Topological inference from measures and vector bundles

Raphaël Tinarrage

université
PARIS-SACLAY

Inférence topologique à partir de mesures et de fibrés vectoriels

Thèse de doctorat de l'Université Paris-Saclay

Ecole Doctorale de Mathématique Hadamard (EDMH) n° 574
Spécialité de doctorat : Mathématiques appliquées
Unité de recherche : Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France.
Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Orsay, le 12/10/2020, par

Raphaël TINARRAGE

Composition du jury :

Patrick MASSOT Professeur, Université Paris-Saclay	Président
Simon MASNOU Professeur, Université C. Bernard Lyon 1	Rapporteur & Examinateur
Shmuel WEINBERGER Professor, University of Chicago	Rapporteur & Examinateur
Dominique ATTALI Directrice de recherche, CNRS	Examinatrice
Blanche BUET Maîtresse de Conférences, Université Paris-Saclay	Examinatrice
Frédéric CHAZAL Directeur de Recherche, INRIA Saclay - Île-de-France	Directeur de thèse
Marc GLISSE Chargé de recherche, INRIA Saclay - Île-de-France	Co-encadrant & Examinateur

Thèse de doctorat

NNT: 2020UPASM001


université
PARIS-SACLAY
FACULTÉ
DES SCIENCES
D'ORSAY


MATHEMATIQUES
ORSAY


Fondation mathématique
FMJH
Jacques Hadamard


Inria

Résumé - Abstract

Résumé. Nous contribuons à l'inférence topologique, basée sur la théorie de l'homologie persistante, en proposant trois familles de filtrations. Nous établissons pour chacune d'elles des résultats de consistance – c'est-à-dire de qualité d'approximation d'un objet géométrique sous-jacent –, et de stabilité – c'est-à-dire de robustesse face à des erreurs de mesures initiales. Nous proposons des algorithmes concrets afin de pouvoir utiliser ces méthodes en pratique.

La première famille, les filtrations-DTM, est une alternative robuste à la filtration de Čech habituelle lorsque le nuage de points est bruité ou contient des points aberrants. Elle repose sur la notion de distance à la mesure, qui permet d'obtenir une stabilité au sens de la distance de Wasserstein.

Deuxièmement, nous proposons les filtrations relevées, qui permettent d'estimer l'homologie des variétés immergées, même quand leur portée est nulle. Nous introduisons la notion de portée normale, et montrons qu'elle permet de contrôler des quantités géométriques associées à la variété. Nous étudions l'estimation des espaces tangents par les matrices de covariance locale.

Troisièmement, nous développons un cadre théorique pour les filtrations de fibrés vectoriels, et définissons les classes de Stiefel-Whitney persistantes. Nous montrons que les classes de Stiefel-Whitney persistantes associées aux filtrations de fibrés de Čech sont consistantes et stables en distance de Hausdorff. Pour permettre leur mise en œuvre algorithmique, nous introduisons la notion de condition étoile faible.

Abstract. We contribute to the theory of topological inference, based on the theory of persistent homology, by proposing three families of filtrations. For each of them, we prove consistency results—that is, quality of approximation of an underlying geometric object—, and stability results—that is, robustness against initial measurement errors. We propose concrete algorithms in order to use these methods in practice.

The first family, the DTM-filtrations, is a robust alternative to the usual Čech filtration when the point cloud is noisy or contains anomalous points. It is based on the notion of distance to measure, which allows to obtain stability in the sense of the Wasserstein distance.

Secondly, we propose the lifted filtrations, which make it possible to estimate the homology of immersed manifolds, even when their reach is zero. We introduce the notion of normal reach, and show that it allows to control geometric quantities associated to the manifold. We study the estimation of tangent spaces by local covariance matrices.

Thirdly, we develop a theoretical framework for vector bundle filtrations, and define the persistent Stiefel-Whitney classes. We show that the persistent classes associated to the Čech bundle filtrations are Hausdorff-stable and consistent. To allow their algorithmic implementation, we introduce the notion of weak star condition.

Remerciements - Acknowledgements

Je souhaite que mes remerciements se dirigent principalement vers mes encadrants Frédéric Chazal et Marc Glisse ; la qualité de leur encadrement a contribué d'une part extrêmement significative dans l'intérêt et la joie que j'ai eu à mener ce travail de thèse.

J'adresse aussi mes remerciements à Shmuel Weinberger et Simon Masnou qui ont bien voulu rapporter ma thèse, ainsi qu'au reste du jury : Dominique Attali, Blanche Buet et Patrick Massot.

Moins personnellement mais tout aussi sincèrement, je remercie les membres des équipes pédagogiques et administratives de l'Inria et du laboratoire de mathématiques d'Orsay.

Enfin, et avec une chaleur que je ne saurais transmettre ici, je remercie ma famille, mes amis et mes collègues.

Contents

Résumé - Abstract **3**

Remerciements - Acknowledgements **5**

Contents **6**

Contents (detailed) **7**

Introduction **13**

I General introduction **13**

II Background **57**

Contributions **101**

III DTM-based filtrations **101**

IV Topological inference for immersed manifolds **131**

V Persistent Stiefel-Whitney classes **197**

Bibliography **237**

Contents (detailed)

Résumé - Abstract **3**

Remerciements - Acknowledgements **5**

Contents **6**

Contents (detailed) **7**

Introduction **13**

I General introduction **13**
I.1 Introduction générale en français 13
I.1.1 Contexte de cette thèse 13
I.1.2 Présentation du chapitre III : Filtrations-DTM 19
I.1.3 Présentation du chapitre IV : Inférence topologique pour les variétés immergées 25
I.1.4 Présentation du chapitre V : Classes de Stiefel-Whitney persistantes 30
I.2 General introduction in English 35
I.2.1 Context of this thesis 35
I.2.2 Presentation of Chapter III: DTM-based filtrations 41
I.2.3 Presentation of Chapter IV: Topological inference for immersed manifolds 46
I.2.4 Presentation of Chapter V: Persistent Stiefel-Whitney classes 51

II Background **57**
II.1 Notations 57
II.2 Background on differential geometry 60
II.2.1 Basic notions of differential geometry 60
II.2.2 Basic notions of simplicial topology 61
II.2.3 Basic notions of vector bundle theory 65
II.2.4 Basic notions of Stiefel-Whitney classes 68
II.2.5 Basic notions of Riemannian geometry 69
II.3 Background on Euclidean geometry of compact sets 71
II.3.1 Basic notions of topology 71
II.3.2 Thickenings and tubular neighborhoods 72
II.3.3 Reach 75

II.3.4 Weak feature size and μ-reach . . . 76
II.4 Background on persistent homology . . . 78
II.4.1 Basic notions of singular and simplicial homology . . . 78
II.4.2 Persistence modules . . . 81
II.4.3 Decomposition of persistence modules . . . 85
II.4.4 Stability of persistence modules . . . 87
II.4.5 Persistent cohomology theory . . . 88
II.5 Background on measure theory . . . 90
II.5.1 Basic notions of measure theory . . . 90
II.5.2 Distance-to-measure . . . 91
II.5.3 Varifolds . . . 92
II.6 Homology inference with Čech filtrations . . . 94
II.6.1 Consistency . . . 94
II.6.2 Stability . . . 96

Contributions 101

III DTM-based filtrations . . . 101
III.1 Weighted Čech filtrations . . . 102
III.1.1 Definition . . . 102
III.1.2 Stability . . . 104
III.1.3 Weighted Vietoris-Rips filtrations . . . 106
III.2 DTM-filtrations . . . 108
III.2.1 The distance to measure (DTM) . . . 109
III.2.2 DTM-filtrations . . . 109
III.2.3 Stability when $p = 1$ . . . 111
III.2.4 Stability when $p > 1$ . . . 116
III.2.5 Proof of Lemma III.20 and Proposition III.21 . . . 120
III.2.6 Consequences under the standard assumption . . . 123
III.3 Conclusion . . . 126
III.A Supplementary results for Section III.1 . . . 126
III.B Supplementary results for Section III.2 . . . 127

IV Topological inference for immersed manifolds . . . 131
IV.1 Preliminaries . . . 132
IV.1.1 Model and hypotheses . . . 132
IV.1.2 Index of constants . . . 134
IV.2 Reach of an immersed manifold . . . 135
IV.2.1 Geodesic bounds under curvature conditions . . . 135
IV.2.2 Normal reach . . . 139
IV.2.3 Probabilistic bounds under normal reach conditions . . . 143
IV.2.4 Quantification of the normal reach . . . 150
IV.3 Tangent space estimation . . . 156

IV.3.1 Local covariance matrices and lifted measure 156
IV.3.2 Consistency of the estimation 158
IV.3.3 Stability of the estimation 163
IV.3.4 An approximation theorem 169
IV.4 Topological inference with the lifted measure 171
IV.4.1 Overview of the method 171
IV.4.2 Homotopy type estimation with the DTM 175
IV.4.3 Persistent homology with DTM-filtrations 177
IV.5 Conclusion 180
IV.A Supplementary material for Section IV.2 180
IV.B Supplementary material for Section IV.3 184

V Persistent Stiefel-Whitney classes 197
V.1 Persistent Stiefel-Whitney classes 198
V.1.1 Definition 198
V.1.2 Čech bundle filtrations 199
V.1.3 Stability 203
V.1.4 Consistency 206
V.2 Computation of persistent Stiefel-Whitney classes 209
V.2.1 Simplicial approximation to Čech bundle filtrations 209
V.2.2 A sketch of algorithm 211
V.3 An algorithm when $d = 1$ 213
V.3.1 The star condition in practice 213
V.3.2 Triangulating the projective spaces 216
V.3.3 Vietoris-Rips version of the Čech bundle filtration 219
V.3.4 Choice of the parameter γ 219
V.4 Conclusion 222
V.A Supplementary material for Section V.1 223
V.B Supplementary material for Section V.3 224
V.B.1 Study of Example V.24 224
V.B.2 Study of Example V.25 226

Bibliography 237

Introduction

I General introduction

I.1 Introduction générale en français

I.1.1 Contexte de cette thèse

Cette thèse s'inscrit dans le contexte de l'inférence topologique à partir de nuages de points. Dans le cadre de l'analyse des données, le nuage de points représente le jeu de données à étudier. Il peut être par exemple le résultat d'une expérience scientifique, ou d'une acquisition de données quelconque. Nous allons étudier ce jeu de données selon les préceptes de l'Analyse Topologique des Données (TDA).

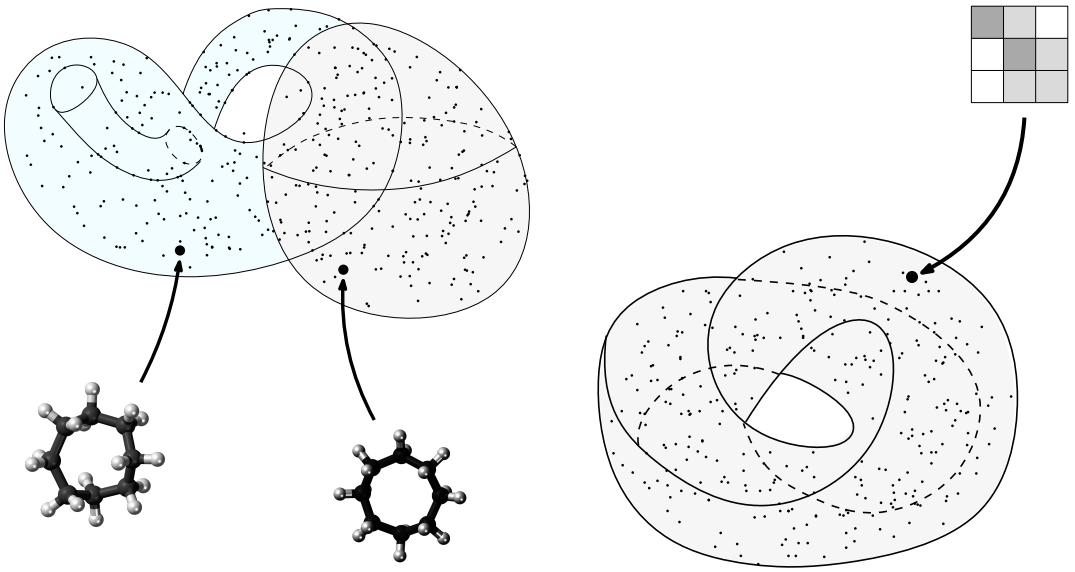
Analyse Topologique des Données. La principe fondamental de la TDA dit que le nuage de points à étudier, vu dans son ensemble, dessine une forme, dont la topologie est intéressante.

