 for each cell $\alpha \in X$ at most one of the inequalities (1) and (2) from Definition 3.2 may actually be an equality, whereas the other one has to be a strict inequality. Therefore, for each cell $\alpha \in X$ such that (1)/(2) is an equality there is exactly one cell $\beta \supset \alpha/\gamma \subset \alpha$ such that $f(\beta) \leq f(\alpha)/f(\alpha) \leq f(\gamma)$. Hence, any discrete Morse function $f\colon F(X) \to \mathbb{R}$ induces a partial matching ∇f on X by matching α with $u(\alpha) := \beta/d(\alpha) := \gamma$. The unmatched cells are called ***critical***. To be precise, ∇f is a partial matching on the Hasse diagram $D(X)$ of the face poset of X:

Definition 3.4. Let X be a CW complex. The ***Hasse diagram*** of the face poset of X is the directed graph $D(X)$ that has vertices $V(D(X)) := X$, that is the cells of X, and an oriented edge (τ, σ) whenever $\sigma \subset \tau$ is a face relation such that there exists no $\theta \in X$ with $\sigma \subset \theta \subset \tau$ (σ is covered by τ).

Moreover, for a discrete Morse function f on X we define the ***modified Hasse diagram*** $D_f(X)$ that arises from $D(X)$ by inverting the edges matched by f.

Remark 3.5. In case of a regular CW complex X, the directed edges correspond to face relations of codimension one. For non-regular CW complexes, edges might correspond to face relations of higher codimension, which makes finding elementary collapses via matchings harder. Furthermore, for non-regular CW complexes, the gradient paths needed for the boundary map of the Morse–Smale complex are harder to spot inside the modified Hasse diagram, since they are then not just zig-zags between any matched edges and unmatched egdges.

In the literature, there is a more specialized definition of discrete Morse functions which was first proposed by Benedetti in Benedetti 2016:

Definition 3.6 (Benedetti 2016, Section 2.1). Let X be a CW complex. A ***Morse–Benedetti*** function on X is a function on the face poset $f\colon F(X) \to \mathbb{R}$ that fulfills for cells σ, τ that:

Monotonicity: If $\sigma \subset \tau$ we have $f(\sigma) \leq f(\tau)$.
Semi-injectivity: $|f^{-1}(\{z\})| \leq 2$ for all $z \in \mathbb{R}$.
Generacity: If $f(\sigma) = f(\tau)$, then either $\sigma \subseteq \tau$ or $\tau \subseteq \sigma$ holds.
Regularity: If $\sigma \subset \tau$ is irregular, then $f(\sigma) < f(\tau)$.

It is straightforward to see that Morse–Benedetti functions are always discrete Morse functions but in general not the other way around.

For some purposes, it is useful to weaken Benedetti's definition:

Definition 3.7. Let X be a CW complex. We call a function $f\colon F(X) \to \mathbb{R}$ a ***weak Morse–Benedetti function*** if it is a discrete Morse function and satisfies semi-injectivity and generacy in the sense of Definition 3.6.

The only difference between weak Morse–Benedetti functions and Morse–Benedetti functions is that matched cells of Morse–Benedetti functions have the same value, whereas for weak Morse–Benedetti functions, the higher-dimensional cell might have a strictly smaller value than the smaller-dimensional cell of the matched pair.

Example 3.8. We consider the following example of three discrete Morse function that induce the same matching.

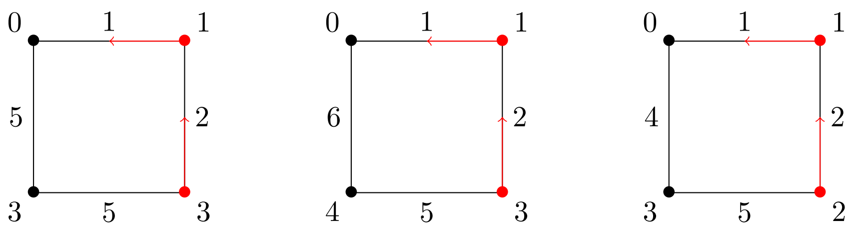


FIGURE 2. Three different dMfs that induce the same matching (indicated in red). The one in the middle is weakly Morse–Benedetti but not Morse–Benedetti. The one on the right is Morse–Benedetti.

The dMf on the left is not weakly Morse–Benedetti due to the two critical edges labeled 5. The dMf in the middle is not Morse–Benedetti because of the matched pair with labels 2 and 3.

The definition of discrete Morse functions gives rise to a subdivision of the space of discrete functions induced by an oriented hyperplane arrangement in $\mathbb{R}^X$ as follows:

We observe that each point $p \in \mathbb{R}^X$ represents a discrete function $f\colon F(X) \to \mathbb{R}$ by $f(\sigma) = p_\sigma$. The conditions (1) and (2) (and (3)) of Definition 3.2 impose conditions on the components of a discrete function $p \in \mathbb{R}^X$ whenever there is a directed edge between the corresponding cells in the Hasse diagram. That is, for the question whether p represents a discrete Morse function it is relevant whether the inequality $p_\sigma < p_\tau$ holds whenever $\sigma \subset \tau$ is represented by a directed edge in $D(X)$. This information is equivalent to the question on which side of the hyperplane H_σ^τ in $\mathbb{R}^X$ given by $x_\sigma = x_\tau$ the point p lies. Hence, the conditions (1) and (2) (and (3)) of Definition 3.2 can be checked by analyzing the position of p relative to the hyperplane arrangement $\mathcal{A}$ given by the equations $x_\sigma = x_\tau$ for all covering face relations $\sigma \subset \tau$ in X. We define the following oriented hyperplane arrangement:

DEFINITION 3.9. Let E be the set of codimension 1 face relations in X. We define the hyperplane arrangement $\mathcal{A}(X) := \{H_\sigma^\tau | \sigma \subset \tau \in E\}$ where H_σ^τ is given by the equation $x_\sigma = x_\tau$. We observe that the assignment $(\sigma \subset \tau) \mapsto H_\sigma^\tau$ defines a bijection $E \cong \mathcal{A}$. Then let $\mathcal{S}(X) := \{+, -, 0\}^E$ be the sign vectors for the oriented hyperplane arrangement $\mathcal{A}^+(X)$ obtained by orienting $\mathcal{A}(X)$ in the following way: For each hyperplane H_σ^τ we define the open half-space defined by $x_\sigma < x_\tau$ to be the positive side of H_σ^τ, the open half-space given by $x_\sigma > x_\tau$ to be the negative side of H_σ^τ, and on H_σ^τ itself we have the value 0. We call $\mathcal{A}^+(X)$ the ***Morse arrangement*** on $\mathbb{R}^X$, respectively on X.

For any sign vector $x \in \mathcal{S}(X)$ we define the sets $x_+ := \{e \in E | x_e = +\}$, $x_- := \{e \in E | x_e = -\}$, and $x_0 := \{e \in E | x_e = 0\}$.

Remark 3.10. The index set E of $\mathcal{S}(X)$ is by construction the set of edges of $D(X)$. Hence, we will use the two different points of view interchangeably without explicitly mentioning it in the notation from now on. Moreover, we drop the $+$ in the notation since there is only one orientation we are interested in. Furthermore, we will denote $\mathcal{A}(X)$ and $\mathcal{S}(X)$ just by $\mathcal{A}, \mathcal{S}$, respectively, when there are not multiple complexes to be confused with.

We interpret an element $x \in \mathcal{S}$ as an equivalence class of discrete functions on some regular CW complex X. For two functions $p, q \in \mathbb{R}^X$ we define $p \sim q$ if and only if for all hyperplanes H, the function p lies on the positive side of H if and only if q lies on the positive side of H. It turns out that this way the regions which contain discrete Morse functions are identified with their equivalence classes given by inducing the same acyclic matching. The positive region, i.e. the region on the positive side of all hyperplanes, turns out to be exactly the set of all critical discrete Morse functions on X. A point $p \in \mathbb{R}^X$ being inside a hyperplane H_σ^τ or on the negative side of H_σ^τ indicates that if p represents a discrete Morse function, then the cells σ and τ are matched by p. Moreover, a function $p \in \mathbb{R}$ is a discrete Morse function if and only if all functions in the same region as p are.

It is immediate that the Morse arrangement $\mathcal{A} \subset \mathcal{H}_{|X|}$ is a subarrangement of the braid arrangement Example 2.8. Moreover, $\mathcal{A}$ is always central but never essential, i.e. the intersection of all hyperplanes of $\mathcal{A}$ always contains $0 \in \mathbb{R}^X$ but never consists of just the origin.

We give a characterization of discrete Morse functions in the language of hyperplane arrangements:

Definition 3.11. Let X be a CW complex and let $\mathcal{A}$ be the corresponding Morse arrangement. A discrete function $p \in \mathbb{R}^X$ is a discrete Morse function if for every cell $\sigma \in X$

(1) $\#\{\tau \in X | (\tau, \sigma) \in D(X),\ p_\tau \leq p_\sigma\} \leq 1$, and
(2) $\#\{\tau \in X | (\sigma, \tau) \in D(X),\ p_\sigma \leq p_\tau\} \leq 1$.
(3) $p_\sigma < p_\tau$ whenever $(\sigma, \tau) \in D(X)$ corresponds to a non-regular cover relation.

We call any region of $\mathcal{A}$ that consists of discrete Morse functions a ***Morse region***. We call the union of all Morse regions of $\mathcal{A}$ the ***parameter space of discrete Morse functions*** on X and denote it by $\mathcal{M}(X)$. We call the unique region $R_{cr}(X)$ that only contains critical discrete Morse functions, i.e. discrete Morse functions for which every cell is critical, the ***critical region*** of $\mathcal{A}$. Moreover, we define the ***essential space of discrete Morse functions*** $\mathcal{M}_{ess}(X)$ as the intersection $\mathcal{M}_{ess}(X) := \overline{R_{cr}(X)} \cap \mathcal{M}(X)$. We define the ***space of Morse–Benedetti functions*** $\mathcal{MB}(X) \subset \mathcal{M}(X)$ as the subspace of those discrete Morse functions that satisfy monotonicity, semi-injectivity and generacy. We define the ***space of weak Morse–Benedetti funtcions*** $\mathcal{MB}_w(X) \subset \mathcal{M}(X)$ as the subspace of those discrete Morse functions that satisfy semi-injectivity and generacy.

Remark 3.12. The Morse arrangement is in general not a reflection arrangement and not simplicial. For example, the critical region of the Morse arrangement on the 2-simplex Δ^2 has 9 faces but it would need to have only 6 faces in order to be simplicial due to the dimension of $\mathcal{Z}(\mathcal{A}(\Delta^2))^\perp$. Hence, the regular CW complex associated to the Morse arrangement, as well as its subcomplex induced by Morse regions, is in general not simplicial.

The reflection arrangement associated to $\mathcal{A}(X)$ for a regular CW complex X is the braid arrangement associated to the symmetric group $\mathfrak{S}_{|X|}$. For arbitrary CW complexes, the story is less well behaved because one has a priori no combinatorial control over cover relations being regular.

Example 3.13. We consider the space of discrete Morse functions on the interval I:

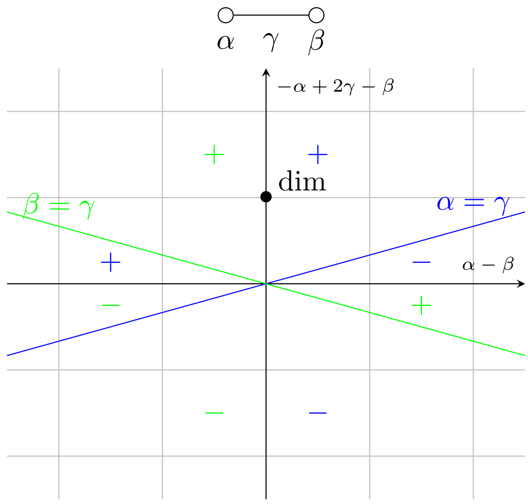


FIGURE 3. The Morse arrangement in the space of discrete functions on I

In order to draw the picture we made use of the fact that summands of multiples of the vector $(1, 1, \dots, 1)$ have no effect on whether a point in $\mathbb{R}^X$ corresponds to a Morse region or not. Hence, we only drew the orthogonal complement $\operatorname{span}\langle(1, 1, \dots, 1)\rangle^{\perp}$. That is, we actually drew the essential arrangement associated to $\mathcal{A}$.