Ce principe a deux conséquences pour les applications. Déjà, en considérant la structure intrinsèque des données, on obtient souvent un objet de petite dimension. Cela permet d'envisager la conception d'algorithmes dont la complexité n'est plus limitée par la dimension de l'espace ambiant, souvent trop grande en pratique. En d'autres termes, on peut espérer échapper à ce que l'on appelle la malédiction de la dimension. Aussi, ce point de vue topologique a le mérite d'illuminer l'analyse des données depuis un angle différent des méthodes habituelles. Plutôt que d'appliquer des modèles rigides aux données, on préserve sa complexité inhérente, que l'on cherche à comprendre au travers d'invariants topologiques. Cela ouvre la porte à de nouvelles intuitions et découvertes.

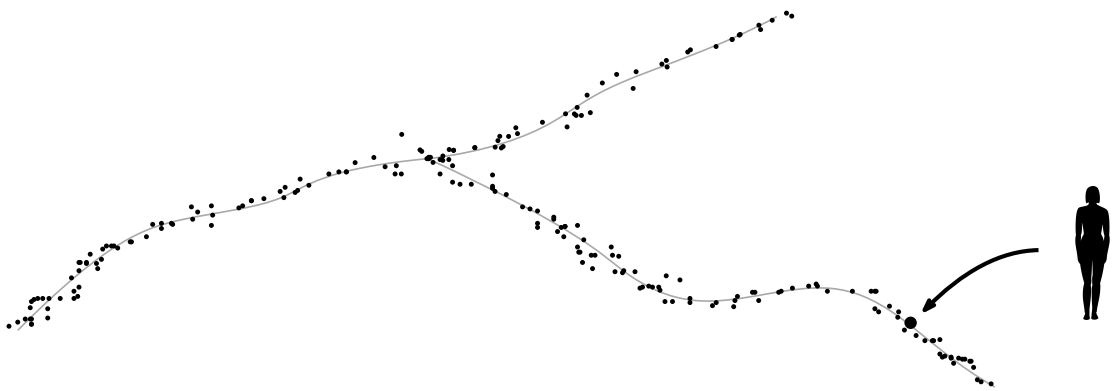
Afin de motiver la TDA, nous proposons trois exemples de jeux de données qui révèlent une topologie intéressante. Le premier exemple vient de la chimie [MTCW10]. La molécule de cyclo-octane C_8H_{16} admet plusieurs configurations stables, c'est-à-dire plusieurs arrangements spatiaux de ses atomes. La configuration d'une telle molécule peut être représentée par 72 variables – les coordonnées 3D de chacun de ses 24 atomes – ou, de manière équivalente, par un point dans $\mathbb{R}^{72}$. En analysant un grand nombre de ces molécules, les auteurs obtiennent un nuage de point dans $\mathbb{R}^{72}$. Dans cet espace de grande dimension, il apparaît que le nuage de points repose sur un objet de dimension bien plus petite : l'union d'une sphère et d'une bouteille de Klein, s'intersectant le long de deux cercles. Ces deux composantes correspondent à des arrangements spatiaux distincts des molécules – conformation *crown* dans la sphère, et conformation *boat-chair* dans la bouteille de Klein. Le comportement des molécules se situant sur l'intersection est encore une question ouverte.

Le deuxième exemple vient de l'analyse d'images [CIDSZ08]. À partir d'une large collection d'images, les auteurs extraient des sous-images (*patches*) de taille

3×3 pixels. Chacun de ces patch, puisqu'il contient 9 pixels, peut être vu comme un point en dimension 9. L'ensemble de ces patches peut ainsi être vu comme un nuage de points dans $\mathbb{R}^9$. Il apparaît que ce nuage se concentre autour d'un objet qui a l'homologie d'une bouteille de Klein. Dans un second temps, les auteurs montrent qu'une partie significative de ces points (60%) sont bien approchés par un plongement de la bouteille de Klein dans $\mathbb{R}^9$. Cette découverte a mené à la conception de méthodes d'analyse d'images basées sur la bouteille de Klein [PC14, CG20].



Le dernier exemple est emprunté à la biomédecine [NLC11]. À partir de tissus de patients infectés par le cancer du sein sont extraites 262 variables génétiques. L'ensemble de ces données forme un nuage de points dans $\mathbb{R}^{262}$. L'analyse proposée ici est différente des deux exemples précédents : il s'agit de réduire la dimension du jeu de données, sans trop changer sa topologie. Plus précisément, les auteurs considèrent sa structure de dimension 1, aussi appelée graphe de Reeb. En pratique, on le calcule grâce à l'algorithme MAPPER [SMC07]. Le résultat est un graphe, qui dans le cas présent s'avère être composé de trois branches. En exploitant cette structure, les auteurs ont découvert un sous-ensemble inattendu de patients : aucun d'entre eux ne décédaient du cancer du sein, ni ne présentaient de métastases. Ces patients correspondent à une signature moléculaire unique, qui fut reconnue comme un nouveau type de cancer du sein ($c\text{-}MYB^+$).



Depuis son émergence dans les années 2000, la TDA a été abondamment utilisée, sur une large gamme de jeux de données. En plus des aperçus précédents, nous pouvons citer des exemples en biomédecine [HGK12, WOC14, ARC14], en génétique [DAE+08, ER14], en physique [GHI+15, SHCP18], en cosmologie [Sou11, SPK11, PEVdW+17] et en analyse d'images et du signal [PD07, PZ16]. Dans chacune de ces études, l'idée est d'extraire de l'information topologique et géométrique du jeu de données, afin de construire de nouveaux descripteurs pour des tâches d'analyse et de classification, ou de découvrir de nouvelles propriétés qui pourraient mener à une meilleure compréhension du problème.

Le travail de cette thèse est motivé par l'enrichissement des fondations mathématiques de ces nouvelles applications. On souhaite obtenir des résultats mathématiques, ainsi que des algorithmes concrets, afin d'estimer des quantités topologiques pour lesquelles aucun algorithme n'a encore été conçu, ou dans des contextes qui n'ont pas encore été étudiés.

Inférence topologique et géométrique. Afin d'apprécier la palette de méthodes qui composent la TDA, nous pouvons jeter un coup d'œil aux recherches récentes. De nombreux travaux ont été portés par l'ambition commune d'estimer des quantités topologiques ou géométriques des jeux de données ; ces travaux forment un ensemble cohérent appelé *l'inférence topologique et géométrique*. La plupart d'entre eux partent du même contexte : les données se présentent sous la forme d'un nuage de points, c'est-à-dire un sous-ensemble fini d'un espace euclidien. Ce nuage de points est vu comme un échantillon d'un objet géométrique régulier, typiquement une sous-variété, ou plus généralement un sous-ensemble compact à portée strictement positive. Dans ce qui suit, le nuage de points est appelé X, et l'objet géométrique sous-jacent $\mathcal{M}$. Dans ce contexte, nous nous attachons à estimer des propriétés de $\mathcal{M}$ à partir de la simple observation de X. Ces propriétés se distinguent en deux classes : topologiques ou géométriques. Pour commencer, les *propriétés topologiques* sont définies comme des invariants de la classe d'homotopie de $\mathcal{M}$. Donnons une idée de la large gamme de propriétés topologiques étudiées jusqu'ici :

- le type d'homotopie de $\mathcal{M}$ [CCSL09, KSC+20],
- sa dimension [Cam03, KRW16, LMR17],
- ses triangulations [Fre02, BF04, BGO09, BG14, BDG18, BDG+, BW],
- son orientabilité [SW11],
- ses groupes d'homologie singulière [NSW08, CCSL09],
- ses opérations cohomologiques [GDR03, Aub11],
- ses classes caractéristiques [Jos04, Gai05, Aub11].

D'autre part, une *propriété géométrique* dépend de la façon dont $\mathcal{M}$ est plongé dans l'espace euclidien ambiant. On peut trouver, dans la recherche récente, l'estimation de :

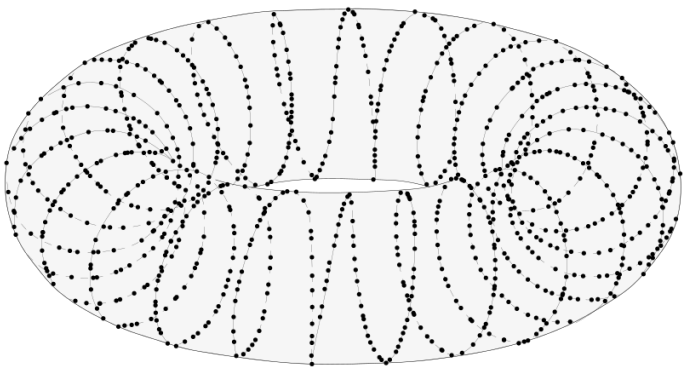
- ses espaces tangents [TVF13, AL19],
- sa distance géodésique [MLX+08, AB18, MSWW19, LD19],
- sa courbure et ses formes fondamentales [CCSL+17, AL19, BLM19b],

- sa portée [AKC+19].

Cette thèse se situe dans la continuité de ces travaux, en particuliers ceux de l'inférence topologique. Nous nous intéressons à l'estimation des types d'homotopie, des groupes d'homologie, des classes caractéristiques de Stiefel-Whitney et des espaces tangents.

Consistance et stabilité. Un problème d'estimation, tels que ceux cités précédemment, est habituellement résolu en proposant un algorithme qui, à partir de l'observation X, retourne un estimateur d'une propriété topologique ou géométrique de $\mathcal{M}$. Il est naturel de se demander d'évaluer la qualité de cet estimateur. Dans cette thèse, nous les évaluerons selon deux perspectives. La première est celle de la *consistance* : si nous observons directement $\mathcal{M}$ plutôt que X, à quel point notre estimateur est proche de la propriété à estimer ? Si elles coïncident, l'estimateur est dit consistant. Autrement, il est biaisé, et il s'agit alors de quantifier ce biais afin de certifier de sa pertinence. Ensuite, la seconde qualité d'approximation est appelée la *stabilité* : a quel point l'estimateur calculé à partir de X est proche de celui calculé à partir de $\mathcal{M}$? Un résultat de stabilité s'écrit typiquement sous la forme d'un borne sur la distance entre ces deux estimateurs, à partir d'un borne sur la distance entre X et $\mathcal{M}$. Cette stabilité est d'une importance critique en pratique : elle signifie que de petites erreurs de mesure dans X sont tolérées, et n'affecteront pas trop l'estimation finale.

Le problème de l'échelle. Lorsque l'on répond à un des problèmes d'estimation mentionnés précédemment, on fait face au problème suivant : les propriétés topologiques et géométriques des jeux de données dépendent fortement de l'échelle à laquelle ils sont considérés. Par exemple, le nuage de points suivant peut être vu comme un échantillon d'une courbe fermée, mais aussi comme un échantillon d'un tore.