Proposition 3.14. *Let X be a regular CW complex, then the space of discrete Morse functions $\mathcal{M}(X)$ is contractible.*

PROOF. We prove the slightly stronger statement that $\mathcal{M}(X)$ is a star domain with the dimension function $p_\sigma = \dim(\sigma)$ as a star point. Let f be a discrete Morse function, we show that $f_t := (1-t) \cdot f + t \cdot \dim$ is a Morse function for all $t \in [0, 1]$. In order to do that, we prove that for any face relation $\sigma \subset \tau$ of codimension one where $f_t(\tau) \leq f_t(\sigma)$ holds, the inequality $f(\tau) \leq f(\sigma)$ must also hold. For $t = 1$ we have $f_1 = \dim$ and the statement is true. For $t < 1$ we have:

$$f(\tau) = \frac{f_t(\tau) - t \cdot (\dim(\sigma) + 1)}{1-t} \leq \frac{f_t(\tau) - t \cdot \dim(\sigma)}{1-t} \leq \frac{f_t(\sigma) - t \cdot \dim(\sigma)}{1-t} = f(\sigma).$$

This implies that f_t is a discrete Morse function for all t if f is a discrete Morse function. Hence, we connected any discrete Morse function to the dimension function by a line segment of discrete Morse functions. In particular, $\mathcal{M}(X)$ is contractible.

□

Proposition 3.15. *Let X be a regular CW complex. Then we have $\mathcal{MB}(X) \subset \overline{R_{cr}(X)}$. Moreover, $\mathcal{MB}_w(X)$ is dense in $\mathcal{M}(X)$ and $\mathcal{MB}(X)$ is dense in $\mathcal{M}_{ess}(X)$.*

PROOF. Since Morse–Benedetti functions[3] are monotone, they can only lie on the positive side of or on a hyperplane $H \in \mathcal{A}(X)$ but never on the negative side. Hence, all Morse–Benedetti functions lie in the closure of the critical region $\overline{R_{cr}(X)}$.

Due to generacy, weak Morse–Benedetti functions cannot have the same value on different cells unless such cells are faces of one another. By semi-injectivity, this can only happen if the two cells are faces of codimension one, i.e. if they are matched. Hence, discrete Morse functions can only fail to be weak Morse–Benedetti functions on intersections of at least one hyperplane of $\mathcal{H} \setminus \mathcal{A}$. Thus, $\mathcal{M}(X) \setminus \mathcal{MB}_w(X)$ is contained in the first stratum of the stratification induced by $\mathcal{H}(X)$ which means that $\mathcal{MB}_w(X)$ is dense in $\mathcal{M}(X)$. It follows that $\mathcal{MB}(X)$ is dense in $\mathcal{M}_{ess}(X)$ by restricting the stratification induced by $\mathcal{H}$ to $\mathcal{M}_{ess}(X)$. □

[3]See Definition 3.6

3. THE PARAMETER SPACE OF DISCRETE MORSE FUNCTIONS

Having established that the space of discrete Morse functions is homotopically trivial, we are going to investigate its combinatorial and geometrical properties. Since properties of objects are ideally investigated via morphisms, we take a look at meaningful notions of morphisms in this context. The goal is to identify morphisms between CW complexes that induce a meaningful notion of morphisms between spaces of discrete Morse functions.

DEFINITION 3.16. Let X, Y be regular CW complexes. We say that a set of 0-cells $\{\sigma_i\} \subset X$ ***spans*** a cell $\tau \in X$ if $\{\sigma_i\} = (\partial\bar{\tau})_0$.

A continuous map $\varphi\colon X \to Y$ is called ***cellular*** if it maps the n-skeleton X_n to the n-skeleton Y_n. A map $\varphi\colon X \to Y$ is called ***order-preserving*** if it corresponds to an order-preserving map $D(\varphi)\colon D(X) \to D(Y)$. A map $\varphi\colon X \to Y$ is called ***simplicial*** if

- it maps n-cells to k-cells for arbitrary $k \leq n$ and
- for any set of 0-cells $\{\sigma_i\} \subset X$ such that $\{\sigma_i\} \subset X$ spans an n-cell σ for some n, the set $\{\varphi(\sigma_i)\} \subset Y$ spans the k-cell $\varphi(\sigma)$ for some $k \leq n$.

Remark 3.17. One central point of this work is the interplay between CW complexes and their face posets. Although discrete Morse theory excels on simplicial complexes, we need to include CW complexes for the comparison to smooth Morse theory in Section 4. Hence, on one hand we need to treat discrete Morse functions on CW complexes and on the other hand we want to treat discrete Morse functions on abstract simplicial complexes. We consider both points of view on regular CW complexes as a middle ground, since homotopy types of regular CW complexes are uniquely determined by their face posets. Nonetheless, it is important to point out that even though the face poset determines its corresponding regular CW complex up to homotopy, geometric cellular morphisms of regular CW complexes are more general than morphisms of the corresponding face posets.

We also remark that in Definition 3.16 cellular maps belong the the framework of CW complexes, whereas order-preserving maps and simplicial maps belong to the realm of face posets.

Example 3.18. In order to see that cellular maps do not necessarily induce maps on face posets, consider a regular CW decomposition of the circle S^1 in two 0-cells a, b and two 1-cells A, B. Let $f\colon S^1 \to S^1$ be given by $f(a) = f(b) = f(A) = a$ and f is given ob B by winding around the circle twice. Then f does not induce a map on face posets because the cell B is mapped to multiple cells, to be precise to all cells of S^1. Nonetheless, f maps the 0-skeleton to the 0-skeleton and the 1-skeleton to the 1-skeleton. Since f winds around the circle twice, the map f is also not homotopic to any map that could induce a map on face posets.

There is a more general notion of equivalence of CW complexes in the literature:

DEFINITION 3.19 (Banyaga and Hurtubise 2004, Definition 6.32). Let X and X' be finite CW complexes. For any cell $e \subset X$ denote by $X(e)$ the smallest subcomplex of X containing e.

The complexes X and X' are called ***cell equivalent*** if and only if there is a homotopy equivalence $h\colon X \to X'$ with the property that there is a bijective correspondence between cells in X and cells in X' such that if $e \subset X$ corresponds to $e' \subset X'$, then h maps $X(e)$ to $X'(e')$ and is a homotopy equivalence of these subcomplexes.

Remark 3.20. Cell equivalence can be seen as a generalization of simplicial equivalences to arbitrary CW complexes. Cell equivalences preserve both, the combinatorial and homotopical structure of CW complexes. In particular, this means that cell equivalences induce isomorphisms of face posets and preserve homotopy classes of characteristic maps.

Cell equivalences seem to be well suited for discrete Morse theory because they preserve both, the combinatorial and the topological structure of CW complexes. Thus, we introduce a notion of morphism that generalizes the notion of cell equivalence while recovering cell equivalences as isomorphisms.

DEFINITION 3.21. Let X and Y be CW complexes. A map $\varphi\colon X \to Y$ is called ***non-degenerate*** if it maps 0-cells to 0-cells, and induces a map $D(\varphi)\colon D(X) \to D(Y)$ that preserves cover relations, and φ induces homotopy equivalences $X(\sigma) \simeq Y(\varphi(\sigma))$ for all $\sigma \in X$.

Remark 3.22. It is straightforward that for regular CW complexes, cell equivalences induce simplicial equivalences. Moreover, a map φ between regular CW complexes is non-degenerate if and only if it maps 0-cells to 0-cells and $D(\varphi)$ preserves cover relations, i.e. face relations of codimension one. In more generality, injective non-degenerate maps between CW complexes are cell equivalences onto their image.

For the context of discrete Morse theory, we prefer the point of view of simplicial maps between regular CW complexes because they focus more on the combinatorial structure at hand. Nonetheless, in the context of arbitrary CW complexes one has to work with cell equivalences because they keep track of attaching maps being irregular.

Furthermore, cell equivalences induce isomorphisms of the corresponding face posets. This way, cell equivalences and non-degenerate maps are the connecting piece between geometric and abstract morphisms of regular CW complexes.

Proposition 3.23. *A map $\varphi\colon X \to Y$ between regular CW complexes is simplicial if and only if it is cellular and order preserving.*

PROOF.

$\Rightarrow$
: Let $\sigma \subseteq \tau$ be a face relation in $D(X)$. The vertices σ_0 form a subset of the vertices τ_0. This implies $\varphi(\sigma_0) \subseteq \varphi(\tau_0)$. Since φ is simplicial, it follows that $\varphi(\sigma_0)$ spans a cell of dimension less or equal to $\dim \tau$. Thus, φ is order-preserving. Moreover, the $\dim \sigma$-skeleton of X is mapped to the $\dim \sigma$-skeleton of Y.

$\Leftarrow$
: The property of φ being cellular in particular implies that φ preserves the 0-skeleton, i.e. maps 0-cells to 0-cells. It follows inductively that for a set of 0-cells $\{\sigma_i\} \subset X$, the set $\{\varphi(\sigma_i)\} \subset Y$ spans a k-cell if $\{\sigma_i\} \subset X$ spans a n-cell for some $k \leq n$. Hence, φ is simplicial.

□

Proposition 3.24. *Non-degenerate maps of regular CW complexes are simplicial and map n-cells to n-cells. Simplicial maps of CW complexes that map n-cells to n-cells are non-degenerate.*

PROOF. Since in regular CW complexes the combinatorial structure of all cells is given by chains of inclusions of the corresponding faces, the first statement follows by induction over chains of non-trivial inclusions of faces of codimension one.

For the second statement, we observe that simplicial maps in particular map 0-simplices to 0-simplices. For the preservation of cover relations we observe that n-cells of regular CW complexes are characterized by chains of cover relations of their faces of length $n+1$. Since these chains need to preserve their length under simplicial maps that map n-cells to n-cells, cover relations also need to be preserved because otherwise the chains would get shorter. □

Having established the notion of non-degenerate maps, we want to use non-degenerate maps to induce maps on spaces of discrete Morse functions. The straightforward way to do that is to use the contravariant functoriality of partially ordered coordinate spaces, see Remark 2.4.

3. THE PARAMETER SPACE OF DISCRETE MORSE FUNCTIONS

Proposition 3.25. *Let $\varphi\colon X \to Y$ be a non-degenerate map of CW complexes. Then there is an induced map $\varphi^*\colon \mathbb{R}^Y \to \mathbb{R}^X$ given by $\varphi^* g := g \circ \varphi$, for $g\colon Y \to \mathbb{R}$.*

If φ is injective, then $\varphi^{*-1}(\mathcal{A}(X)) \subset \mathcal{A}(Y)$ is a subarrangement. Moreover, φ^* restricts to a map $\mathcal{M}(Y) \to \mathcal{M}(X)$.

Proof. The first statement is true because non-degenerate maps induce maps on face posets. The second statement follows because non-degenerate maps preserve cover relations. The third statement follows by injectivity. □

Since we require the relatively strong notion of non-degenerate maps, we also get a map in the other direction.

Lemma 3.26. *Let $\varphi\colon X \to Y$ be a non-degenerate map between CW complexes. Then φ induces a linear map $\varphi_*\colon \mathbb{R}^X \to \mathbb{R}^Y$ that maps $\mathcal{A}^+(X)$ to $\mathcal{A}^+(Y)$. If φ is injective, then φ_* embeds $\mathcal{A}^+(X)$ into $\mathcal{A}^+(Y)$.*

Proof. By Proposition 3.24, φ maps n-cells to n-cells and, therefore, induces a linear map $\varphi_*\colon \mathbb{R}^X \to \mathbb{R}^Y$. Since φ preserves cover relations in $D(X)$ and subcomplexes up to homotopy, φ preserves the property of cover relations being regular. Hence, φ_* maps $\mathcal{A}^+(X)$ to $\mathcal{A}^+(Y)$. If φ is injective, so is $D(\varphi)$ and it follows that φ_* embeds $\mathcal{A}^+(X)$ into $\mathcal{A}^+(Y)$. □

Remark 3.27. The term non-degenerate comes from the fact that the same construction with degenerate simplicial maps might lead to mapping discrete Morse functions to discrete functions with degenerate cells.

In fact, if φ is a degenerate simplicial map, $D(\varphi)$ does not preserve cover relations. Since $D(\varphi)$ is a poset map between finite posets, the fiber $D(\varphi)^{-1}(y)$ of a point $y \in D(Y)$ is either empty or an interval in $D(X)$. Intervals in $D(X)$ correspond to subspaces in $\mathbb{R}^X$ which means that φ_* collapses the subspace corresponding to $D(\varphi)^{-1}(y)$ to one coordinate corresponding to y. Geometrically, this corresponds to collapsing[4] the corresponding subcomplex $U \subset X$ to a point.

In order to define an induced map on the space of discrete functions one needs to make a choice for a linear map $\tilde{\varphi}_*\colon \mathbb{R}^U \to \mathbb{R}^y$.

Canonical options like projections to one of the coordinates or taking the sum of the coordinates in general do not preserve the property of being a discrete Morse function. The only canonical possibilities that induce a map on the space of discrete Morse functions are the ones which ensure that the images of the collapsed subcomplexes become critical. If one wants to preserve matched pair of simplices along the collapsed subcomplex, one needs to apply more elaborate constructions that depend on the given input Morse function.

Definition 3.28. Let X be a CW complex. For $\sigma \in X$, we define

- $\mathrm{Face}(\sigma) := \{\tau \in X \mid \tau \subset \sigma \text{ arbitrary face relation}\}$,
- $\mathrm{Coface}(\sigma) := \{\tau \in X \mid \tau \supset \sigma \text{ arbitrary face relation}\}$,
- $\mathrm{Face}_1(\sigma) := \{\tau \in X \mid \tau \subset \sigma \text{ cover relation in } D(X)\}$, and
- $\mathrm{Coface}_1(\sigma) := \{\tau \in X \mid \tau \supset \sigma \text{ cover relation in } D(X)\}$.

Definition 3.29. Let $\varphi\colon X \to Y$ be a non-degenerate map between CW complexes. We define an induced map $\varphi_*\colon \mathbb{R}^X \to \mathbb{R}^Y$ on the spaces of discrete functions. We define $\varphi_*(f)$ for any discrete function $f \in \mathbb{R}^X$ as follows:

(1) for every $\tau \in Y$ such that $|\varphi^{-1}(\tau)| = 1$, i.e. $\varphi^{-1}(\tau) = \{\sigma\}$, we define
$$\varphi(f)_\tau := f_\sigma$$
for any $f \in \mathbb{R}^X$.

[4]in the sense of collapsing a topological space, not necessarily in the sense of simple collapses

(2) for any $\tau \in Y$ such that $|\varphi^{-1}(\tau)| \neq 1$ we proceed as follows:
We define

$$\mathrm{up}(f_{|\mathrm{Coface}_1(\tau)}) := \begin{cases} \min(f_{|\mathrm{Coface}_1(\tau)}) & \text{if } f \text{ is defined on some cell of } \mathrm{Coface}_1(\tau). \\ \min(f_{|\mathrm{Coface}(\tau)}) & \text{if } f \text{ is only defined on cells of } \mathrm{Coface}(\tau) \setminus \mathrm{Coface}_1(\tau) \\ \max(f_{|\mathrm{Face}_1(\tau)}) + 2 & \text{if } f \text{ is not defined on any cell of } \mathrm{Coface}(\tau). \end{cases}$$

Here, we slightly abuse notation because $f_{|\mathrm{Coface}_1(\tau)}$ might not be defined for all cells of $\mathrm{Coface}_1(\tau)$. If $f_{|\mathrm{Coface}_1(\tau)}$ is only defined for some cells of $\mathrm{Coface}_1(\tau)$, we take the minimum on those cells. Start with the minimal dimensional τ such that $|\varphi^{-1}(\tau)| \neq 1$.

- If $\mathrm{Face}_1(\tau) = \emptyset$ and $\mathrm{Coface}_1(\tau) = \emptyset$, then define
$$\varphi_*(f)_\tau := 0.$$
- If $\mathrm{Face}_1(\tau) = \emptyset$ and $\mathrm{Coface}_1(\tau) \neq \emptyset$, then define
$$\varphi_*(f)_\tau := \begin{cases} \mathrm{up}(f_{|\mathrm{Coface}_1(\tau)}) - 1 & \text{if } f \text{is defined on any cell of } \mathrm{Coface}(\tau) \\ 0 & \text{else} \end{cases}.$$
- If $\mathrm{Face}_1(\tau) \neq \emptyset$ and $\mathrm{Coface}_1(\tau) = \emptyset$, then define
$$\varphi_*(f)_\tau := \max(f_{|\mathrm{Face}_1(\tau)}) + 1.$$
- If $\mathrm{Face}_1(\tau) \neq \emptyset$ and $\mathrm{Coface}_1(\tau) \neq \emptyset$, then define
$$\varphi_*(f)_\tau := \frac{\max(f_{|\mathrm{Face}_1(\tau)}) + \mathrm{up}(f_{|\mathrm{Coface}_1(\tau)})}{2}.$$

We apply this definition inductively over the dimension of the cells for which φ_* is not yet defined.

Remark 3.30. The set $\mathrm{Face}_1(\sigma)$ corresponds to the set of elements τ of $D(X)$ which are covered by σ, whereas the cells of $\mathrm{Coface}_1(\sigma)$ are exactly the cells τ of X which cover σ. If one uses the covering relation instead of the codimension, there is a straightforward generalization to non-regular CW complexes.

Proposition 3.31. *The induced map φ_* is affine-linear. Moreover, φ_* is linear if and only if $\mathrm{Face}_1(\tau) = \emptyset$ and $\mathrm{Coface}_1(\tau) = \emptyset$ for all τ such that $|\varphi^{-1}(\tau)| \neq 1$. In particular, φ_* is always linear if φ is injective.*

PROOF. The proof is straightforward from the definition. □

Proposition 3.32. *The map φ_* restricts to a map of spaces of discrete Morse functions $\varphi_*\colon \mathcal{M}(\mathcal{X}) \to \mathcal{M}(Y)$. Moreover, φ_* preserves induced matchings if φ is injective.*

PROOF. If φ in injective, then X can be interpreted as a subcomplex of Y. The map $\varphi_*(f)$ is then just an extension of f to Y that is critical on $Y \setminus \varphi(X)$. By construction, we have $\varphi_*(f)_{|\varphi(X)} = f \circ \varphi^{-1}$ and, therefore, the second statement holds.

For cells τ such that $|\varphi^{-1}(\tau)| \neq 1$, we observe that according to Definition 3.29 $\varphi_*(f)$ is constructed such that every face that is covered by τ gets a strictly smaller function value than τ and every coface that covers τ gets a strictly greater function value than τ under $\varphi_*(f)$. Hence, every τ such that $|\varphi^{-1}(\tau)| \neq 1$ is by construction a critical cell of $\varphi_*(f)$. In particular, $\varphi_*(f)$ is a discrete Morse function. □

Example 3.33. We consider the following simplicial map between two regular complexes:

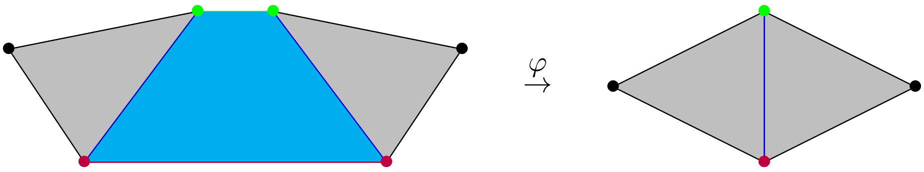
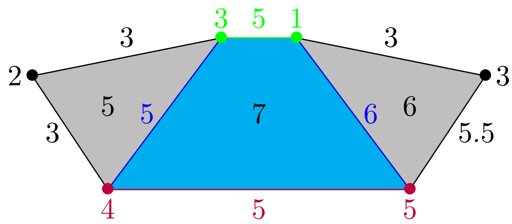
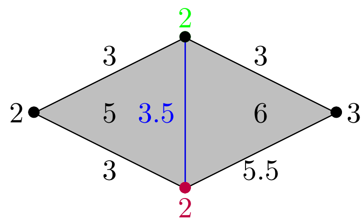


FIGURE 4. A simplicial map between regular CW complexes

We consider the following discrete Morse function f:



The induced discrete Morse function $\varphi_*(f)$ is:



Remark 3.34. As we see in Example 3.33, the map φ_* only preserves matchings where it is injective. Therefore, $\varphi_*(f)$ might in general induce a Morse complex with more cells than f does even if φ collapses a subcomplex of X.

4. Relationship to Smooth Morse Theory

In this section we want to recall the description of the space of Morse functions on a smooth manifold given by Cerf in Cerf 1970. Moreover, we will explain the relationship between certain neighborhoods of Morse functions and the space of discrete Morse functions on their induced CW decompositions, inspired by Cerf 1970, Section 3.2.

In order to get going, we start with a description of the space of smooth Morse functions on a given manifold M and its relationship to Cerf's stratifications Cerf 1970 of the space of smooth functions and the space of Morse functions. Throughout this section, we assume any smooth manifold M to be finite-dimensional, compact, and with a fixed Riemannian metric.

DEFINITION 4.1 (Cerf 1970, Definition 1, 2, and 3). Let M be a smooth manifold. Denote by $C^\infty(M)$ the space of smooth functions from M to $\mathbb{R}$ together with the $\mathcal{C}^\infty$ topology. Let $f \in C^\infty(M)$ be a smooth function, let $p \in M$ be a critical point of f, and let $a \in \mathbb{R}$ be a critical value of f.

The ***codimension of*** $p \in M$ is defined as $\operatorname{codim}(\partial_\mathrm{p}(\mathrm{f}) \subset C^\infty_{0,p}(M)) :=$ $\dim(C^\infty_{0,p}(M)/(\partial_\mathrm{p}(\mathrm{f})))$, where $C^\infty_{0,p}(M)$ denotes the ring[5] of germs of smooth functions vanishing at p and $(\partial_\mathrm{p}(\mathrm{f}))$ denotes the ideal generated by the partial derivatives of f at the point p.

The ***codimension of*** a is defined as $\operatorname{codim}(a) := |\{p \in f^{-1}(a) | p \text{ critical}\}| - 1$.

[5]These rings do not necessarily contain a unit.

The ***codimension of*** f is defined as

$$\operatorname{codim}(f) := \sum_{p \in M \text{ critical}} \operatorname{codim}(p) + \sum_{a \in \mathbb{R} \text{ critical}} \operatorname{codim}(a).$$

Moreover, we define $\mathcal{F} := C^\infty(M)$ and $\mathcal{F}^j := \{f \in \mathcal{F} | \operatorname{codim}(f) = j\}$. We call $\{\mathcal{F}^j\}_{j \in \mathbb{N}}$ the ***natural stratification of*** $\mathcal{F}$.

We recall that for a function $f \in \mathcal{F}$ having a higher codimension than 0 means either that f has degenerate critical points or multiple critical points of the same value. We denote the space of (smooth) Morse functions by $\mathcal{M}_{\text{smo}}(M) \subset \mathcal{F}$.

Next, we consider Cerf's comparison to the space of discrete functions.

We refer to Sharko 1993, Chapter 1, §3 and Cerf 1970 for background information on the natural stratification of the space of smooth functions $\mathcal{F}$. Before proceeding, we want to recall from Sharko 1993, Chapter 1, §3 that $\mathcal{F}$ is a smoothly path connected smooth Fréchet manifold. From now on, all paths will be at least C^1 and path components will refer to C^1 path components.

Proposition 4.2 (Cerf 1970, Section 3.2: Definition 5, Lemma 1)**.** *Let $f \in \mathcal{M}_{smo}(M)$ be a smooth Morse function with q critical points. Given a choice of an ordering of the critical points $c_1, \ldots, c_q$, there are open neighborhoods U_i with $c_i \in U_i \subset M$ for which there is an open neighborhood $\mathcal{V}$ with $f \in \mathcal{V} \subset \mathcal{F}$ such that all $\tilde{f} \in \mathcal{V}$ are Morse and have exactly one critical point $\tilde{c}_i \in U_i$ of the same index as c_i for all $1 \leq i \leq q$. Moreover, the ordering of the critical points of f defines a topological submersion*

$$\eta \colon \mathcal{V} \to \mathbb{R}^q$$
$$\tilde{f} \mapsto (\tilde{f}(\tilde{c}_1), \ldots, \tilde{f}(\tilde{c}_q))$$

such that the restriction of the natural stratification of $\mathcal{F}$ to $\mathcal{V}$ is the preimage of the stratification induced by the hyperplane arrangement $\mathcal{H}_q$ (see Example 2.8).