Afin d'illustrer cette difficulté sur un problème d'estimation particulier, donnons nous comme tâche d'estimer les groupes d'homologie singulière de $\mathcal{M}$ à partir de l'observation X – un problème qui occupe une proportion significative de cette thèse. Nous remarquons que le nuage de points X, comme donné en pratique, est un espace topologique discret. Ainsi, il ne contient aucune topologique intéressante en lui-même. Toutefois, si il est échantillonné suffisamment proche de $\mathcal{M}$, il existe un construction qui permet de retrouver le type d'homotopie de $\mathcal{M}$ à partir de X, et par conséquent ses groupes d'homologie.

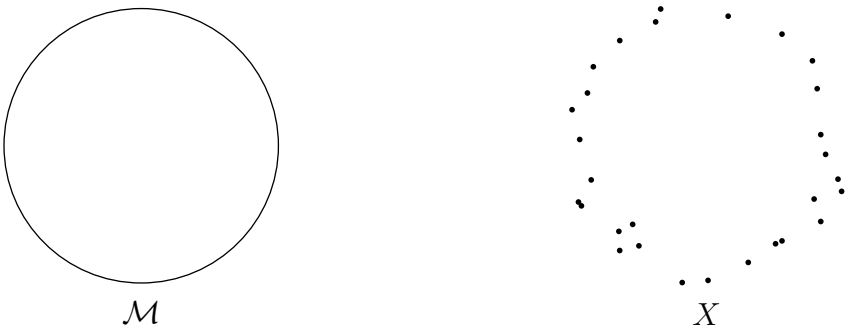
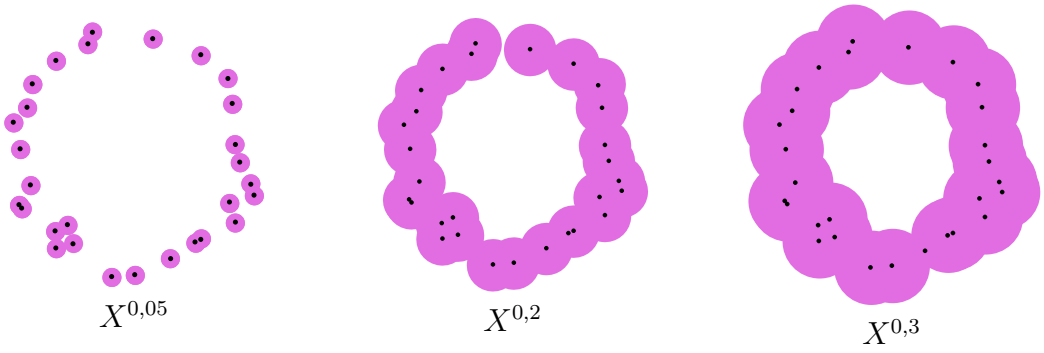


Figure I.1: L'objet sous-jacent et l'observation.

Cette construction consiste en l'épaississement de X. Pour tout $t \geq 0$, le t-épaississement de l'ensemble X, noté X^t, est l'ensemble des points de l'espace ambiant à distance au plus t de X.



On peut observer que la dernière figure est un épaississement qui, en tant qu'espace topologique, est du même type d'homotopie que $\mathcal{M}$. Afin d'estimer les propriétés topologiques associées à $\mathcal{M}$, il s'agit alors de sélectionner ces valeurs de t telles que X^t et $\mathcal{M}$ soient dans la même classe d'homotopie. En d'autres termes, on cherche à quelle échelle X doit être regardé. Si l'on a accès à certaines quantités géométriques, telles que la portée de $\mathcal{M}$ et la densité de X, alors il existe des procédures qui permettent de sélectionner de tels épaississement [NSW08, CCSL09]. Si l'on a pas accès à ces quantités, nous nous en remettons à l'homologie persistante.

Homologie persistante. Illustrons son utilisation, dans le contexte de l'estimation du $i^{\text{ème}}$ groupe d'homologie singulière $H_i(\mathcal{M})$. Dans le paragraphe précédent, nous avons finalement obtenu une collection d'épaississements, parmi lesquels certains ont le type d'homotopie de $\mathcal{M}$. Au lieu de sélectionner ces épaississements, l'idée de l'homologie persistante est de les considérer tous ensemble, et de récupérer ensuite de leur collection les groupes d'homologie de $\mathcal{M}$.

La famille de tous les épaississements est appelée la filtration de Čech et est notée $V[X] = (X^t)_{t\geq 0}$. C'est une famille croissante de sous-ensembles

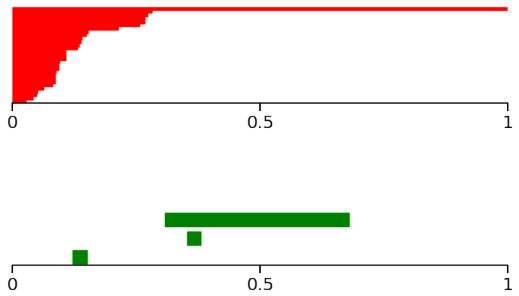
$$... \subset X^{t_1} \subset X^{t_2} \subset X^{t_3} \subset ...$$

En appliquant le $i^{\text{ème}}$ foncteur d'homologie singulière sur un corps, on obtient un diagramme d'espaces vectoriels

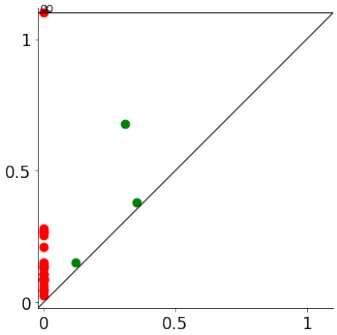
$$... \longrightarrow H_i(X^{t_1}) \longrightarrow H_i(X^{t_2}) \longrightarrow H_i(X^{t_3}) \longrightarrow ...$$

Leur collection forme un *module de persistance*, noté $\mathbb{V}[X] = (H_i(X^t))_{t\geq 0}$. Il s'agit de l'objet principal de la théorie de l'homologie persistante. Il rassemble l'homologie de X à chaque échelle.

Un travail considérable a été mené afin de comprendre la structure algébrique des modules de persistance. Dans certains cas, quand X est fini par exemple, le module de persistance $\mathbb{V}[X]$ peut être écrit comme une somme de modules de persistance indécomposables. Dans le cas fini, c'est une conséquence du théorème de Gabriel sur la décomposition des carquois ; ce résultat a ensuite été généralisé à une plus large gamme de modules de persistance [BCB20]. En général, la décomposition d'un module de persistance en modules indécomposables est unique, et donne un description complète du module de persistance sous une forme commode, appelée *code-barres.* Les codes-barres résument l'évolution des groupes d'homologie tout au long du processus d'épaississement. Ils peuvent être représentés de deux façons : comme des codes-barres ou comme des diagrammes.



Un code-barres de persistance



Un diagramme de persistance

Sur un code-barres, on lit des barres, qui correspondent aux cycles de $H_i(X^t)$ pour diverses valeurs de t. Plus une barre est longue, plus le cycle correspondant persiste à travers la filtration.

La figure précédente montre le code-barres de l'homologie de la filtration de Čech sur un échantillon X de $\mathcal{M}$, où $\mathcal{M}$ est le cercle unité de $\mathbb{R}^2$, comme sur la figure I.1. Le code-barres du module de persistance correspondant au $0^{\text{ème}}$ groupe d'homologie est représenté en rouge, et celui correspondant au 1^{er} groupe d'homologie en vert. On peut identifier deux barres proéminentes : une rouge et une verte. Elles correspondent à des cycles persistants, ce qui suggère que l'objet sous-jacent admet un $0^{\text{ème}}$ groupe d'homologie de dimension 1, ainsi qu'un 1^{er} groupe d'homologie de dimension 1. Ceci est cohérent avec l'homologie du cercle. Dit autrement, le code-barres construit à partir de X donne une idée de l'homologie de $\mathcal{M}$. plus formellement, dans cette thèse, nous construirons des codes-barres, et expliqueront comment y lire les groupes d'homologie.

Contributions. Ce manuscrit présente des résultats concernant l'inférence topologique à partir de nuages de points, basés sur la théorie de l'homologie persistante. Il consiste en trois travaux distincts bien que connectés. Nous en donnons une brève description ici, et une présentation plus précise dans les trois sous-sections suivantes. Premièrement, nous considérons des jeux de données X corrompus par des points aberrants. Dans ce cas, la filtration de Čech de X ne permet plus d'estimer

les groupes d'homologie de $\mathcal{M}$. À la place, nous proposons les filtrations-DTM, et montrons que l'homologie de $\mathcal{M}$ peut être lue dans leurs codes-barres. La stabilité de ces filtrations s'exprime en distance de Wasserstein – une distance entre mesure de probabilités qui autorise la présence de points aberrants dans le jeu de données. Notre deuxième travail consiste en l'étude des jeux de données qui échantillonnent une variété immergée. Quand l'objet sous-jacent $\mathcal{M}$ n'est pas une sous-variété, mais l'immersion d'une variété $\mathcal{M}_0$, la filtration de Čech ne permet pas d'estimer les groupes d'homologie de $\mathcal{M}_0$. Nous proposons une façon d'adapter cette méthode. Notre dernier travail prend place dans le contexte où l'objet sous-jacent $\mathcal{M}$ est muni d'une structure de fibré vectoriel. Dans ce cas, ses classes de Stiefel-Whitney sont définies, mais il n'existe pas d'algorithme général pour les calculer. Nous développons un cadre théorique persistant pour les classes de Stiefel-Whitney, et montrons qu'elles peuvent être estimées à partie d'un échantillon X du fibré vectoriel.

Puisque ces trois travaux sont envisagés comme des méthodes à appliquer en pratique, nous fournissons une feuille de programmation `Jupyter notebook` pour chacune d'entre elles. Elles contiennent des implémentations de nos méthodes, en langage `Python`. Elles sont disponibles aux liens suivants :

- DTM-based filtrations (Chapter III): `https://github.com/raphaeltinarrage/DTM-Filtrations/blob/master/Demo.ipynb`
- Topological inference for immersed manifolds (Chapter IV): `https://github.com/raphaeltinarrage/ImmersedManifolds/blob/master/Demo.ipynb`
- Persistent Stiefel-Whitney classes (Chapter V): `https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Demo.ipynb`

De plus, toutes ces fonctions sont rassemblées dans le package `velour`, disponible à l'adresse `https://github.com/raphaeltinarrage/velour`.

I.1.2 Présentation du chapitre III : Filtrations-DTM

Le problème des points aberrants. Retournons à notre cadre de travail initial : $\mathcal{M} \subset \mathbb{R}^n$ est un sous-ensemble dont l'homologie est à estimer, en se basant sur l'observation d'un autre sous-ensemble $X \subset \mathbb{R}^n$. La procédure de l'homologie persistante consiste à construire une filtration à partir de X – c'est-à-dire une famille croissante de sous-ensembles – puis de la convertir en un module de persistance en appliquant le foncteur d'homologie. Ce module de persistance est finalement résumé en un code-barres, depuis lequel on lit les propriétés topologiques de X.

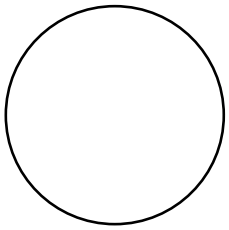
Parmi les filtrations disponibles à l'utilisateur, la plus commune est la filtration de Čech $V[X]$, définie comme la collection de tous les épaississement de X. Le module de persistance correspondant à son $i^{\text{ème}}$ homologie est noté $\mathbb{V}[X]$. En pratique, on utilise aussi la filtration de Vietoris-Rips, une version simpliciale de la filtration de Čech, moins coûteuse à calculer. L'avantage de ces filtrations est de produire des diagrammes de persistance qui sont robustes aux petites variations de X en distance de Hausdorff. Ce résultat s'appelle le théorème de stabilité

[CSEH07, CDSO14] : si $\mathbb{V}[X]$ et $\mathbb{V}[\mathcal{M}]$ sont les modules de persistance associés aux filtrations de Čech de X et $\mathcal{M}$, alors

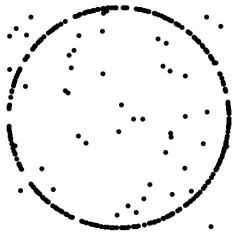
$$\mathrm{d_i}\left(\mathbb{V}[X], \mathbb{V}[\mathcal{M}]\right) \leq \mathrm{d_H}\left(X, \mathcal{M}\right),$$

où $\mathrm{d_i}\left(\cdot,\cdot\right)$ est la distance d'entrelacement entre les modules de persistance, et $\mathrm{d_H}\left(\cdot,\cdot\right)$ la distance de Hausdorff entre les sous-ensembles de $\mathbb{R}^n$. À côté de ce résultat de stabilité, nous avons un résultat de consistance : le module de persistance $\mathbb{V}[\mathcal{M}]$ contient de l'information concernant les groupes d'homologie de $\mathcal{M}$, pourvu que $\mathcal{M}$ soit une sous-variété ou un sous-ensemble à portée strictement positive [CCSL09]. Par conséquent, si $\mathrm{d_H}\left(X, \mathcal{M}\right)$ est petite, alors la procédure habituelle de l'homologie persistante permet de retrouver l'homologie de $\mathcal{M}$ à partir de l'observation X.

Malheureusement, si le nuage de points X contient des points aberrants, alors la distance de Hausdorff $\mathrm{d_H}\left(X, \mathcal{M}\right)$ peut être grande, et le théorème de stabilité ne donnerait pas une borne intéressante. L'homologie de $\mathcal{M}$ ne peut alors plus être lue dans la filtration de Čech $V[X]$. En d'autres termes, ces filtrations usuelles sont sensibles à la présence de points aberrants dans le jeu de données, ce qui les rend inappropriées dans ce contexte. Il est important de voir que la distance de Hausdorff $\mathrm{d_H}\left(X, \mathcal{M}\right)$ n'est plus une mesure pertinente de la proximité entre X et $\mathcal{M}$.



$\mathcal{M}$



X

Afin de résoudre ce problème, nous proposons de réduire l'importance des points aberrants dans la filtration. Nous voudrions obtenir une filtration qui se comporterait comme si les points aberrants étaient absents. Pour parvenir à une telle construction, nous allons quantifier le degré d'aberrance des points, via une estimation de densité locale. Cette idée demande de voir le nuage de points X comme une mesure.

Un point de vue mesure. Voyons maintenant les sous-ensembles X et $\mathcal{M}$ comme des mesures de probabilités, notées μ et ν. Par exemple, si X est fini, μ pourrait être la mesure empirique sur X, et si $\mathcal{M}$ est une sous-variété compacte, ν pourrait être la mesure uniforme sur $\mathcal{M}$. La *distance de Wasserstein* $\mathrm{W}_2\left(\mu,\nu\right)$ entre ces mesures permet de quantifier la proximité entre X et $\mathcal{M}$, tout en autorisant quelques points aberrants. Est-il possible de construire des filtrations $W[\mu]$ et $W[\nu]$ telles que l'on ait une stabilité de la forme

$$\mathrm{d_i}\left(W[\mu], W[\nu]\right) \leq \text{constante} \cdot \mathrm{W}_2\left(\mu,\nu\right) \ ? \tag{I.1}$$

Différents travaux ont été menés ces dernières années dans cette direction. Parmi eux, les filtrations par les sous-niveaux de la distance-à-la-mesure (DTM), introduites dans [CCSM11], et certaines de ses variantes [PWZ15], permettent de capturer des informations topologiques de l'espace sous-jacent.

Malheureusement, le calcul des sous-niveaux de la DTM fait intervenir le calcul d'un diagramme de Voronoï d'ordre k, qui est souvent trop coûteux en pratique. Pour contourner ce problème, [GMM13] introduit la k-distance témoignée, dont la persistance est plus légère à calculer. Les auteurs montrent que la k-distance témoignée est une bonne approximation de la DTM, à une constante additive près. On trouve aussi dans [BL19] une étude d'une autre approximation de la DTM, appelée la k-PDTM. Dans [BCOS16, Buc14], une version pondérée de la filtration de Vietoris-Rips est proposée afin d'approcher la persistance de la DTM, et plusieurs résultats de stabilités et d'approximation sont établis, dans la même veine que [GMM13].

Nous proposons une solution alternative à ce problème, dans l'intention de combiner les deux aspects des méthodes présentées précédemment : stabilité de la DTM, et légèreté algorithmique de ses variantes. À cette fin, nous définissions les *filtrations-DTM*. Ces filtrations se situent à mi-chemin entre les filtrations usuelles et la filtration par les sous-niveaux de la DTM : elles sont calculables en pratique, et elles sont robustes aux points aberrants.

DTM-filtrations. Afin de construire les filtrations-DTM, nous commençons par généraliser la notion de *filtration de Čech pondérée*. Nous rappelons au lecteur que la filtration de Čech de X à l'instant t est définie comme le sous-ensemble suivant de $\mathbb{R}^n$:

$$V^t[X] = \bigcup_{x \in X} \overline{\mathcal{B}}\,(x, t)\,,$$

où $\overline{\mathcal{B}}\,(x,t)$ est la boule fermée de centre x et de rayon t. Afin de réduire l'importance des point aberrants x de X, nous proposons de retarder l'apparition de la boule $\overline{\mathcal{B}}\,(x,t)$ dans la filtration. Ceci est accompli par l'utilisation de boules modifiées $\overline{\mathcal{B}}_f(x,t)$. Elles sont définies comme

$$\overline{\mathcal{B}}_f(x,t) = \begin{cases} \emptyset & \text{si } t < f(x), \\ \overline{\mathcal{B}}\left(x, \left(t^p - f(x)^p\right)^{\frac{1}{p}}\right) & \text{sinon.} \end{cases}$$

Cette boule dépend de deux paramètres : un réel $f(x) \geq 0$ qui contrôle le délai d'apparition de la boule dans la filtration, et un paramètre $p \in [0, \infty]$ qui contrôle sa vitesse de grossissement. Une fois que l'on a choisi une fonction $f\colon X \to \mathbb{R}^+$, appelée la *fonction de pondération*, et un paramètre p, on peut définir la filtration de Čech pondérée. Elle est notée $V[X, f, p]$, et est définie comme la collection des sous-ensembles

$$V^t[X, f, p] = \bigcup_{x \in X} \overline{\mathcal{B}}_f(x, t).$$

Quand p vaut 2, cette construction apparaît déjà dans [Buc14, BCOS16]. En utilisant des résultats classiques, nous montrons que ces filtrations sont stables aux petites perturbations de X en distance de Hausdorff, et aux petites perturbations de f en norme sup (propositions III.4 et III.5).

Maintenant, pour une fonction quelconque f, ces résultats de stabilité ne sont pas satisfaisants face jeux de données avec points aberrants. Il s'agit de choisir une

fonction de pondération f qui attribue des valeurs plus grandes aux points aberrants x, de sorte à ce que leurs boules $\overline{\mathcal{B}}_f(x,t)$ apparaissent tard dans la filtration. Une fonction qui remplit ce critère est la DTM. La DTM dépend d'une mesure μ et d'un paramètre $m \in (0,1)$, et est notée $\mathrm{d}_{\mu,m}\colon \mathbb{R}^n \to \mathbb{R}^+$. Étant donné une mesure μ et deux paramètres $m \in (0,1)$ et $p \in [0,+\infty]$, on définit la *filtration-DTM* $W[\mu,m,p]$ comme la collection des sous-ensembles

$$W^t[\mu,m,p] = \bigcup_{x\in\mathrm{supp}(\mu)} \overline{\mathcal{B}}_{\mathrm{d}_{\mu,m}}(x,t),$$

où $\mathrm{supp}(\mu)$ représente le support de la mesure μ. En d'autres termes, la filtration-DTM est la filtration de Čech pondérée, avec pour paramètres l'ensemble $X = \mathrm{supp}(\mu)$, la fonction de pondération $f = \mathrm{d}_{\mu,m}$ et le paramètre p. Le module de persistance correspondant est noté $\mathbb{W}[\mu,m,p]$. Si μ est la mesure empirique sur un ensemble fini X, on écrit simplement $W[X,m,p]$ et $\mathbb{W}[X,m,p]$.

En guise d'illustration, nous considérons sur la figure I.2 un ensemble X qui est un échantillon du cercle unité avec des points aberrants. En utilisant la filtration-DTM, on voit que les boules grossissent plus vite sur le cercle sous-jacent que sur les points aberrants.

Figure I.2: Gauche : La filtration de Čech usuelle $V^t[X]$ au temps $t = 0{,}3$. **Droite :** La filtration-DTM $W^t[X,m,p]$ avec paramètres $m = 0{,}1$ et $p = 1$.

Comme contrepartie de la simplicité de leur construction, les filtrations-DTM n'héritent pas d'une stabilité comme décrite par l'équation (I.1). On obtient toutefois une autre formulation de leur stabilité : la proximité entre deux filtrations-DTM $W[\mu,m,p]$ et $W[\nu,m,p]$ repose sur l'existence de mesures intermédiaires qui sont proches de μ et ν pour la métrique de Wasserstein.

Théorème III.22. Soient μ et ν deux mesures de probabilité sur $\mathbb{R}^n$, $m \in (0,1)$ et $p \in [1,+\infty]$. Pour toutes mesures de probabilité μ', ν' telles que $\mathrm{supp}(\mu') \subset \mathrm{supp}(\mu)$ et $\mathrm{supp}(\nu') \subset \mathrm{supp}(\nu)$, nous avons

$$\mathrm{d_i}\left(\mathbb{W}[\mu,m,p], \mathbb{W}[\nu,m,p]\right) \leq m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu,\mu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu',\nu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\nu',\nu\right) + c(\mu',m,p) + c(\nu',m,p).$$

La notation $\mathrm{d_i}\left(\cdot,\cdot\right)$ fait référence ici à la distance d'entrelacement, et les termes $c(\mu',m,p)$, $c(\nu',m,p)$ à des quantités qui témoignent de la régularité des mesures μ' et ν' (définies aux équations (III.2) et (III.3)). Le théorème donne une borne petite lorsque les mesures μ' et ν' sont choisies comme des versions de μ et ν nettoyées

de leurs points aberrants. Lorsque $p = 1$ et que la mesure initiale ν satisfait à la condition de régularité (a, d)-standard (introduite à la sous-section III.2.6), on obtient une stabilité plus explicite :

Corollaire III.24. Si la mesure ν est (a, d)-standard, alors

$$\mathrm{d_i}\left(W[\mu, m, 1], W[\nu, m, 1]\right) \leq \text{constante} \cdot \left(\left(\frac{\mathrm{W}_2\left(\mu, \nu\right)}{m} \right)^{\frac{1}{2}} + m^{\frac{1}{d}} \right).$$

Ces résultats montrent que la filtration-DTM construite sur μ est un estimateur fiable de celle construite sur ν, à condition que la distance de Wasserstein $\mathrm{W}_2\left(\mu, \nu\right)$ et que le paramètre m soient petits.

Enfin, on montre un résultat de consistance : les filtrations-DTM sont proches des filtrations par les sous-niveaux de la DTM.