Our plan for this section is to identify the space $\mathbb{R}^q$ in Proposition 4.2 with the space of discrete functions on the CW decomposition of M induced by f, and use the stability of Morse–Smale functions to extend η to path components. This allows for comparison maps from path components of the space of smooth Morse functions on M to spaces of discrete Morse functions on certain cellular decompositions of M. For that, we recall the notion of Morse–Smale functions.

Definition 4.3 (Banyaga and Hurtubise 2004, Definition 6.1). A Morse function $f \colon M \to \mathbb{R}$ on a finite dimensional smooth Riemannian manifold (M, g) is said to satisfy the ***Morse–Smale transversality condition*** if and only if the stable manifold of p and unstable manifold of q with respect to f intersect transversally for all pairs of critical points p, q of f. A Morse function that satisfies the Morse–Smale transversality condition is called a ***Morse–Smale function***.

It is well known (see e.g. Banyaga and Hurtubise 2004, Theorem 6.34) that Morse functions only induce a cellular decomposition M_f of M, unique up to cell equivalence[6], if they fulfill the Morse–Smale property and are compatible with the given Riemannian metric. However, it is proved in Franks 1979, Proposition 1.6 that the gradient of any Morse–Smale function can be perturbed to be compatible with the Riemannian metric without leaving its path component in the space of Morse–Smale vector fields, in particular while preserving topological conjugacy. Thus, we have a well-defined CW decomposition of M induced by any Morse–Smale function.

Alternatively, one can associate gradient-like Morse–Smale vector fields to non-Morse–Smale functions instead (see Matsumoto 2002, Chapter 3 and Section 4.2). Since we want to

[6]Recall the Definition 3.19 of cell equivalences.

take the point of view of the parameter space of Morse functions, we have fixed a Riemannian metric and consider the subspace $\mathcal{MS}(M) \subset \mathcal{M}_{\text{smo}}(M)$ of Morse–Smale functions on M.

Definition 4.4. Let Z be either the space of Morse functions $\mathcal{M}_{\text{smo}}(M) \subset \mathcal{F}$ or the space of Morse–Smale functions $\mathcal{MS}(M) \subset \mathcal{M}_{\text{smo}}(M) \subset \mathcal{F}$ and let $f \in Z$. We denote by $\mathcal{N}(f) \subset Z$ the path component of f in Z[7].

It is a classical result that the property of being a Morse function is stable, i.e. for any Morse function f, there is an open neighborhood $U(f) \subset C^\infty(M)$ such that $U(f) \subset \mathcal{M}_{\text{smo}}(M)$ and for every path $f_\_ \colon [0,1] \to U(f)$ there is an $\epsilon > 0$ such that f_t is Morse for every $t < \epsilon$ (see Banyaga and Hurtubise 2004, Corollary 5.24). This classical result is also used in the proof of *Proposition* 4.2. Moreover, in Franks 1979, § the stability of Morse–Smale functions is stated in an even stronger way, which we use for the following theorem: if ϕ is a Morse–Smale flow and ϕ' is a sufficiently small C^1 perturbation of ϕ, then there is a homeomorphism $M \to M$ carrying orbits of ϕ to orbits of ϕ' and preserving their orientation. In other words: ϕ and ϕ' are ***topologically conjugate***.

Theorem 4.5. *Let f, g be Morse–Smale functions on M which are contained in the same path component, i.e. $g \in \mathcal{N}(f) \subset \mathcal{MS}(M)$. Then there is a bijection $\psi \colon \operatorname{Cr}(f) \xrightarrow{\cong} \operatorname{Cr}(g)$ such that there is an orbit of ϕ_f connecting two critical points $c_1, c_2 \in \operatorname{Cr}(f)$ if and only if there is an orbit of ϕ_g connecting $\psi(c_1), \psi(c_2) \in \operatorname{Cr}(g)$.*

Proof. This result is a consequence of the structural stability of Morse–Smale functions, respectively Morse–Smale vector fields, respectively Morse–Smale flows. Let f_t be a path from f to g in $\mathcal{MS}(M)$. We choose an open cover on I that consists of such open neighborhoods $U(t_i)$ of time steps t_i that the restriction of f_t to $U(t_i)$ is a small enough C^1 perturbation so that stability in the sense of Franks 1979 holds. Due to I being compact, there is a finite subcover of neighborhoods where the restriction of f_t induces topologically conjugate flows. Thus, f_t induces topologically conjugate flows for all t. Hence, the statement follows. □

Corollary 4.6. *For any compact finite-dimensional Riemannian manifold M, all Morse–Smale functions which are in the same path component of $\mathcal{MS}(M)$ induce the cell-equivalent CW decompositions M_f of M. That is, there is an up to cell equivalence well-defined CW decomposition $M_{\mathcal{N}}$ associated to any path component $\mathcal{N}$ of the space of Morse–Smale functions on M.*

Corollary 4.7. *It follows from Theorem 4.5 that Cerf's map from Proposition 4.2 extends to a map $\eta \colon \mathcal{N} \to \mathcal{M}(M_{\mathcal{N}})$, where $\mathcal{M}(M_{\mathcal{N}})$ denotes the CW decomposition of M induced by any Morse function in the path component $\mathcal{N}$ of $\mathcal{M}(M)$. Moreover, Cerf's proof of Proposition 4.2 shows that this instance of η is a topological submersion and compatible with the respective stratifications, too.*

Building on Corollary 4.6, we can extend the map η from Proposition 4.2 to the path component $\mathcal{N}(f)$ of a given Morse–Smale function f. By Proposition 4.2, it still follows that $\eta \colon \mathcal{N}(f) \to \mathcal{M}(M_{\mathcal{N}(f)}) \coloneqq \mathcal{M}(M_f)$ is a topological submersion which is compatible with the respective stratifications.

Using Cerf's stratification of the space of Morse functions, especially the connection between different path components given by cancellations and creations of pairs of critical points, we want to further extend η to certain systems of path components:

Definition 4.8. Consider the set $\pi_0(\mathcal{MS}(M))$ of path components of the space of Morse–Smale functions on M. Let $\mathcal{N}_1, \mathcal{N}_2 \in \pi_0(\mathcal{MS}(M))$ be path components. Then we define a partial order $\leq$ on $\pi_0(\mathcal{MS}(M))$ generated by $\mathcal{N}_1 < \mathcal{N}_2$ if for any $f \in \mathcal{N}_2$ there is a $g \in \mathcal{N}_1$ such

[7] It will be clear from the context which of the two possibilities for Z will be used.

that g arises from f by a cancellation of a pair of critical points. We call $\leq$ the ***cancellation order*** on $\pi_0(\mathcal{MS}(M))$.

Remark 4.9. It is shown in Cerf 1970, Section 3 that if two generic Morse functions, i.e. elements of $\mathcal{F}^0$, are related to each other by a cancellation of one pair of critical points, then all paths in $\mathcal{F}$ between them are homotopic in $\mathcal{F}$ to one which traverses the first stratum $\mathcal{F}^1$ in one isolated point and is otherwise entirely contained in $\mathcal{F}^0$. Hence, we can think about two path components $\mathcal{N}_1, \mathcal{N}_2$ such that $\mathcal{N}_1 < \mathcal{N}_2$ as being neighbors, only separated by a codimension one stratum which corresponds to a hyperplane in $\mathcal{M}(\mathcal{N}_2)$ via η.

Proposition 4.10. *Let $f, g \in \mathcal{MS}(M)$ be two Morse–Smale functions such that f arises from g by a cancellation of one pair of critical points. Then M_f and M_g are simply homotopy equivalent by a simple collapse $M_g \to M_f$/ a simple extension $M_f \to M_g$ of the two cells which correspond to the canceled critical points at the corresponding sublevel complex. In particular, the map $M_f \to M_g$ is non-degenerate*[8].

Proof. This statement is basically a rephrased version of Matsumoto 2002, Theorem 3.34 (canceling handles: rephrased). In Matsumoto 2002, Theorem 3.34, building on techniques from Milnor 1965, Proof of Theorem 5.4 the authors argue that a cancellation of a pair of critical points induces a pair of handles in the induced handle body decomposition such that the handlebody decomposition induced by f is diffeomorphic to the one induced by g. The statement then follows in a straightforward manner from the construction of the CW decomposition associated to a handlebody decomposition Matsumoto 2002, Theorem 4.18. □

Corollary 4.11. *Let $\mathcal{N}_1 < \mathcal{N}_2 \in \pi_0(\mathcal{MS}(M))$ be path components, which are related by cancellations of critical points. Then $\mathcal{A}(M_{\mathcal{N}_1})$ canonically embeds into $\mathcal{A}(M_{\mathcal{N}_2})$ by Lemma 3.26. Moreover, we have an induced map $\mathcal{M}(M_{\mathcal{N}_1})) \to \mathcal{M}(M_{\mathcal{N}_2}))$ given by Definition 3.29.*

5. Relationship to Other Concepts in Persistent Topology

5.1. The space of discrete Morse matchings. In the literature, there is already work on a so-called space, or complex, of discrete Morse functions. In fact, this name has been used with a slight abuse of terminology since these works refer to a complex whose simplices correspond to Morse matchings rather than Morse functions.

Definition 5.1 (Chari and Joswig 2005, Section 2). Let X be a simplicial complex. The ***complex of discrete Morse matchings*** $\mathfrak{M}(X)$ has as vertices cover relations in $D(X)$ and simplices sets of cover relations in $D(X)$ that correspond to acyclic matchings on $D(X)$.

It is straightforward how this notion should be generalized to arbitrary CW complexes:

Definition 5.2. Let X be a CW complex. The ***complex of discrete Morse matchings*** $\mathfrak{M}(X)$ on X has vertices cover relations in $D(X)$ that correspond to regular faces and simplices sets of vertices that correspond to acyclic matchings on $D(X)$.

Since the geometric structure of the simplices, i.e. their points, do not have any obvious meaning for discrete Morse matchings, we will not distinguish between $\mathfrak{M}(X)$ and its face poset $D(\mathfrak{M}(X))$ in our notation.

Proposition 5.3. *Let X be a CW complex and let $\mathcal{A}$ be the Morse arrangement on $\mathbb{R}^X$. Then there is a canonical embedding of posets $\mathfrak{M}(X) \subset \mathcal{L}(\mathcal{A})$.*

Proof. The minimal elements of both posets are in canonical bijection because in both cases they correspond to the cover relations in $D(X)$ that correspond to regular face relations. Then collections of cover relations that form an acyclic matching in $\mathfrak{M}(X)$ are mapped to intersections of hyperplanes in $\mathcal{L}(\mathcal{A})$ that correspond to the same acyclic matching. □

[8]Recall Definition 3.21 for the definition of non-degenerate maps.

5.2. The parameter space of merge trees. In order to give a model for the parameter space of merge trees, we introduce a notion of edit morphisms to relate different combinatorial merge trees to each other.

DEFINITION 5.4. An ***edit move*** on a poset of merge tree type P is one of the following operations:

The identity:
: The poset P is left unchanged.

Adding a leaf:
: A new minimal element x is added in one of two ways: either, we just introduce the relation $x \leq y$ for a pre-existing non-minimal element $y \in P$, or we add a new non-minimal element y' between two pre-existing elements $z_1 \leq y \leq z_2$ and add the relation $x \leq y$. As an exceptional case, if P is the one element poset $\{r\}$, in particular, there are no inner nodes, we allow to add two leaves $x < r, y < r$ simultaneously.

Removing a leaf:
: A pre-existing minimal element $x \in P$ is removed. If xs parent node y then only has one child node x', we also remove y.

Splitting an inner node
: If a node $z \in P$ has at least 3 child nodes, z can be split into two nodes $z_1 \leq z_2$ such that there is no $x \in P$ with $z_1 \leq x \leq z_2$. The previous child nodes of z are then distributed among z_1 and z_2 such that z_1 gets at least two child nodes and z_2 gets at least one child node other than z_2.

Merging two adjacent inner nodes:
: If there are two non-minimal elements $z_1 \leq z_2 \in P$ such that there is no element between them, i.e. there exists no $x : z_1 \leq x \leq Z_2$, we construct the quotient $P/_{z_1=z_2}$. This creates an element z instead of z_1 and z_2 such that all previous child nodes of z_1 and z_2 become child nodes of z.

An ***edit morphism*** $\eta\colon P \to P'$ between posets of merge tree type is a composition of edit moves together with an automorphism of P'. An edit morphism is called ***elementary*** if it performs only one of the edit moves. An edit morphism is called ***trivial*** if it has the identity as its only corresponding edit move.