Corollaire III.25. Soit V la filtrations par les sous-niveaux de la DTM $\mathrm{d}_{\nu,m}$. Si la mesure ν est (a, d)-standard, alors

$$d_i(V,\ W[\nu, m, 1]) \leq \text{constante} \cdot m^{\frac{1}{d}}.$$

En combinant ces résultats de stabilité et de consistance, et en choisissant un paramètre m de l'ordre de $\mathrm{W}_2\left(\mu, \nu\right)^{\frac{d+1}{2}}$, on obtient

$$d_i(V,\ W[\mu, m, 1]) \leq \text{constante} \cdot \mathrm{W}_2\left(\mu, \nu\right)^{\frac{d+1}{2d}}.$$

Ainsi, les filtrations-DTM peuvent être utilisés comme des estimateurs robustes de l'homologie de l'objet sous-jacent aux observations. Ceci est illustré par la figure I.3, qui représente les diagrammes de persistance des filtrations de la figure I.2 (une filtration de Čech et une filtration-DTM). Seulement le second diagramme révèle l'homologie d'un cercle. En effet, parmi les deux points éloignés de la diagonale, le point rouge indique un $0^{\text{ème}}$ groupe d'homologie de dimension 1, et le point vert un 1^{er} groupe d'homologie de dimension 1 également.

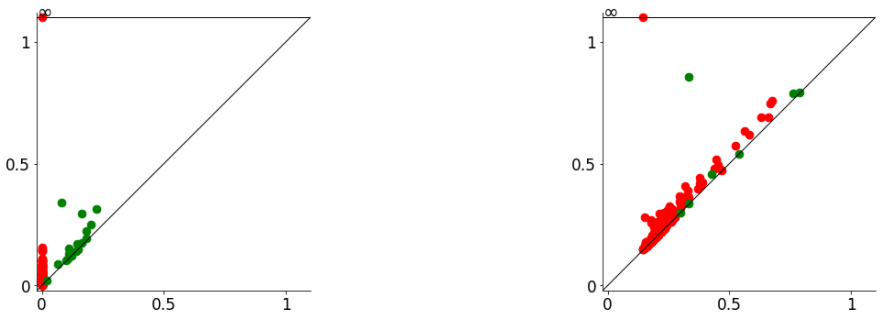


Figure I.3: Gauche : Diagramme de persistance de la filtration de Čech $V[X]$. **Droite :** Diagramme de persistance de la filtration-DTM $W[X, m, 1]$.

Motivations appliquées. Les filtrations-DTM on été initialement expérimentées dans le cadre d'un projet de recherche industriel dont le but était la détection d'anomalies à partir de capteurs inertiels posés sur des ponts et des bâtiments [Lab18]. Dans cette étude, les données se présentaient sous la forme de séries

temporelles, mesurant l'accélération de capteurs attachés aux ponts ou bâtiments à étudier. En utilisant des fenêtres glissantes et un *time-delay embedding*, ces séries temporelles ont été converties en nuages de points de cardinal fixé dans $\mathbb{R}^n$. Des filtrations ont ensuite été construites à partir de ces nuages, et leur persistance a été calculée, donnant naissance à des diagrammes de persistance dépendant du temps. Ils ont ensuite été analysés de façon à détecter des anomalies ou des caractéristiques temporelles particulières [SDB16, Ume17]. Dans ce cadre appliqué, les filtrations-DTM se sont avérés plus robustes au bruit, mais aussi capables de mieux révéler les propriétés topologiques du jeu de données que la filtration de Vietoris-Rips habituelle.

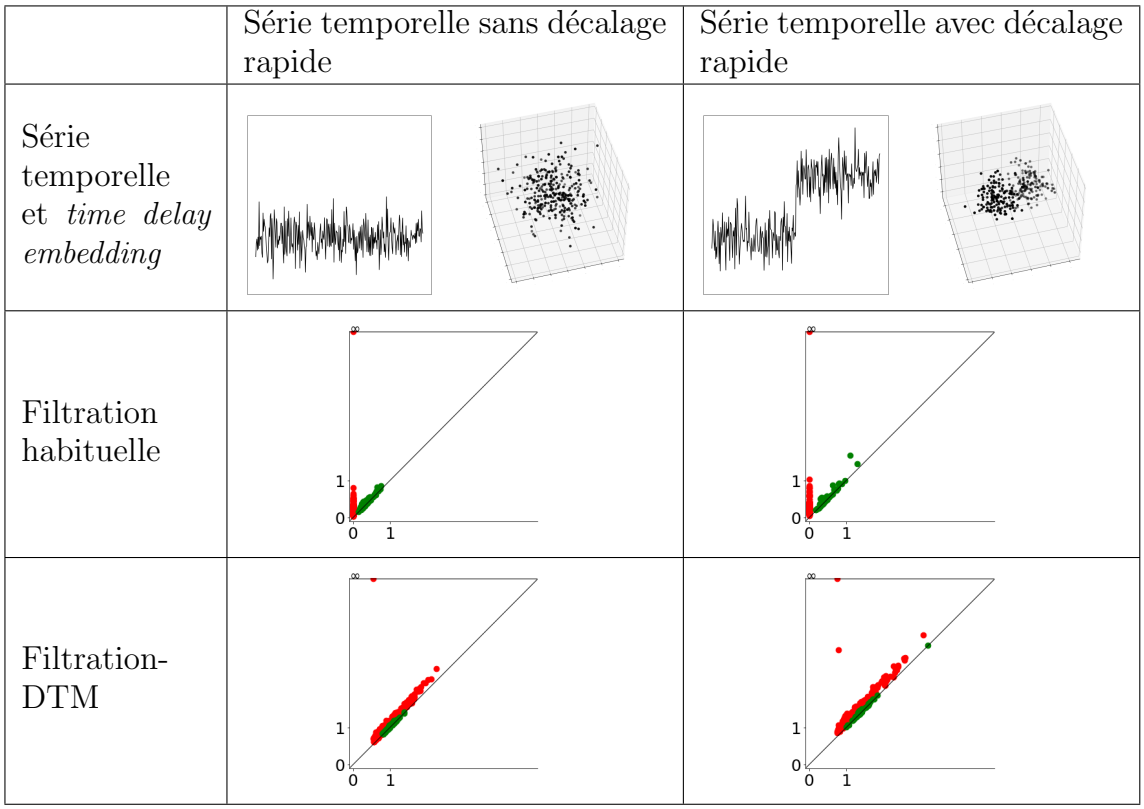


Figure I.4: Comparaison entre la filtration de Vietoris-Rips et la filtration-DTM.

La figure I.4 est un exemple synthétique qui compare la filtration de Vietoris-Rips et la filtration-DTM. La première ligne représente deux séries temporelles qui présentent des comportements très différents, ainsi que leur plongement dans $\mathbb{R}^3$. Ici, une série $(x_1, x_2, \ldots, x_n)$ est transformée en le nuage de points

$$\{(x_1, x_2, x_3), (x_2, x_3, x_4), \ldots, (x_{n-2}, x_{n-1}, x_n)\}.$$

La deuxième ligne contient les diagrammes de persistance des filtrations de Vietoris-Rips construites sur ces deux nuages de points (les points rouges et verts correspondant respectivement aux cycles de degré 0 et de degré 1). On voit que ces diagrammes ne permettent pas de détecter clairement les comportements différents de ces deux séries temporelles. Sur la troisième ligne sont représentés les diagrammes de persistance de filtrations-DTM construites sur ces nuages de points. Un points rouge sur

le second diagramme apparaît clairement éloigné de la diagonale, ce qui est une conséquence du décalage rapide de la seconde série temporelle.

I.1.3 Présentation du chapitre IV : Inférence topologique pour les variétés immergées

Inférence topologique pour les variétés immergées. On sait que si l'objet $\mathcal{M}$ à estimer est une sous-variété, et que si le nuage de points observé X est suffisamment proche en distance de Hausdorff, alors l'homologie persistante de X permet d'estimer les groupes d'homologie de $\mathcal{M}$. Dans ce contexte, on utilise la filtration de Čech de X. Toutefois, en pratique, il peut être trop limitant de demander à ce que l'objet sous-jacent soit une sous-variété. Certains travaux on été menés afin de réduire la régularité de $\mathcal{M}$. Par exemple, si nous supposons seulement que $\mathcal{M}$ soit un sous-ensemble à portée strictement positive, ou même à μ-portée strictement positive, alors il est prouvé que la filtration de Čech de X permet encore de retrouver l'homologie de $\mathcal{M}$ [CCSL09].

Dans ce travail, nous étudions un cas de régularité différent : $\mathcal{M}$ est une variété immergée, non plongée. Formellement, nous considérons une variété abstraite $\mathcal{M}_0$, immergée dans un espace euclidien par une application $u\colon \mathcal{M}_0 \to \mathbb{R}^n$, dont l'image est $\mathcal{M}$. Comme précédemment, nous observons un sous-ensemble $X \subset \mathbb{R}^n$, que nous supposons proche de $\mathcal{M}$. Notre objectif est d'estimer l'homologie de $\mathcal{M}_0$ à partir de X.

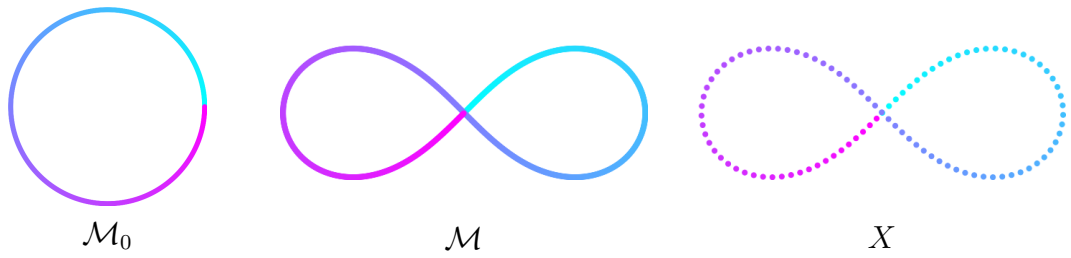


Figure I.5: Gauche : La variété abstraite $\mathcal{M}_0$, un cercle. **Centre :** L'immersion $\mathcal{M} \subset \mathbb{R}^2$, connue sous le nom de lemniscate de Bernoulli. **Droite :** L'observation X.

En tant qu'immersion, $\mathcal{M}$ peut s'auto-intersecter, et les ensembles $\mathcal{M}_0$ et $\mathcal{M}$ peuvent ne pas partager le même type d'homotopie. La filtration de Čech de $\mathcal{M}$, ou de X, révélerait l'homologie de $\mathcal{M}$, et non celle de $\mathcal{M}_0$. Par conséquent, l'approche usuelle basée sur la filtration de Čech ne s'applique plus ici, et de nouvelles méthodes doivent être développées.

Parmi les travaux qui concernent les variétés immergés, nous citons [ACLZ17], qui se place dans le contexte où $\mathcal{M}$ est une union de sous-variétés qui s'intersectent. Ainsi, $\mathcal{M}$ n'est pas une sous-variété, mais une variété immergée. Les auteurs proposent un algorithme pour classifier les différentes composantes de $\mathcal{M}$, c'est-à-dire les composantes connexes de $\mathcal{M}_0$. Leur algorithme est basé sur l'estimation d'espaces tangents, afin de séparer l'ensemble $\mathcal{M}$ là où il s'intersecte ; c'est un point de vue que nous adoptons aussi.

Relever les variétés immergées. Afin d'estimer l'homologie d'une variété à partir de l'observation d'une de ses immersions, nous proposons d'estimer son fibré tangent, vu comme sous-ensemble d'un autre espace euclidien. Comme nous le verrons, pendant la procédure d'estimation de ce fibré tangent, nous commettrons des erreurs, qui résulteront en des points aberrants. Ce problème sera résolu en utilisant les filtrations-DTM, décrites précédemment. Nous considérons ainsi dès à présent un point de vue mesure. Nous présentons maintenant notre méthode.