Remark 5.5. Let P, P' be posets of merge tree type such that P' arises from P by either an addition of a leaf or a splitting of an inner node. Then any corresponding edit morphism induces a poset morphism $\eta\colon P \to P'$ given by the inclusion of the preexisting nodes.

Additionally, we want to remark that for each finite poset of merge tree type P, there is only a finite number of elementary edit moves that can be applied to P. Those are:

(1) Adding one leaf to any inner node or edge,
(2) Removing at most one leaf per leaf (in general not possible for all leaves.)

We identify spaces of discrete functions, as introduced in Definition 3.1, with sets of set maps $\mathrm{Map}(P, \mathbb{R})$ for posets P of CW type. Moreover, we identify sets of merge trees with the same underlying combinatorial merge tree P with sets of order-preserving maps $\mathrm{Map}_{\leq}(P, \mathbb{R})$ for posets P of merge tree type. We use the following model[9] for the class of posets of merge tree type:

DEFINITION 5.6. We construct the ***set of combinatorial merge trees*** Mer_{comb} as follows: Let $n \in \mathbb{N}_{\geq 1}$. We choose a set, denoted by Mer^n_{comb}, of representatives for posets of merge tree type with root node 1 on the set $\{1, 2, \ldots, n\}$. That is, for every poset of merge tree type P with n elements, there is exactly one element of Mer^n_{comb} that has the same

[9]By model we mean a set that has one specific poset of merge tree typ for every isomorphism class of posets of merge tree type.

isomorphism type as P. This is possible because partial orders on $\{1,\dots,n\}$ are a subset of $\{1,\dots,n\}\times\{1,\dots,n\}$. We define

$$Mer_{comb} \coprod_{n\in\mathbb{N}_{\geq 1}} Mer^n_{comb}.$$

We do **not** claim that the choice of representatives in Mer_{comb} is canonical in any way - **it is not**. However, this is fine for our purposes, as we see later on.

Definition 5.7. We define the ***set of merge trees*** Mer as the disjoint union

$$Mer := \coprod_{P\in Mer_{comb}} \mathrm{Map}_{\leq}(P,\mathbb{R}).$$

For a merge tree $\theta \in Mer$ we denote the underlying combinatorial merge tree, i.e. the source of θ, by $s(\theta)$.

Moreover, we define the ***set of strict merge trees*** Mer_{st} by

$$Mer_{st} := \coprod_{PMer_{comb}} \mathrm{Map}_{<}(P,\mathbb{R}),$$

i.e. we here demand that the height functions on the combinatorial merge trees are strictly monotone.

Furthermore, we define the ***set of well-branched merge trees*** $Mer_{wb} \subset Mer_{st}$ by

$$Mer_{wb} := \{\theta \in Mer_{st} | \forall y \in s(\theta) \text{ the restriction } \theta_{|s(\theta)_{\leq y}} \text{ has a minimum.}\},$$

where $s(\theta)_{\leq y}$ denotes the rooted subtree of $s(\theta)$ with root y.

Definition 5.8. Let X be a regular CW complex. We define a map $M\colon \mathcal{M}(X) \to Mer$ as follows: Let $f \in \mathcal{M}(X)$ be a discrete Morse function and let $\mathrm{Cr}_i(f)$ the set of critical cells of dimension i of f. We create a poset of merge tree type $\tilde{P}$ as follows: for each critical vertex $\sigma \in \mathrm{Cr}_0(f)$ we add an element σ to $\tilde{P}$. For each time connected components $\sigma_1,\dots,\sigma_k$ of sublevel complexes merge within the filtration, we add an element τ to $\tilde{P}$ and add the relations $\sigma_i \leq \tau$. The elements τ correspond to certain critical edges $\tau \in \mathrm{Cr}_1(f)$. Moreover, we identify $\tilde{P}$ with the unique representative $P \in Mer_{comb}$ of the isomorphism type of $\tilde{P}$. Then we define $M(f)(\sigma) := f_\sigma$, i.e. the elements of P get their functions values under $M(f)$ from the critical values of the corresponding critical cells.

Proposition 5.9. *The image of $\mathcal{MB}_w(X)$ is contained in Mer_{wb}.*

Proof. Since weak Morse–Benedetti functions cannot obtain the same value on different critical cells, the even stronger statement holds that $M(\mathcal{MB}_w(X))$ only contains injective maps for any regular CW complex X. □

We proceed by considering how crossing hyperplanes of $\mathcal{A}$ and $\mathcal{H}$ relates to edit moves between the induced combinatorial merge trees.

Lemma 5.10. *Let X be a regular CW complex and let H^τ_σ be a hyperplane of $\mathcal{A}(X)$. If $\dim(\sigma) = 0$, then passing from a Morse region R_+ on the positive side of H^τ_σ to a Morse region R_- on the negative side without crossing any other hyperplane of $\mathcal{H}$ corresponds to the removal of a leaf. Moreover, passing through H^τ_σ the other way around corresponds to adding a leaf.*

Remark 5.11. We observe that a crossing as mentioned in Lemma 5.10 does not exist for all Morse regions on the positive side of H^τ_σ. Even if it is possible to cancel σ and τ without violating the Morse condition, it might happen that one has to cross other regions of $\mathcal{H}$ or even other Morse regions along the way.

5. RELATIONSHIP TO OTHER CONCEPTS IN PERSISTENT TOPOLOGY

Proof of Lemma 5.10. Since, by assumption, both R_+ and R_- are Morse regions and it is possible to pass from R_+ to R_- without crossing any hyperplane of $\mathcal{H}$ other than H_σ^τ, this means that for any Morse function $f \in R_+$, we have $f(\sigma) < f(\tau)$ and there is no cell with a function value between them. Since $\dim(\sigma) = 0$, at filtration level $f(\sigma)$ a new connected component arises, which is merged with another connected component at level $f(\tau)$ via τ. Because passing to R_- corresponds to matching σ and τ, moving to R_- removes the leaf of $M(f)$ that corresponds to σ. It is immediate that passing from R_- to R_+ adds a new leaf to the induced merge tree. □

Remark 5.12. It is well known, e.g. Forman 1998, Theorem 11.1, that whenever a Morse function $f \in \mathcal{M}(X)$ induces a unique gradient path between a critical d cell σ and a critical $d+1$ cell τ such that all face relations along the gradient path are regular, the gradient path can be inverted, matching σ and τ in the process. If σ and τ have consecutive function values, then σ is a face of τ and we are in the situation of Lemma 5.10. Otherwise, there are other cells with function values between $f(\sigma)$ and $f(\tau)$ and we have to cross hyperplanes of $\mathcal{A}$ corresponding to all face relations along said unique gradient path between σ and τ as well as hyperplanes of $\mathcal{H} \setminus \mathcal{A}$ corresponding to cells that are not contained in the gradient path with function values between $f(\sigma)$ and $f(\tau)$. It should be remarked that during this procedure, we cannot cross all hyperplanes of $\mathcal{A}$ corresponding to the face relations along the unique gradient path simultaneously because otherwise we would leave the space of discrete Morse functions.

Lemma 5.13. *Let X be a regular CW complex, let $\sigma, \tau \in X$ be two 1 cells, and let H_σ^τ be the corresponding hyperplane in $\mathcal{H}$. Let $f \in \mathcal{M}$ be a discrete Morse function inside a region R_1 of $\mathcal{H}$ such that σ and τ are critical and merge connected components, i.e. correspond to inner nodes of $M(f)$. Assume there is a region R_2 adjacent to R_1 via $H_\sigma^\tau \in \mathcal{H}$, i.e. there is a path from R_1 to R_2 that only crosses H_σ^τ and no other hyperplane of $\mathcal{H}$. If*

(1) σ and τ correspond to adjacent inner nodes of $M(f)$, then moving through H_σ^τ first merges the corresponding nodes and then splits them the other way around.

(2) the nodes corresponding to σ and τ are not comparable in $M(f)$, passing through H_σ^τ corresponds to moving their associated values past each other without changing the underlying combinatorial merge tree.

Remark 5.14. In the context of Lemma 5.13, if the nodes that correspond to σ and τ are comparable but not adjacent in $M(F)$, then it is not possible to cross H_σ^τ without crossing other hyperplanes of $\mathcal{H}$ because then there are other nodes with values in between due to the monotonicity of values associated to nodes of $M(f)$.

Proof of Lemma 5.13. (1) If σ and τ correspond to adjacent inner nodes of the induced merge tree, then σ and τ subsequently merge three previously disconnected connected components into one. Crossing H_σ^τ means that the components are merged simultaneously while inside H_σ^τ and the merging order is reversed after exiting H_σ^τ on the other side. This corresponds by definition to first merging the nodes and then splitting them the other way around.

(2) If the nodes corresponding to σ and τ are not comparable then σ and τ belong to different connected components which are, if at all, merged at higher level via a third 1-cell. Hence, crossing H_σ^τ only changes the levels at which σ and τ merge their respective components but not the underlying combinatorial merge tree.

□

Definition 5.15. In order to topologize the space of merge trees Mer, we introduce the ***euclidean edit distance***: We first define the metric for merge trees, whose underlying combinatorial merge trees are related by compositions of elementary edit morphisms (EEM) in the same direction, i.e. either a composition of adding leaves and splitting inner nodes or removing leaves and merging inner nodes:

Let $\theta \in \mathrm{Map}_{\leq}(P, \mathbb{R}), \theta' \in \mathrm{Map}_{\leq}(P', \mathbb{R})$ be merge trees such that P and P' are related by a composition of EEMs in the same direction $\eta\colon P \to P'$. Then by Remark 5.5 we have an induced morphism $\eta\colon P \to P'$ if η is a composition of adding leaves and splitting inner nodes. We extend the inclusion $\eta\colon P \to P'$ associated to any composition of adding leaves and splitting inner nodes to a map $\eta_*\colon \mathrm{Map}_{\leq}(P, \mathbb{R}) \times \mathrm{Map}_{\leq}(P', \mathbb{R}) \to \mathrm{Map}_{\leq}(P', \mathbb{R})$ by

$$(1) \qquad \eta_*(\theta, \theta')_x := \begin{cases} \theta_{\eta^{-1}(x)} \text{ if } x \in \eta(P), \\ \theta_{\eta^{-1}(y)} \text{ if } x \notin \eta(P),\ x \text{ a leaf, } y \in \eta(P) \text{ is } x\text{'s parent in } P' \\ \theta'_x \text{ if } x \notin \eta(P) \text{ was added as a new parent node for a new leaf} \\ \theta'_y \text{ if } x \notin \eta(P),\ x \text{ a leaf, } y \notin \eta(P) \text{ is } x\text{'s parent in } P' \\ \theta'_y \text{ if } x \notin \eta(P),\ x \text{ an inner node } y \text{ the node } x \text{ was split from} \end{cases} .$$

If η is a composition of removing leaves and merging inner nodes, we have an induced inclusion $\eta\colon P' \to P$ in the other direction. Then we define

$$(2) \qquad \mathrm{d}(\theta, \theta') := \mathrm{d}^{euc}(\eta_*(\theta, \theta'), \theta')$$

.

We define the ***euclidean edit distance*** by

$$(3) \qquad \mathrm{d}(\theta, \theta') := \inf_{\sigma}\{\sum_i \mathrm{d}(\theta_i, \theta_{i+1})\},$$

where $\sigma = (\theta = \theta_1, \dots, \theta_n = \theta')$ are sequences of merge trees related by compositions of elementary edit morphisms in the same direction.

Proposition 5.16. *The euclidean edit distance* d *from Definition 5.15 is a pseudo-metric on* Mer, $Mer_<$, *and* Mer_{wb}.

Proof. The equality $\mathrm{d}(\theta, \theta) = 0$ holds because the identity is an edit move. Symmetry holds because every edit move has an inverse: adding leaves is inverse to removing them and splitting inner nodes in inverse to merging them. Sequences of edit morphisms can be inverted stepwise by inverting elementary edit morphisms.

The triangle inequality follows from the triangle equality of the euclidean metric for trivial edit morphisms and by composition of the sequences in Equation (3) otherwise.

□

Remark 5.17. The euclidean edit distance cannot be a metric on Mer because of the possibility of constant functions. Moreover, the euclidean edit distance cannot be a metric on $Mer_<$, and Mer_{wb} either due to symmetries of the underlying combinatorial merge trees. It is possible to fix the latter issue by consideration of ordered underlying combinatorial merge trees, i.e. total order on the sets of children for each inner node. Then it is possible to fix representatives by demanding a suitable notion of compatibility between the orders on the sets of children and function values on the nodes.

Theorem 5.18. *Let* X *be a regular CW complex. Then the map* $M\colon \mathcal{M}(X) \to Mer$ *from Definition 5.8 that maps a discrete Morse function to its induced merge tree is continuous.*

Proof. Let $f_0 \in \mathcal{M}(X)$ be an arbitrary discrete Morse function and let $\epsilon > 0$. We have to show that there is a $\delta > 0$ such that for all $f \in \mathcal{M}(X)$ with $\mathrm{d}(f, f_0) < \delta$ we have $\mathrm{d}(M(f), M(f_0)) < \epsilon$.

For that, we observe that for a fixed discrete Morse function f_0, the map to the induced merge tree is basically an orthogonal projection to the subspace of $\mathbb{R}^X$ which is spanned by the critical cells of dimensions 0 and 1 that correspond to the nodes of the induced merge tree. If f_0 lies in the interior of a region of $\mathcal{H}$, then we can choose δ small enough such that all f with $\mathrm{d}(f, f_0) < \delta$ lie in the same region of $\mathcal{H}(X)$ as f_0. Then all functions in the δ ball $U_\delta(f_0)$ around f_0 induce the same underlying combinatorial merge tree and $M_{|U_\delta(f_0)}$ is just the orthogonal projection, which is continuous.

If f_0 lies on any of the hyperplanes of $\mathcal{H}$, then Lemma 5.10 and Lemma 5.13 describe how perturbations of f_0 affect the induced merge tree. That is, all merge trees in an ϵ neighborhood around $M(f_0)$ that are induced by any discrete Morse function on X are realized via a small perturbation of f_0. This is because merge trees in $U_\epsilon(f_0)$ are related to f_0 by edit moves that correspond to crossing the hyperplanes that f_0 lies on. In this context, it is no problem that a priori $M(f_0)$ admits more edit moves than $\mathcal{H}$ has hyperplanes because by Lemma 5.10 and Lemma 5.13, only the edit moves that correspond to a crossing a hyperplanes in the space of discrete Morse functions actually lead to merge trees that have preimages under the induced merge tree map. □

We define three hyperplane arrangements on the space of merge trees, which are related to the Morse arrangement and the braid arrangement on the space of Morse functions on any given CW complex:

Definition 5.19. Let P be a poset of merge tree type. We define the ***leaf arrangement*** $\mathcal{A}(P)$ in $\mathrm{Map}_\leq(P, \mathbb{R})$ by $\mathcal{A}(P) \coloneqq \{H_x^y | x \text{ is a leaf and } y \text{ is the parent node of } x\}$.

We define the ***order arrangement*** $\mathcal{O}(P)$ by $\mathcal{O}(P) \coloneqq \{H_x^y | x \leq y\}$.

Moreover, we define the ***braid arrangement*** $\mathcal{H}(P)$ in $\mathrm{Map}_\leq(P, \mathbb{R})$ by $\mathcal{H}(P) \coloneqq \{H_x^y | x, y \in P\}$.

Proposition 5.20. *Let P be a poset of merge tree type and let $f\colon X \to \mathbb{R}$ a discrete Morse function on a CW complex X such that P is isomorphic to the underlying combinatorial merge tree of $M(f)$. Then the following holds:*

(1) The hyperplanes of $\mathcal{A}(P)$ correspond to families of maximal gradient paths between critical 0 and 1 cells induced by f. If f is weakly Morse–Benedetti, then this correspondence is bijective.

(2) The hyperplanes of $\mathcal{O}(P)$ correspond to a collection of zig-zags of maximal gradient paths between cells of dimension ≤ 1 induced by f that begin and end at maximal critical cells of sublevel complexes without leaving the sublevel complex corresponding to the higher of the two cells.

(3) The hyperplanes of $\mathcal{H}(P)$ correspond to a collection of zig-zags of maximal gradient paths between cells of dimension ≤ 1 induced by f.

In particular $\mathcal{A}(P) \subset \mathcal{O}(P) \subset \mathcal{H}(P)$ holds.

Proof. (1) By Definition 5.8, the leaves of P correspond to critical 0 cells of f. Moreover, the parent relation $x \prec y$ of any leaf x corresponds to the collection of critical 1 cells that merge the connected component that corresponds to x, say $X_{M(f)_y - \varepsilon}[\sigma]$, where σ is the unique critical 0 cell, with the connected components that correspond to y's other child nodes. Let τ be one of these merging 1 cells. Then one of τ's boundary 0 cells σ_0 belongs to $X_{M(f)_y - \varepsilon}[x]$ and there is a gradient path γ in $X_{M(f)_y}[\tau]$ that begins with $\tau \supset \sigma_0$. Since maximal gradient paths lead to critical cells and cannot lead through cells of increasing dimension, γ has to lead to σ. If f is weakly Morse–Benedetti, then each inner node y corresponds to a unique merging 1 cell τ and because, by assumption, only one of τ's boundary 0 cells σ_0 belongs to $X_{M(f)_y - \varepsilon}[x]$, the gradient path starting at $\tau \sup \sigma_0$ is unique because it is contained in the 1 skeleton of X.

(2) The hyperplanes of $\mathcal{O}(M(f)$ correspond by definition to inclusions of connected components sublevel complexes of f. Similar to the proof of (1), cover relations

$x \prec y$ of nodes in $M(f)$ correspond to gradient paths from merging 1 cells that correspond to y to any critical 0 cell of the connected component that corresponds to x. If $x \leq y$ is not a cover relation, then the statement follows inductively because the interval $[x, y]$ is a chain due to P being of merge tree type.

(3) Since posets of merge tree type have a unique maximum, namely the root, by definition, all elements of P are related by zig-zags of comparison relations. It follows from (2) that cells that correspond to elements of P are related by zig-zags of zig-zags of gradient paths, which are zig-zags of gradient paths.

□

Remark 5.21. Since discrete gradient paths are sequences of matched cells, gradient paths correspond to intersections of hyperplanes of $\mathcal{A}(X)$. Even though this observation makes the correspondence between families of gradient paths and hyperplanes in Proposition 5.20 (1) into a correspondence between certain intersections of hyperplanes in $\mathcal{A}(X)$ and hyperplanes in $\mathcal{A}(M(f)$, that does not in general lead to a correspondence between cancellations of gradient paths and removals of leaves.

Corollary 5.22. *If $f: X \to \mathbb{R}$ is a weak Morse–Benedetti function, then cancellations of critical 0 cells and 1 cells are in 1-1 correspondence with removals of leaves of the induced merge tree. If f is not weakly Morse–Benedetti, cancellations of critical 0 cells and 1 cells still induce removals of leaves of the induced merge tree but not necessarily the other way around.*

Corollary 5.23. *Let P be a poset of merge tree type such that $P \cong M(f)$ for a weak Morse–Benedetti function $f: X \to \mathbb{R}$, then $M^{-1}(\mathcal{A}(P)) \subset \mathcal{A}(X)$. Moreover, $M^{-1}(\mathcal{A}(P))$ agrees with $\mathcal{A}(X)$ on the 0 skeleton of X.*

5.3. The parameter space of barcodes. We identify sets of barcodes with sets of order-preserving maps $\mathrm{Map}_{\leq}(P, \mathbb{R})$ for posets P of barcode type.

DEFINITION 5.24. We define the ***set of barcodes*** by

$$Bar := \coprod_{P \text{ of barcode type}} \mathrm{Map}_{\leq}(P, \mathbb{R}).$$

Remark 5.25. Technically, we have to solve the same set theoretic issues as for the space of merge trees again for the space of barcodes. They can be solved in the same way as for merge trees.

DEFINITION 5.26. Let P, P' be two posets of barcode type.

(1) If $P \cong P'$ as posets, we call an isomorphism $\eta: P \to P'$ a ***reordering*** of the bars. We call id: P → P the ***trivial reordering***.
(2) We say that P' arises from P' by ***collapsing a bar*** if P' has one bar fewer than P. In that case, a ***collapse*** of a bar is an injective poset map $\eta: P' \to P$. Moreover, we say that P arises from P' by ***creation*** of a new bar.

We refer to all of the mentioned operations as ***edit moves***. An ***elementary edit morphism*** between posets of barcode type is a composition of an edit move with a reordering of the bars. An ***edit morphism*** is a composition of elementary edit morphisms.

Associated to an edit morphism $\eta: P \to P'$, we define a map $\eta_*: \mathrm{Map}_{\leq}(P, \mathbb{R}) \times \mathrm{Map}_{\leq}(P', \mathbb{R}) \to \mathrm{Map}_{\leq}(P', \mathbb{R})$ by

$$\eta_*(\beta, \beta')_x := \begin{cases} \theta_x & \text{if } x \in P, \\ \theta'_x & \text{if } x \notin P \end{cases}.$$

DEFINITION 5.27. The ***euclidean edit distance*** on Bar is defined as follows:

5. RELATIONSHIP TO OTHER CONCEPTS IN PERSISTENT TOPOLOGY

Let P, P' be posets of barcode type. Let $\beta_1 \in \mathrm{Map}_{\leq}(P, \mathbb{R}), \beta_2 \in \mathrm{Map}_{\leq}(P', \mathbb{R})$. If P' arises from P by an edit morphism $\eta\colon P \to P'$, we define $\mathrm{d}(\beta_1, \beta_2) \coloneqq \mathrm{d}^{\mathrm{euc}}(\eta_*(\beta, \beta'), \beta')$ In the general case, we define the ***euclidean edit distance*** by

$$\mathrm{d}(\beta, \beta') \coloneqq \inf_{\sigma}\{\sum_i \mathrm{d}(\beta_i, \beta_{i+1})\},$$

where $\sigma = (\beta = \beta_1, \ldots, \beta_n = \beta')$ are sequences of barcodes related by edit morphisms.

Proposition 5.28. *The euclidean edit distance is a pseudo metric on Bar*

Proof. The proof is analogous to the one of Proposition 5.16. □

As for merge trees, the euclidean edit distance on barcodes cannot be a metric due to symmetries. Again, it is possible to deal with the symmetries by using ordered barcodes in order to choose specific representatives.

We topologize the space of barcodes with the euclidean edit distance.

Definition 5.29. Let P be a poset of barcode type. We define the ***order arrangement*** $\mathcal{O}(\mathcal{P})$ in $\mathrm{Map}_{\leq}(P, \mathbb{R})$ by $\mathcal{O}(P) \coloneqq \{H_x^y | x \leq y\}$.

Moreover, we define the ***braid arrangement*** $\mathcal{H}(P)$ in $\mathrm{Map}_{\leq}(P, \mathbb{R})$ by $\mathcal{H}(P) \coloneqq \{H_x^y | x, y \in P\}$.

In order to define a map from the space of strict merge trees to the space of barcodes, we apply the combinatorial elder rule, see Curry et al. 2024, Definition 2.15

Definition 5.30. We define a map $B\colon Mer_{wb} \to Bar$ as follows:

Let $\theta \in Mer_{wb}$ be a well-branched merge tree with an underlying combinatorial merge tree P. We define a poset $\tilde{B}(\theta)$ inductively: Add for each inner node y of P an interval $[x, y]$ to $\tilde{B}(\theta)$, where x denotes leaf of $P_{\leq y}$ with the minimal value θ_x. For all leaf nodes x that are not matched by the rule above, we add an interval $[x, y]$ to $\tilde{B}(\theta)$ where y is x's parent node.

By construction, we have a canonical map $i\colon \tilde{B}(\theta) \to P$ given by mapping the elements of $\tilde{B}(\theta)$ to their corresponding nodes in P. We define $B(\theta) \coloneqq i^*(\theta) = \theta \circ i$.

Proposition 5.31. *The map $B\colon Mer_{wb} \to Bar$ from Definition 5.30 is continuous.*

Proof. Let $\theta_0 \in Mer_{wb}$ be an arbitrary well-branched merge tree and let $\epsilon > 0$. We have to show that there is a $\delta > 0$ such that for all $\theta \in Mer_{wb}$ with $\mathrm{d}(\theta, \theta_0) < \delta$ we have $\mathrm{d}(B(\theta), B(\theta_0)) < \epsilon$.

Since θ_0 is by assumption strict, it must lie in the interior of some region of $\mathcal{O}(s(\theta))$. Thus, δ can be chosen small enough such that all θ with $\mathrm{d}(\theta, \theta_0) < \delta$ have the same underlying combinatorial merge tree $s(\theta)$ and induce the same combinatorial barcode $\tilde{B}(\theta)$. Therefore, $B_{|U_\delta(\theta_0)}$ is a linear map between finite-dimensional vectorspaces and, therefore, continuous. □

Bibliography

Aguiar, Marcelo and Swapneel Mahajan (2017). *Topics in Hyperplane Arrangements*. Vol. 226. Mathematical Surveys and Monographs. American Mathematical Society.