Soient $\mathcal{M}_0$ une variété compacte de régularité $\mathcal{C}^2$ et de dimension d, et μ_0 une mesure de probabilité de Radon sur $\mathcal{M}_0$ de support $\mathcal{M}_0$. Soit $u\colon \mathcal{M}_0 \to \mathbb{R}^n$ une immersion. Nous faisons l'hypothèse que cette immersion est telle que les points aux auto-intersections correspondent à des espaces tangents différents. Autrement dit, pour tout $x_0, y_0 \in \mathcal{M}_0$ tels que $x_0 \neq y_0$ et $u(x_0) = u(y_0)$, les espaces tangents $d_{x_0}u(T_{x_0}\mathcal{M}_0)$ et $d_{y_0}u(T_{y_0}\mathcal{M}_0)$ de $\mathcal{M}_0$, vus dans $\mathbb{R}^n$, sont différents. Définissons l'image de l'immersion $\mathcal{M} = u(\mathcal{M}_0)$ et la mesure poussée en avant $\mu = u_*\mu_0$. On suppose que l'on observe la mesure μ, ou une mesure ν proche. Notre objectif est d'estimer l'homologie singulière de $\mathcal{M}_0$ (avec coefficients dans $\mathbb{Z}_2$ par exemple) à partir de ν.

Afin de revenir à $\mathcal{M}_0$, nous procédons comme suit : soit $\mathrm{M}(\mathbb{R}^n)$ l'espace vectoriel des matrices $n \times n$, et $\check{u}\colon \mathcal{M}_0 \to \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ l'application

$$\check{u}\colon x_0 \longmapsto \left(u(x_0),\ \frac{1}{d+2} p_{T_x\mathcal{M}}\right),$$

où $p_{T_x\mathcal{M}}$ est la matrice de projection orthogonale sur l'espace tangent $T_x\mathcal{M} \subset \mathbb{R}^n$, écrite dans la base canonique de $\mathbb{R}^n$. L'ensemble $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$ est une sous-variété de $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, difféomorphique à $\mathcal{M}_0$. Il est appelé le *relevé* de $\mathcal{M}_0$. L'espace $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ est appelé *l'espace de relèvement*.

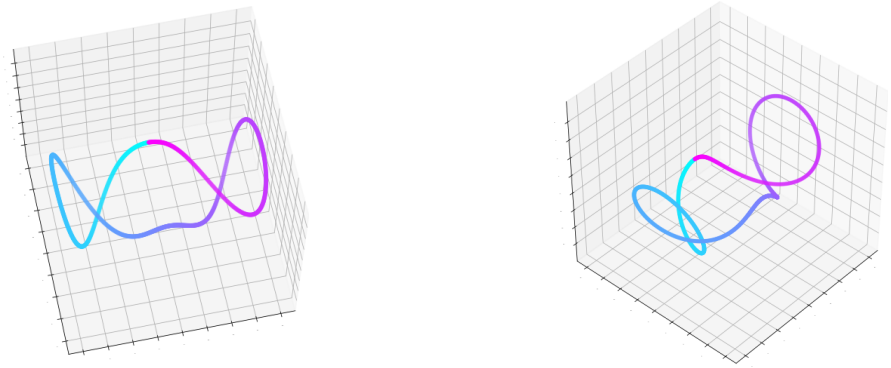


Figure I.6: Deux visualisations de la sous-variété $\check{\mathcal{M}} \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, projetée dans un sous-espace de dimension 3 par PCA. On voit qu'elle ne s'auto-intersecte pas. L'ensemble initial $\mathcal{M}$ est représenté en figure I.5.

Supposons que l'on sache estimer $\check{\mathcal{M}}$ à partir de ν. Alors on pourrait considérer l'homologie persistance d'une filtration basée sur $\check{\mathcal{M}}$ – disons la filtration de Čech de $\check{\mathcal{M}}$ dans l'espace de relèvement $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ par exemple – et espérer lire l'homologie de $\mathcal{M}_0$ dans le code-barres correspondant.

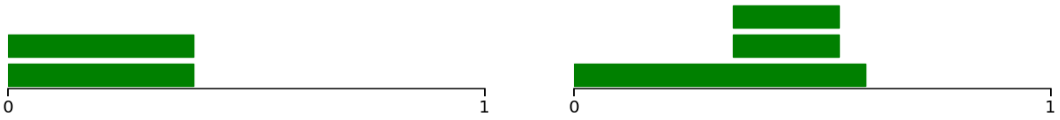


Figure I.7: Gauche : Code-barres de la 1-homologie de la filtration de Čech de $\mathcal{M}$ dans l'espace ambiant $\mathbb{R}^2$. On y lit l'homologie d'un lemniscate. **Droite :** Code-barres de la 1-homologie de la filtration de Čech de $\check{\mathcal{M}}$ dans l'espace de relèvement $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$. On y lit l'homologie d'un cercle (au début de la filtration).

Malheureusement, nous n'allons pas être capables de donner une bonne estimation de $\check{\mathcal{M}}$. En effet, les espaces tangents $p_{T_x\mathcal{M}}$, que l'on estime par des matrices de covariance locale, ne seront pas estimés correctement si x est trop proche d'une auto-intersection de $\mathcal{M}$. Ainsi, plutôt que d'estimer l'ensemble relevé $\check{\mathcal{M}}$, nous proposons d'estimer la *mesure relevée exacte* $\check{\mu}_0$, définie comme $\check{\mu}_0 = \check{u}_*\mu_0$. C'est une mesure sur l'espace de relèvement $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, de support $\check{\mathcal{M}}$. En utilisant des filtrations pour mesures – comme les filtrations-DTM – on peut espérer retrouver l'homologie de $\mathcal{M}_0$.

Il est à noter que $\check{\mathcal{M}}$ peut être naturellement vu comme une sous-variété de $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^n)$, où $\mathcal{G}_d(\mathbb{R}^n)$ dénote la grassmannienne des sous-espaces linéaires de dimension d de $\mathbb{R}^n$. De ce point de vue, $\check{\mu}_0$ peut être vue comme une mesure sur $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^n)$, autrement dit, un *varifold*. Toutefois, pour des raisons informatiques, nous préférons travailler dans l'espace matriciel $\mathrm{M}(\mathbb{R}^n)$ plutôt que $\mathcal{G}_d(\mathbb{R}^n)$.

Voici une définition alternative de $\check{\mu}_0$: pour toute fonction test $\phi\colon \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n) \to \mathbb{R}$,

$$\int \phi(x, A)\mathrm{d}\check{\mu}_0(x, A) = \int_{\mathcal{M}_0} \phi\left(u(x_0), \frac{1}{d+2} p_{T_x\mathcal{M}}\right) \mathrm{d}\mu_0(x_0).$$

De retour à la mesure observée ν, nous proposons d'estimer $\check{\mu}_0$ avec la *mesure relevée* $\check{\nu}$, définie comme suit : pour toute fonction test $\phi\colon \mathbb{R}^n \times \mathrm{M}(E) \to \mathbb{R}$,

$$\int \phi(x, A)\mathrm{d}\check{\nu}(x, A) = \int_{\mathcal{M}} \phi\Big(x, \overline{\Sigma}_\nu(x)\Big)\mathrm{d}\nu(x),$$

où $\overline{\Sigma}_\nu(x)$ est la matrice de covariance locale renormalisée (définie à la section IV.3). Elle dépend d'un paramètre $r > 0$. Nous montrons que $\overline{\Sigma}_\nu(x)$ peut être utilisée pour estimer les espaces tangents $\frac{1}{d+2} p_{T_x\mathcal{M}}$ de $\mathcal{M}$. Toutefois, cette estimation est biaisée autour des auto-intersections de $\mathcal{M}$, comme montré en figure I.8. Afin de quantifier la qualité de cette approximation, nous définissons une nouvelle quantité géométrique.

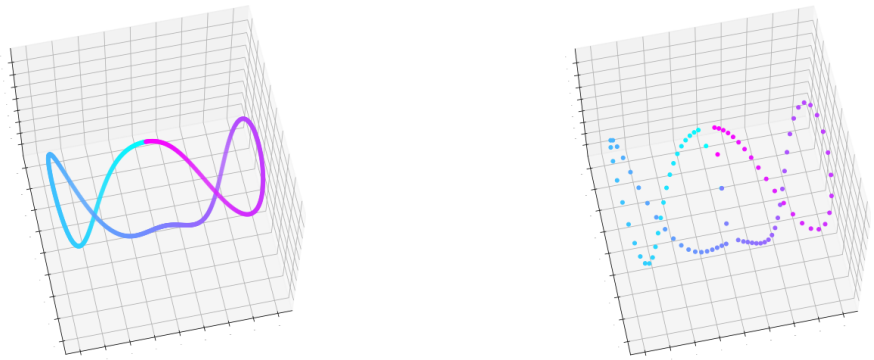


Figure I.8: Gauche : Les ensembles $\mathrm{supp}(\mu) = \mathcal{M}$ et $\mathrm{supp}(\check{\mu}_0) = \check{\mathcal{M}}$, où μ est la mesure uniforme sur $\mathcal{M}$ (voir figure I.5). **Droite :** Les ensembles $\mathrm{supp}(\nu)$ et $\mathrm{supp}(\check{\nu})$, où ν est la mesure empirique sur X. Paramètres $\gamma = 2$ and $r = 0{,}1$.

Portée normale. Puisque la notion usuelle de portée n'est plus pertinente dans le cas des variétés immergées, nous introduisons la *portée normale* – une fonction qui indique localement à quel point la variété immergée peut être vue comme une sous-variété. Elle est notée $\lambda \colon \mathcal{M} \to [0, +\infty)$. Nos résultats s'expriment aussi en fonction de ρ, une mesure de la courbure de $\mathcal{M}_0$.

Théorème IV.11. Soit $\mathcal{M}$ une variété immergée de classe $\mathcal{C}^2$. Soit $x \in \mathcal{M}$ et $r < \min\left\{\frac{1}{4\rho}, \lambda(x)\right\}$. Alors $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ est un ensemble de portée au moins $\frac{1-2\rho r}{\rho}$.

Quand $\mathcal{M}$ est une variété plongée, nous connectons la portée normale avec la notion habituelle de portée.

Proposition IV.10. Soit $\mathcal{M}$ une variété plongée de classe $\mathcal{C}^2$. On a

$$\mathrm{reach}(\mathcal{M}) = \min\left\{\frac{1}{\rho_*}, \frac{1}{2}\lambda_*\right\},$$

où ρ_* est le maximum des normes d'opérateur des secondes formes fondamentales de $\mathcal{M}_0$, et $\lambda_* = \inf_{x \in \mathcal{M}} \lambda(x)$ est le minimum de sa portée normale.

La portée normale permet de quantifier la qualité d'approximation de la mesure relevée exacte $\check{\mu}_0$ par la mesure relevée $\check{\nu}$. Nous prouvons un résultat d'estimation globale, de la forme suivante : $\check{\mu}_0$ et $\check{\nu}$ sont proches en distance de Wasserstein, tant que μ et ν le sont. Nous utilisons une version modifiée de la distance de Wasserstein dans l'espace de relèvement, notée $\mathrm{W}_{p,\gamma}(\cdot, \cdot)$, qui dépend d'un paramètre $\gamma > 0$ (défini à la sous-section IV.3.1). Ce paramètre est conçu pour équilibrer l'importance donnée à l'information spatiale (composante $\mathbb{R}^n$) et l'information des espaces tangents (composante $\mathrm{M}(\mathbb{R}^n)$) dans $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$. De plus, nos résultats reposent sur les hypothèses techniques 1, 2, 3 et 4, que nous décrivons à la sous-section IV.1.1.