Banyaga, A. and D. Hurtubise (2004). *Lectures on Morse Homology*. Texts in the Mathematical Sciences. Springer Netherlands. ISBN: 9781402026959.

Benedetti, Bruno (2016). "Smoothing discrete Morse theory". In: *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 16, pp. 335–368.

Björner, Anders et al. (1999). *Oriented Matroids*. Second. Encyclopedia of Mathematics and its Applications. Cambridge University Press. ISBN: 13 978-0-521-77750-6.

Brück, Benjamin and Adélie Garin (2023). "Stratifying the space of barcodes using Coxeter complexes". In: *J. Appl. and Comput. Topology* 7, pp. 369–395. DOI: https://doi.org/10.1007/s41468-022-00104-7.

Brüggemann, Julian (2023). *On the moduli space of discrete Morse functions*. arXiv: 2312.14032 [math.AT].

Capitelli, Nicolas Ariel and Elias Gabriel Minian (2017). "A Simplicial Complex is Uniquely Determined by Its Set of Discrete Morse Functions". In: *Discrete and Computational Geometry* 58, pp. 144–157. DOI: https://doi.org/10.1007/s00454-017-9865-z.

Catanzaro, Michael J. et al. (2020). "Moduli spaces of morse functions for persistence". In: *J Appl. and Comput. Topology* 4, pp. 353–385. DOI: https://doi.org/10.1007/s41468-020-00055-x.

Cerf, Jean (Feb. 1970). "La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie". In: *Publications Mathematiques de L'Institut des Hautes Scientifiques* 39, pp. 7–170. DOI: 10.1007/BF02684687.

Chari, Manoj K. and Michael Joswig (2005). "Complexes of discrete Morse functions". In: *Discrete Mathematics* 302.1. Structural Combinatorics, pp. 39–51. ISSN: 0012-365X. DOI: https://doi.org/10.1016/j.disc.2004.07.027. URL: https://www.sciencedirect.com/science/article/pii/S0012365X05002906.

Curry, Justin et al. (2024). "From trees to barcodes and back again II: Combinatorial and probabilistic aspects of a topological inverse problem". In: *Computational Geometry* 116, p. 102031. ISSN: 0925-7721. DOI: https://doi.org/10.1016/j.comgeo.2023.102031.

Cyranka, Jacek, Konstantin Mischaikow, and Charles Weibel (2020). "Contractibility of a persistence map preimage". In: *J Appl. and Comput. Topology* 4, pp. 509–523. DOI: Contractibilityofapersistencemappreimage.

Forman, Robin (Mar. 1998). "Morse theory for cell complexes". In: *Adv. Math.* 134.1, pp. 90–145. ISSN: 0001-8708. DOI: 10.1006/aima.1997.1650.

Franks, John M. (1979). "Morse-Smale flows and homotopy theory". In: *Topology* 18.3, pp. 199–215. ISSN: 0040-9383. DOI: https://doi.org/10.1016/0040-9383(79)90003-X.

Hatcher, Allen E. and John B. Wagoner (1973). "Pseudo-isotopies of compact manifolds". In: *Asterisque* 6. DOI: 10.1007/BF02684687.

Leygonie, Jacob and Ulrike Tillmann (2022). "The fiber of persistent homology for simplicial complexes". In: *Journal of Pure and Applied Algebra* 226.12. ISSN: 0022-4049. DOI: https://doi.org/10.1016/j.jpaa.2022.107099. URL: https://www.sciencedirect.com/science/article/pii/S0022404922000950.

Lin, Maxwell and Nicholas A. Scoville (2021). "On the automorphism group of the Morse complex". In: *Advances in Applied Mathematics* 131, p. 102250. ISSN: 0196-8858. DOI: https://doi.org/10.1016/j.aam.2021.102250.

Matsumoto, Yukio (2002). *An Introduction to Morse Theory*. Vol. 208. American Mathematical Society. ISBN: 978-0-8218-1022-4.

Milnor, John (1965). *Lectures on the h-Cobordism Theorem*. Princeton: Princeton University Press. ISBN: 9781400878055. DOI: doi:10.1515/9781400878055. URL: https://doi.org/10.1515/9781400878055.

Sharko, Vladimir V. (1993). *Functions on manifolds: algebraic and topological aspects*. Vol. 131. Translations of Mathematical Monographs. American Mathematical Society.

Wetzels, F. and C. Garth (Oct. 2022). "A Deformation-based Edit Distance for Merge Trees". In: *2022 Topological Data Analysis and Visualization (TopoInVis)*. Los Alamitos, CA, USA: IEEE Computer Society, pp. 29–38. DOI: 10.1109/TopoInVis57755.2022.00010. URL: https://doi.ieeecomputersociety.org/10.1109/TopoInVis57755.2022.00010.

Bibliography

Adiprasito, Karim and Bruno Benedetti (2019). *Collapsibility of CAT(0) spaces*. arXiv: 1107.5789 [math.MG].

Aguiar, Marcelo and Swapneel Mahajan (2017). *Topics in Hyperplane Arrangements*. Vol. 226. Mathematical Surveys and Monographs. American Mathematical Society.

Alexandroff, Pavel (1937). "Diskrete Räume". In: *Recueil Mathémathique* 2(44).3, pp. 501–519. URL: %7Bhttps://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=sm&paperid=5579&option%5C_lang=eng%7D.

Alexandrov, Pawel S. (1956). *Combinatorial Topology*. Graylock Press.

Arone, Gregory Z. and D. Lukas B. Brantner (2021). "The action of Young subgroups on the partition complex". In: *Publationes mathématiques de l'IHÉS* 133, pp. 47–156. DOI: 10.1007/s10240-021-00123-7.

Ballester, Rubén, Carles Casacuberta, and Sergio Escalera (2024). *Topological Data Analysis for Neural Network Analysis: A Comprehensive Survey*. arXiv: 2312.05840 [cs.LG].

Banyaga, A. and D. Hurtubise (2004). *Lectures on Morse Homology*. Texts in the Mathematical Sciences. Springer Netherlands. ISBN: 9781402026959.

Baryshnikov, Yuliy (2019). "Time Series, Persistent Homology and Chirality". In: arXiv:1909.09846 [math.PR]. arXiv: arXiv:1909.09846 [math.PR].

Benedetti, Bruno (2016). "Smoothing discrete Morse theory". In: *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 16, pp. 335–368.

Bickle, Allan (2020). *Fundamentals of graph theory*. Vol. 43. Pure and Applied Undergraduate Texts. American Mathematical Society, Providence, RI, pp. xv+336. ISBN: 978-1-4704-5342-8.

Björner, Anders et al. (1999). *Oriented Matroids*. Second. Encyclopedia of Mathematics and its Applications. Cambridge University Press. ISBN: 13 978-0-521-77750-6.

Botnan, Magnus Bakke and Michael Lesnick (2023). *An Introduction to Multiparameter Persistence*. arXiv: 2203.14289 [math.AT].

Brück, Benjamin and Adélie Garin (2023). "Stratifying the space of barcodes using Coxeter complexes". In: *J. Appl. and Comput. Topology* 7, pp. 369–395. DOI: https://doi.org/10.1007/s41468-022-00104-7.

Brüggemann, Julian (Nov. 2022). "On Merge Trees and Discrete Morse Functions on Paths and Trees". In: *J Appl. and Comput. Topology*. DOI: https://doi.org/10.1007/s41468-022-00101-w.

— (2023). *On the moduli space of discrete Morse functions*. arXiv: 2312.14032 [math.AT].

Brüggemann, Julian and Nicholas A. Scoville (2023). *On cycles and merge trees*. arXiv: 2301.01316 [math.AT].

Capitelli, Nicolas Ariel and Elias Gabriel Minian (2017). "A Simplicial Complex is Uniquely Determined by Its Set of Discrete Morse Functions". In: *Discrete and Computational Geometry* 58, pp. 144–157. DOI: https://doi.org/10.1007/s00454-017-9865-z.

Cardona, Robert et al. (2022). "The universal ℓ^p-metric on merge trees". In: *38th International Symposium on Computational Geometry*. Vol. 224. LIPIcs. Leibniz Int. Proc. Inform. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, Art. No. 24, 20. DOI: 10.4230/lipics.socg.2022.24.

Carlsson, Gunnar (Apr. 2009). “Topology and Data”. In: *Bulletin of The American Mathematical Society - BULL AMER MATH SOC* 46, pp. 255–308. DOI: 10.1090/S0273-0979-09-01249-X.

Carlsson, Gunnar and Mikael Vejdemo-Johansson (2022). *Topological data analysis with applications*. Cambridge University Press, Cambridge, pp. xi+220. ISBN: 978-1-108-83865-8. DOI: 10.1017/9781108975704.

Carr, Hamish, Jack Snoeyink, and Ulrike Axen (2003). “Computing contour trees in all dimensions”. In: *Computational Geometry: Theory and Applications*, pp. 75–94.

Carr, Hamish, Jack Snoeyink, and Michiel van de Panne (2010). “Flexible isosurfaces: Simplifying and displaying scalar topology using the contour tree”. In: *Computational Geometry* 43.1. Special Issue on the 14th Annual Fall Workshop, pp. 42–58. ISSN: 0925-7721. DOI: https://doi.org/10.1016/j.comgeo.2006.05.009. URL: https://www.sciencedirect.com/science/article/pii/S0925772109000455.

Catanzaro, Michael J. et al. (2020). “Moduli spaces of morse functions for persistence”. In: *J Appl. and Comput. Topology* 4, pp. 353–385. DOI: https://doi.org/10.1007/s41468-020-00055-x.

Cerf, Jean (Feb. 1970). “La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie”. In: *Publications Mathematiques de L’Institut des Hautes Scientifiques* 39, pp. 7–170. DOI: 10.1007/BF02684687.

Chari, Manoj K. and Michael Joswig (2005). “Complexes of discrete Morse functions”. In: *Discrete Mathematics* 302.1. Structural Combinatorics, pp. 39–51. ISSN: 0012-365X. DOI: https://doi.org/10.1016/j.disc.2004.07.027. URL: https://www.sciencedirect.com/science/article/pii/S0012365X05002906.

Chazal, Frédéric and Bertrand Michel (2021). “An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists”. In: *Frontiers in Artificial Intelligence* 4. ISSN: 2624-8212. DOI: 10.3389/frai.2021.667963. URL: https://www.frontiersin.org/articles/10.3389/frai.2021.667963.

Curry, Justin (Feb. 2019). “The fiber of the persistence map for functions on the interval”. In: *J Appl. and Comput. Topology*, pp. 301–321.

Curry, Justin, Jordan DeSha, et al. (2024). “From trees to barcodes and back again II: Combinatorial and probabilistic aspects of a topological inverse problem”. In: *Computational Geometry* 116, p. 102031. ISSN: 0925-7721. DOI: https://doi.org/10.1016/j.comgeo.2023.102031.

Curry, Justin, Haibin Hang, et al. (2022). “Decorated merge trees for persistent topology”. In: *J. Appl. Comput. Topol.* 6.3, pp. 371–428. ISSN: 2367-1726. DOI: 10.1007/s41468-022-00089-3.

Cyranka, Jacek, Konstantin Mischaikow, and Charles Weibel (2020). “Contractibility of a persistence map preimage”. In: *J Appl. and Comput. Topology* 4, pp. 509–523. DOI: Contractibilityofapersistencemappreimage.

Delfinado, Cecil Jose A. and Herbert Edelsbrunner (1993). “An incremental algorithm for Betti numbers of simplicial complexes”. In: *Proceeding SCG ’93 Proceedings of the ninth annual symposium on Computational geometry*. Springer International Publishing, pp. 232–239.

— (1995). “An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere”. In: *Computer Aided Geometric Design* 12.7. Grid Generation, Finite Elements, and Geometric Design, pp. 771–784. ISSN: 0167-8396. DOI: https://doi.org/10.1016/0167-8396(95)00016-Y. URL: https://www.sciencedirect.com/science/article/pii/016783969500016Y.

Dłotko, Paweł and Hubert Wagner (2012). *Computing homology and persistent homology using iterated Morse decomposition*. arXiv: 1210.1429 [math.AT].