Théorème IV.33. Soit $\mathcal{M}$ une variété immergée qui vérifie les hypothèses

1, 2 et 3. Soient $\gamma, r > 0$. Si $\mathrm{W}_p(\mu,\nu)$ et r sont assez petits, alors

$$\mathrm{W}_{p,\gamma}(\check{\nu},\check{\mu}_0) \leq \text{constante} \cdot \gamma \cdot \left(\mu(\lambda^r)^{\frac{1}{p}} + r + \left(\frac{\mathrm{W}_p(\mu,\nu)}{r^{d+1}} \right)^{\frac{1}{2}} \right) + 2\mathrm{W}_p(\mu,\nu).$$

La quantité $\mu(\lambda^r)$ réfère à $\mu(\lambda^{-1}([0,r]))$, c'est-à-dire la mesure des points de $\mathcal{M}$ de petite portée normale. En ajoutant une hypothèse sur cette quantité, on obtient un résultat plus simple.

Corollaire IV.35. De plus, si l'immersion $\mathcal{M}$ satisfait l'hypothèse 4, alors

$$\mathrm{W}_{p,\gamma}(\check{\nu},\check{\mu}_0) \leq \big(1 + \text{constante} \cdot \gamma\big) r^{\frac{1}{p}}.$$

Comme conséquence de ce résultat, nous pouvons estimer le type d'homotopie de $\mathcal{M}_0$ (corollaire IV.38). En utilisant des filtrations-DTM, nous pouvons aussi retrouver l'homologie de $\mathcal{M}_0$ (corollaire IV.42).

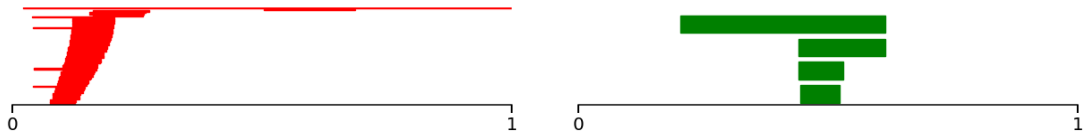


Figure I.9: Code-barres de la 0-homologie **(gauche)** et 1-homologie **(droite)** de la filtration-DTM sur la mesure relevée $\check{\nu}$. On peut y lire l'homologie d'un cercle (une grande barre rouge et une grande barre verte). Paramètres $\gamma = 2$, $r = 0{,}1$ et $m = 0{,}01$.

Estimation des espaces tangents. Notre méthode repose sur l'estimation des espaces tangents de $\mathcal{M}$ à partir de X, via les matrices de covariance locale. Leur stabilité est exprimée en distance de Wasserstein. Nous les étudions par restriction des mesure μ et ν aux boules $\overline{\mathcal{B}}(x,r)$. Les mesures de probabilités correspondantes sont notées $\overline{\mu_x}$ et $\overline{\nu_x}$. Nous montrons que les mesures de probabilité restreintes héritent de la distance de Wasserstein initiale (équation (IV.25)). La stabilité locale des mesures a aussi été étudiée dans [MMM18, MSW19] et l'estimation des espaces tangents par les matrices de covariance locale dans [ACLZ17]. Nos résultats améliorent ceux de [MSW19].

Comme conséquence de la stabilité, et en utilisant la portée normale, nous prouvons que $\overline{\Sigma}_\mu(x)$ est un estimateur consistant des espaces tangents $\frac{1}{d+2} p_{T_x\mathcal{M}}$ de $\mathcal{M}$, et qu'elle est robuste à des petites variations de μ.

Proposition IV.24 et équation (IV.26). Soient $\mathcal{M}$ une variété immergée et μ une mesure qui satisfait les hypothèses 2 et 3. Soient $x \in \mathcal{M}$ et $r < \min\left\{\frac{1}{2\rho}, \lambda(x)\right\}$. On a

$$\left\| \overline{\Sigma}_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq \text{constante} \cdot r.$$

De plus, si ν est une autre mesure de probabilité, alors pour $\mathrm{W}_1(\mu,\nu)$ assez

petit, on a

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq \text{constante} \cdot \left(\frac{\mathrm{W}_1(\mu,\nu)}{r^{d+1}}\right)^{\frac{1}{2}}.$$

En choisissant un rayon r de l'ordre de $\mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}$, on obtient

$$\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} \leq \text{constante} \cdot \mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}.$$

En d'autres termes, les matrices de covariance locale de ν estiment les espaces tangents de $\mathcal{M}$ à vitesse $\mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}$. Cette borne n'est pas aussi bonne que l'état de l'art de méthodes d'estimation d'espaces tangents [AL19]. Toutefois, notre résultat est très général, puisque nous supposons seulement que l'observation ν est proche de μ en distance de Wasserstein.

I.1.4 Présentation du chapitre V : Classes de Stiefel-Whitney persistantes

D'autres invariants topologiques. Comme précédemment, considérons $\mathcal{M}$ et X des sous-ensembles de $\mathbb{R}^n$, où X est vu comme un échantillon de $\mathcal{M}$. Jusqu'à présent, nous nous sommes attachés à estimer les groupes d'homologie de $\mathcal{M}$ à partir de X. Toutefois, il existe en topologie algébrique de nombreux autres invariants associés à $\mathcal{M}$, qu'il est intéressant d'estimer. La structure d'anneau de cohomologie de $\mathcal{M}$, ou les opérations cohomologiques dans $H^*(\mathcal{M})$, en font partie. Leur estimation a été étudiée dans le contexte de la théorie de la persistance. Dans [Yar10], l'auteur propose un algorithme pour calculer le cup-product, et dans [Aub11] un algorithme pour calculer la carrés de Steenrod.

Les classes de Stiefel-Whitney sont un autre exemple d'invariant. Elles sont associées à tout espace topologique muni d'une structure de fibré vectoriel. Nous considérons le problème d'inférence suivant : $\mathcal{M}$ est muni d'un fibré vectoriel, et nous souhaitons estimer ses classes de Stiefel-Whitney à partir d'une observation X. Notre approche se place dans un cadre persistant : nous définissons les *classes de Stiefel-Whitney persistantes*, et donnons des résultats de stabilité et de consistance.

À notre connaissance, le problème d'estimation des classes de Stiefel-Whitney n'a pas été considéré dans le contexte d'une observation sous forme de nuage de points. Toutefois, dans [Aub11], l'auteure propose un algorithme pour calculer les classes de Stiefel-Whitney, dans le cas particulier du fibré tangent d'un espace Euler mod-2 (c'est-à-dire un complexe simplicial qui triangule une variété).

Les classes de Stiefel-Whitney comme raffinement de la cohomologie. En pratique, les classes de Stiefel-Whitney fournissent plus d'information topologique que les groupes de cohomologie seuls. Afin d'illustrer cela, considérons le tore et la bouteille de Klein, notés $\mathcal{M}_0$ et $\mathcal{M}_0'$. Seule l'une de ces deux surfaces est orientable, et elles ne sont donc pas homéomorphes. Soit $\mathbb{Z}_2$ le corps à deux éléments. Observons

que les groupes de cohomologie de $\mathcal{M}_0$ et $\mathcal{M}'_0$ sur $\mathbb{Z}_2$ sont égaux :

$$\begin{aligned} H^0(\mathcal{M}_0) &= H^0(\mathcal{M}'_0) = \mathbb{Z}_2, \\ H^1(\mathcal{M}_0) &= H^1(\mathcal{M}'_0) = \mathbb{Z}_2 \times \mathbb{Z}_2, \\ H^2(\mathcal{M}_0) &= H^2(\mathcal{M}'_0) = \mathbb{Z}_2. \end{aligned}$$

Ainsi, les groupes de cohomologie sur $\mathbb{Z}_2$ ne permettent pas de distinguer les variétés $\mathcal{M}_0$ et $\mathcal{M}'_0$. Afin de les différencier, divers raffinements issus de la topologie algébrique peuvent être utilisés. Ici, nous étudions les classes de Stiefel-Whitney. Si nous équipons $\mathcal{M}_0$ et $\mathcal{M}'_0$ de leurs fibrés tangents, leurs premières classes de Stiefel-Whitney sont distinctes : seule l'une d'elles est nulle. Nous sommes ainsi capables de distinguer ces deux variétés.

$H^*(\mathcal{M}_0) = \mathbb{Z}_2[x,y]/\langle x^2, y^2\rangle$
$w_1(\tau_{\mathcal{M}_0}) = 0$

$H^*(\mathcal{M}'_0) = \mathbb{Z}_2[x,y]/\langle x^3, y^2, x^2y\rangle$
$w_1(\tau_{\mathcal{M}'_0}) = x$

Figure I.10: Les anneaux de cohomologie de $\mathcal{M}_0$ et $\mathcal{M}'_0$ sur $\mathbb{Z}_2$, et les premières classes de Stiefel-Whitney de leurs fibrés tangents respectifs $\tau_{\mathcal{M}_0}$ et $\tau_{\mathcal{M}'_0}$.

Plus généralement, les classes de Stiefel-Whitney ont été abondamment utilisés en topologie différentielle, par exemple dans des problèmes d'immersion de variétés dans des espaces de petite dimension, ou dans des problèmes de cobordisme. Notre travail a pour intention de proposer cet outil à la théorie de l'homologie persistante.

Classes de Stiefel-Whitney persistantes. En toute généralité, si X est un espace topologique muni d'une structure de fibré vectoriel ξ de dimension d, il existe une collection de classes de cohomologie $w_1(\xi), ..., w_d(\xi)$, les classes de Stiefel-Whitney, telles que $w_i(\xi)$ soit un élément du groupe de cohomologie $H^i(X)$ sur $\mathbb{Z}_2$ pour $i \in [1, d]$. Afin de définir les classes de Stiefel-Whitney dans un cadre persistant, nous allons utiliser une définition commode des fibrés vectoriels : la donnée d'un fibré vectoriel sur un espace compact X est équivalente à la donnée d'une application continue $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ pour m assez grand, où $\mathcal{G}_d(\mathbb{R}^m)$ est la grassmannienne des sous-espaces linéaires de dimension d de $\mathbb{R}^m$. Une telle application est appelée une *application classifiante* pour ξ. Elle est fortement apparentée à l'application de Gauss des sous-variétés orientables de $\mathbb{R}^3$, comme expliqué à la figure I.11.

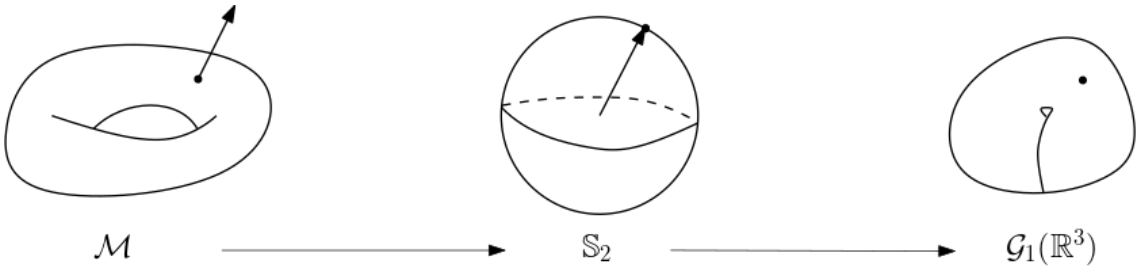


Figure I.11: Si $\mathcal{M}$ est une surface orientable de $\mathbb{R}^3$, l'application de Gauss $g\colon \mathcal{M} \to \mathbb{S}_2$ associe chaque $x \in \mathcal{M}$ a un vecteur normal de $\mathcal{M}$ en x. En post-composant cette application avec l'application quotient habituelle $\mathbb{S}_2 \to \mathcal{G}_1(\mathbb{R}^3)$, on obtient une application classifiante $p\colon \mathcal{M} \to \mathcal{G}_1(\mathbb{R}^3)$ pour le fibré normal de $\mathcal{M}$.