Edelsbrunner, Herbert, David Letscher, and Afra Zomorodian (2002). "Topological Persistence and Simplification". In: *Discrete Comput Geom* 28, pp. 511–533. DOI: https://doi.org/10.1007/s00454-002-2885-2.

Elbers, Willem and Rien van de Weygaert (Jan. 2023). "Persistent topology of the reionization bubble network - II. Evolution and classification". In: *Monthly Notices of the Royal Astronomical Society* 520.2, pp. 2709–2726. ISSN: 0035-8711. DOI: 10.1093/mnras/stad120.

Engelke, Wito et al. (2021). "Topology-Based Feature Design and Tracking for Multi-center Cyclones". In: *Topological Methods in Data Analysis and Visualization VI*. Ed. by Ingrid Hotz et al. Cham: Springer International Publishing, pp. 71–85. ISBN: 978-3-030-83500-2.

Ferri, Massimo (2015). "Persistent Topology for Natural Data Analysis – a Survey". In: Lecture Notes in Artificial Intelligence 10344, pp. 117–133.

Forman, Robin (Mar. 1998). "Morse theory for cell complexes". In: *Adv. Math.* 134.1, pp. 90–145. ISSN: 0001-8708. DOI: 10.1006/aima.1997.1650.

— (2002). "A user's guide to discrete Morse theory". In: *Sém. Lothar. Combin.* 48, Art. B48c, 35.

Franks, John M. (1979). "Morse-Smale flows and homotopy theory". In: *Topology* 18.3, pp. 199–215. ISSN: 0040-9383. DOI: https://doi.org/10.1016/0040-9383(79)90003-X.

Gasparovic, Ellen et al. (2022). *Intrinsic Interleaving Distance for Merge Trees*. arXiv: 1908.00063 [cs.CG].

Ghrist, Robert (2008). "Barcodes: The Persistent Topology of Data". In: *Bulletin of the American Mathematical Society* 45, pp. 61–75.

Goerss, Paul G. and John F. Jardine (2009). *Simplicial Homotopy Theory*. Modern Birkhäuser Classics. Birkhäuser Basel. ISBN: 978-3-0346-0189-4. DOI: https://doi.org/10.1007/978-3-0346-0189-4.

Gromov, Michail (1987). "Hyperbolic Groups". In: Mathematical Sciences Research Institute Publications 8, pp. 75–264.

Hajebi, Sahab and Ramin Javadi (2023). "On the parameterized complexity of the acyclic matching problem". In: *Theoretical Computer Science* 958, p. 113862. ISSN: 0304-3975. DOI: https://doi.org/10.1016/j.tcs.2023.113862. URL: https://www.sciencedirect.com/science/article/pii/S0304397523001755.

Hatcher, Allen E. and John B. Wagoner (1973). "Pseudo-isotopies of compact manifolds". In: *Asterisque* 6. DOI: 10.1007/BF02684687.

Hausmann, Jean-Claude (1996). "On the Vietoris-Rips Complexes and a Cohomology Theory for Metric Spaces". In: Annals of Mathematics Studies 138, pp. 175–188.

Heine, Christian et al. (July 2016). "A Survey of Topology-based Methods in Visualization". In: *Computer Graphics Forum* 35, pp. 643–667.

Hensel, Felix, Michael Moor, and Bastian Rieck (2021). "A Survey of Topological Machine Learning Methods". In: *Frontiers in Artificial Intelligence* 4. ISSN: 2624-8212. DOI: 10.3389/frai.2021.681108. URL: https://www.frontiersin.org/articles/10.3389/frai.2021.681108.

Hovey, Mark (2007). *Model Categories*. Mathematical surveys and monographs. American Mathematical Society. ISBN: 9780821843611.

Jardine, John F. (2020). *Persistent homotopy theory*. arXiv: 2002.10013 [math.AT].

Johnson, Benjamin and Nicholas A. Scoville (July 2022). "Merge trees in discrete Morse theory". In: *Res. Math. Sci.* 9, Paper No. 49.

Klette, Reinhard and Azriel Rosenfeld (2004). *Digital Geometry*. The Morgan Kaufmann Series in Computer Graphics. San Francisco: Morgan Kaufmann. ISBN: 978-1-55860-861-0. DOI: https://doi.org/10.1016/B978-1-55860-861-0.50027-4.

Kweon, In So and Takeo Kanade (Mar. 1994). "Extracting Topographic Terrain Features from Elevation Maps". In: *CVGIP: Image Understanding* 59, pp. 171–182.

Lewiner, Thomas, Hélio Lopes, and Geovan Tavares (2003a). "Optimal discrete Morse functions for 2-manifolds". In: *Computational Geometry* 26.3, pp. 221–233. ISSN: 0925-7721. DOI: https://doi.org/10.1016/S0925-7721(03)00014-2.

— (2003b). "Toward Optimality in Discrete Morse Theory". In: *Experimental Mathematics* 12.3, pp. 271–285. DOI: 10.1080/10586458.2003.10504498.

Leygonie, Jacob and Ulrike Tillmann (2022). "The fiber of persistent homology for simplicial complexes". In: *Journal of Pure and Applied Algebra* 226.12. ISSN: 0022-4049. DOI: https://doi.org/10.1016/j.jpaa.2022.107099. URL: https://www.sciencedirect.com/science/article/pii/S0022404922000950.

Lin, Maxwell and Nicholas A. Scoville (2021). "On the automorphism group of the Morse complex". In: *Advances in Applied Mathematics* 131, p. 102250. ISSN: 0196-8858. DOI: https://doi.org/10.1016/j.aam.2021.102250.

Liu, Shusen et al. (Dec. 2016). "Visualizing High-Dimensional Data: Advances in the Past Decade". In: *IEEE Transactions on Visualization and Computer Graphics*, pp. 1249–1268.

Lundell, Albert T. and Stephen Weingram (1969). *The Topology of CW Complexes*. The university series in higher mathematics. Springer New York, NY. ISBN: 978-1-4684-6256-2. DOI: doi:10.1007/978-1-4684-6254-8. URL: https://doi.org/10.1007/978-1-4684-6254-8.

Maria, Clément and Hannah Schreiber (2023). "Discrete Morse Theory for Computing Zigzag Persistence". In: *Discrete and Computational Geometry*. DOI: 10.1007/s00454-023-00594-x.

Mathar, Rudolf et al. (2020). *Fundamentals of Data Analytics*. Springer Cham. ISBN: 978-3-030-56831-3.

Matsumoto, Yukio (2002). *An Introduction to Morse Theory*. Vol. 208. American Mathematical Society. ISBN: 978-0-8218-1022-4.

Milnor, John (1965). *Lectures on the h-Cobordism Theorem*. Princeton: Princeton University Press. ISBN: 9781400878055. DOI: doi:10.1515/9781400878055. URL: https://doi.org/10.1515/9781400878055.

Mischaikow, Konstantin and Vidit Nanda (2013). "Morse Theory for Filtrations and Efficient Computation of Persistent Homology". In: *Discrete and Computational Geometry* 50, pp. 330–353. DOI: 10.1007/s00454-013-9529-6.

Morozov, Dmitriy, Kenes Beketayev, and Gunther Weber (2013). "Interleaving distance between merge trees". In: *Discrete and Computational Geometry* 49.22-45, p. 52.

Nanda, Vidit, Dai Tamaki, and Kohei Tanaka (Oct. 2018). "Discrete Morse theory and classifying spaces". In: *Advances in Mathematics* 340, pp. 723–790.

Oesterling, Patrick, Christian Heine, Gunther H. Weber, Dimitry Morozov, et al. (2017). "Computing and Visualizing Time-Varying Merge Trees for High-Dimensional Data". In: *Topological Methods in Data Analysis and Visualization*, pp. 87–101.

Oesterling, Patrick, Christian Heine, Gunther H. Weber, and Gerik Scheuermann (Mar. 2013). "Visualizing nD Point Clouds as Topological Landscape Profiles to Guide Local Data Analysis". In: *IEEE Transactions on Visualization and Computer Graphics* 19, pp. 514–526.

Oudot, Steve Y. (2015). *Persistence theory: from quiver representations to data analysis*. Vol. 209. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, pp. viii+218. ISBN: 978-1-4704-2545-6. DOI: 10.1090/surv/209.

Paolini, Giovanni and Mario Salvetti (2021). "Proof of the $K(\pi, 1)$ conjecture for all affine Artin groups". In: *Inventiones Mathematicae* 224, pp. 487–572. DOI: 10.1007/s00222-020-01016-y.

Polterovich, Leonid et al. (2020). *Topological persistence in geometry and analysis*. Vol. 74. University Lecture Series. American Mathematical Society, Providence, RI, pp. xi+128. ISBN: 978-1-4704-5495-1.

Quillen, Daniel G. (1976). *Homotopical Algebra*. Lecture Notes in Mathematics. Heidelberg: Springer Berlin. ISBN: 978-3-540-03914-3. DOI: https://doi.org/10.1007/BFb0097438.

Rand, Ian and Nicholas A. Scoville (2020). "Discrete Morse functions, vector fields, and homological sequences on trees". In: *Involve* 13.2, pp. 219–229. DOI: DOI:10.2140/involve.2020.13.219.

Robins, Vanessa (2002). "Computational Topology for Point Data: Betti Numbers of α-Shapes". In: *Morphology of Condensed Matter: Physics and Geometry of Spatially Complex Systems*. Ed. by Klaus Mecke and Dietrich Stoyan. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 261–274. ISBN: 978-3-540-45782-4. DOI: {10.1007/3-540-45782-8_11}. URL: %7Bhttps://doi.org/10.1007/3-540-45782-8%5C_11%7D.

Robins, Vanessa, Peter John Wood, and Adrian P. Sheppard (2011). "Theory and Algorithms for Constructing Discrete Morse Complexes from Grayscale Digital Images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.8, pp. 1646–1658. DOI: 10.1109/TPAMI.2011.95.

Sharko, Vladimir V. (1993). *Functions on manifolds: algebraic and topological aspects*. Vol. 131. Translations of Mathematical Monographs. American Mathematical Society.

Shinagawa, Yoshihisa, Tosiyasu L. Kunii, and Yannick L. Kergosien (1991). "Surface Coding Based on Morse Theory". In: *IEEE Computer Graphics and Applications*, pp. 66–78.

Tarasov, Sergey P. and Michael N. Vyalyi (June 1998). "Construction of contour trees in 3D in O(n log n) steps". In: *SCG 98: Proceedings of the fourteenth annual symposium on Computational geometry*, pp. 68–75.

tom Dieck, Tammo (2008). *Algebraic Topology*. Textbooks in Mathematics. European Mathematical Society. ISBN: 978-3-03719-048-7.

Tralie, Christopher J. et al. (2022). *The DOPE Distance is SIC: A Stable, Informative, and Computable Metric on Time Series And Ordered Merge Trees*. arXiv: 2212.01648 [cs.IR].

Tyrtyshnikov, Eugene E. (1997). *A Brief Introduction to Numerical Analysis*. Springer Science+Business Media, LLC. ISBN: 978-1-4612-6413-2.

van Kreveld, Marc et al. (Aug. 1997). "Contour trees and small seed sets for isosurface traversal". In: *SCG 97: Proceedings of the thirteenth annual symposium on Computational geometry*, pp. 212–220.

Vietoris, Leopold (1927). "Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen". In: *Mathematische Annalen* 97, pp. 454–755.

Weber, Gunther H., Peer-Timo Bremer, and Valerio Pascucci (Sept. 2007). "Topological Landscapes: A Terrain Metaphor for Scientific Data". In: *IEEE Transactions on Visualization and Computer Graphics* 13, pp. 1416–1423.

Wetzels, F. and C. Garth (Oct. 2022). "A Deformation-based Edit Distance for Merge Trees". In: *2022 Topological Data Analysis and Visualization (TopoInVis)*. Los Alamitos, CA, USA: IEEE Computer Society, pp. 29–38. DOI: 10.1109/TopoInVis57755.2022.00010. URL: https://doi.ieeecomputersociety.org/10.1109/TopoInVis57755.2022.00010.

Whitehead, J. H. C. (1949). "Combinatorial homotopy. I". In: *Bulletin of the American Mathematical Society* 55.3.P1, pp. 213–245.

Yan, Lin et al. (Aug. 2019). "A Structural Average of Labeled Merge Trees for Uncertainty Visualization". In: *IEEE Transactions on Visualization and Computer Graphics* 26, pp. 832–842.