Étant donné une application classifiante $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ pour un fibré vectoriel ξ, les classes de Stiefel-Whitney $w_1(\xi), ..., w_d(\xi)$ peuvent être définies comme les images de certaines classes de cohomologie de la grassmannienne via l'application induite en cohomologie $p^*\colon H^*(X) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^m))$. Si w_i est la $i^{\text{ème}}$ classe de Stiefel-Whitney de la grassmannienne, alors la $i^{\text{ème}}$ classe de Stiefel-Whitney du fibré vectoriel ξ est

$$w_i(\xi) = p^*(w_i). \tag{I.2}$$

Afin de déplacer ces considérations dans un cadre persistant, supposons que l'on ait un jeu de données de la forme (X, p), où X est un sous-ensemble fini de $\mathbb{R}^n$, et p est une application $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$. Notons $(X^t)_{t\geq 0}$ la filtration de Čech de X, c'est à dire la collection des t-épaississements X^t de X dans l'espace ambiant $\mathbb{R}^n$. Pour définir des classes de Stiefel-Whitney *persistantes*, on voudrait étendre l'application $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ en $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$. Cependant, nous n'avons pas trouvé de façon intéressante d'étendre cette application. À la place, nous proposons de regarder le jeu de données d'un point de vue différent. Transformons le fibré vectoriel (X, p) en le sous-ensemble de $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$ défini par

$$\check{X} = \{(x, p(x)), x \in X\}.$$

La grassmannienne $\mathcal{G}_d(\mathbb{R}^m)$ se plonge naturellement dans $\mathrm{M}(\mathbb{R}^m)$, l'espace des matrices $m \times m$. Dans cette perspective, $\check{X}$ peut être vu comme un sous-ensemble de $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Soit $(\check{X}^t)_{t\geq 0}$ la filtration de Čech de $\check{X}$ dans l'espace ambiant $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Une application naturelle $p^t\colon \check{X}^t \to \mathcal{G}_d(\mathbb{R}^m)$ peut être définie : associe chaque point $(x, A) \in \check{X}^t$ à la projection de A sur $\mathcal{G}_d(\mathbb{R}^m)$, vue comme sous-ensemble de $\mathrm{M}(\mathbb{R}^m)$:

$$p^t\colon (x, A) \in \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m) \longmapsto \mathrm{proj}\,(A, \mathcal{G}_d(\mathbb{R}^m)).$$

La projection est bien définie si A n'appartient pas à l'axe médian de $\mathcal{G}_d(\mathbb{R}^m)$. Nous montrons que cette condition est vérifiable en pratique (lemme V.3). La filtration de Čech de $\check{X}$, munie des applications de projection étendues $(p^t\colon \check{X}^t \to \mathcal{G}_d(\mathbb{R}^m))_t$, est appelée la *filtration de fibré vectoriel de Čech*. Maintenant, nous

pouvons définir la $i^{\text{ème}}$ classe de Stiefel-Whitney persistante comme la collection de classes de cohomologie $w_i(X) = (w_i^t(X))_t$, où $w_i^t(X)$ est l'image

$$w_i^t(X) = (p^t)^*(w_i),$$

et où w_i est la $i^{\text{ème}}$ classe de Stiefel-Whitney de la grassmannienne (nous invitons le lecteur à comparer cette définition avec l'équation (I.2)). Nous résumons l'information donnée par une classe de Stiefel-Whitney persistante dans un diagramme, que nous appelons une *barre de vie*.

Stabilité et consistance. Afin d'illustrer nos résultats, considérons le plongement du tore $u\colon \mathcal{M}_0 \to \mathcal{M} \subset \mathbb{R}^3$ représenté sur la figure I.12. Dénotons par $p_{T_x\mathcal{M}}$ la matrice de projection sur l'espace tangent de $\mathcal{M}$ en x. L'ensemble $\check{\mathcal{M}} = \{(x, p_{T_x\mathcal{M}}), x \in \mathcal{M}\}$ est un sous-ensemble de $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$.

Figure I.12: La sous-variété $\mathcal{M} \subset \mathbb{R}^3$, et la sous-variété $\check{\mathcal{M}} \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3) \simeq \mathbb{R}^{12}$, projetée dans un sous-espace de dimension 3 par PCA.

La barre de vie de la première classe de Stiefel-Whitney persistante de ce fibré est représentée en figure I.13. La barre est hachurée, ce qui signifie que la classe est nulle tout au long de la filtration. Ceci est cohérent avec la première classe de Stiefel-Whitney habituelle du fibré tangent du tore, qui est nulle.

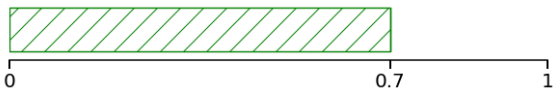


Figure I.13: Barre de vie de la première classe de Stiefel-Whitney persistante de $\check{\mathcal{M}}$. Elle est définie sur l'intervalle $\left[0, \frac{\sqrt{2}}{2}\right)$ seulement (voir la définition V.4).

Pour continuer, considérons l'immersion de la bouteille de Klein $u'\colon \mathcal{M}'_0 \to \mathcal{M}' \subset \mathbb{R}^3$ représentée sur la figure I.14. Pour $x_0 \in \mathcal{M}'_0$, notons $p_{T_{x_0}\mathcal{M}}$ la matrice de projection sur l'espace tangent de $\mathcal{M}'_0$ en x_0, vu dans $\mathbb{R}^3$. L'ensemble $\check{\mathcal{M}}' = \left\{\left(u(x_0), p_{T_{x_0}\mathcal{M}}\right), x_0 \in \mathcal{M}'_0\right\}$ est un sous-ensemble de $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$. Nous remarquons que $\check{\mathcal{M}}'$ est une sous-variété (difféomorphique à la bouteille de Klein), alors que $\mathcal{M}'$ ne l'est pas.

Figure I.14: L'ensemble $\mathcal{M}' \subset \mathbb{R}^3$, et la sous-variété $\check{\mathcal{M}}' \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3) \simeq \mathbb{R}^{12}$, projetée dans un sous-espace de dimension 3 par PCA.

Comme précédemment, nous calculons les classes de Stiefel-Whitney persistantes sur la filtration de Čech de $\check{\mathcal{M}}'$. La figure I.15 représente la barre de vie de la première classe de Stiefel-Whitney persistante de cette filtration. La barre est remplie, ce qui signifie que la classe est non-nulle tout au long de la filtration. Ceci est cohérent avec la première classe de Stiefel-Whitney habituelle du fibré tangent de la bouteille de Klein.

Figure I.15: Barre de vie de la première classe de Stiefel-Whitney persistante $\check{\mathcal{M}}'$.

La construction que nous proposons est définie pour tout sous-ensemble $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. En particulier, elle peut être appliquée à tout échantillon fini de $\check{\mathcal{M}}$ et $\check{\mathcal{M}}'$.

> **Corollaire V.8.** Soient deux sous-ensembles $X, Y \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ tels que $\mathrm{d_H}(X, Y) \leq \epsilon$. Alors il existe un ϵ-entrelacement entre les modules de cohomologie persistante de leurs filtrations de fibrés vectoriels de Čech, tel que les classes de Stiefel-Whitney persistantes soient envoyés sur les classes de Stiefel-Whitney persistantes.

On montre aussi que les classes de Stiefel-Whitney persistantes sont des estimateurs consistants des classes de Stiefel-Whitney.

> **Corollaire V.11.** Soient $\mathcal{M}_0 \to \mathcal{M} \subset \mathbb{R}^n$ une immersion, $p\colon \mathcal{M}_0 \to \mathcal{G}_d(\mathbb{R}^m)$ un fibré vectoriel, et $\check{\mathcal{M}} \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ l'ensemble relevé correspondant. Soit $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ un sous-ensemble tel que $\mathrm{d_H}(X, \mathcal{M}) \leq \epsilon$. Alors pour tout $t \in [4\epsilon, \mathrm{reach}(\check{\mathcal{M}}) - 3\epsilon)$, la composition des inclusions $\mathcal{M}_0 \hookrightarrow \check{\mathcal{M}} \hookrightarrow X^t$ induit un isomorphisme $H^*(\mathcal{M}_0) \leftarrow H^*(X^t)$ qui envoie la $i^{\text{ème}}$ classe de Stiefel-Whitney persistante $w_i^t(X)$ de la filtration de fibré vectoriel de Čech de X sur la $i^{\text{ème}}$ classe de Stiefel-Whitney de $(\mathcal{M}_0, p)$.

En guise d'illustration, nous représentons sur la figure I.16 les barres de vie des premières classes de Stiefel-Whitney persistantes des échantillons X et X' de $\check{\mathcal{M}}$ et $\check{\mathcal{M}}'$. Nous observons qu'elles sont proches des barres de vie initiales.

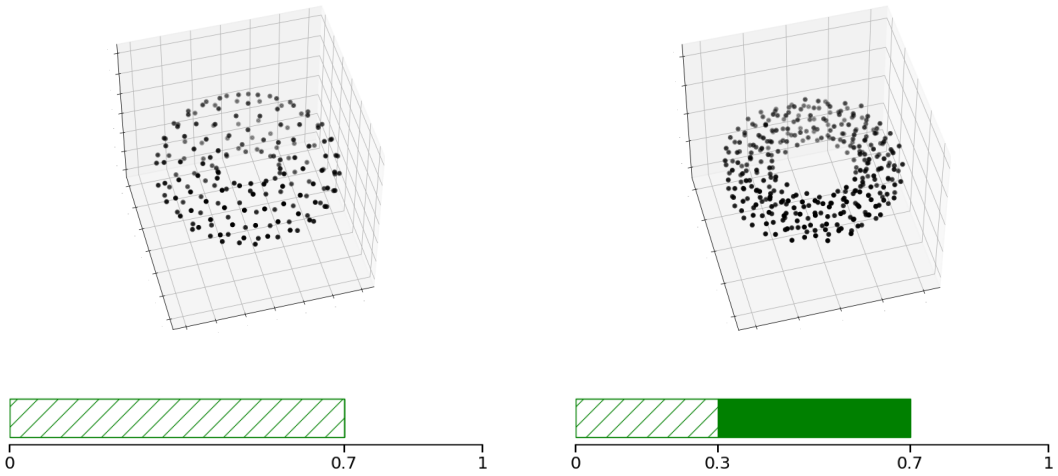


Figure I.16: Gauche : Un échantillon de $\check{\mathcal{M}} \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$, projeté dans $\mathbb{R}^3$, et la barre de vie de sa première classe de Stiefel-Whitney persistante. **Droite :** De même pour $\check{\mathcal{M}}'$.

Mise en œuvre algorithmique. Nous proposons un algorithme concret afin de calculer les classes de Stiefel-Whitney persistantes. Cet algorithme repose sur plusieurs ingrédients, dont la triangulation des espaces projectifs, et la méthode de l'approximation simpliciale.

L'approximation simpliciale, bien que très utilisée théoriquement, ne peut s'appliquer que si le complexe simplicial est suffisamment fin, une propriété qui est certifiée par la *condition étoile*. Cependant, cette condition ne peut pas être vérifiée en pratique. Nous contournons ce problème en introduisant la *condition étoile faible*, une variante qui ne dépend que de la structure combinatoire du complexe simplicial. Lorsque le complexe simplicial est suffisamment fin, la condition étoile et la condition étoile faible deviennent des notions équivalentes (proposition V.20).

I.2 General introduction in English

I.2.1 Context of this thesis

This thesis takes place in the context of topological inference from point clouds. From a data analysis point of view, the point cloud represents the input dataset. For instance, it may be the result of a scientific experiment, or any acquisition of data. We will study this dataset according to the precepts of Topological Data Analysis (TDA).

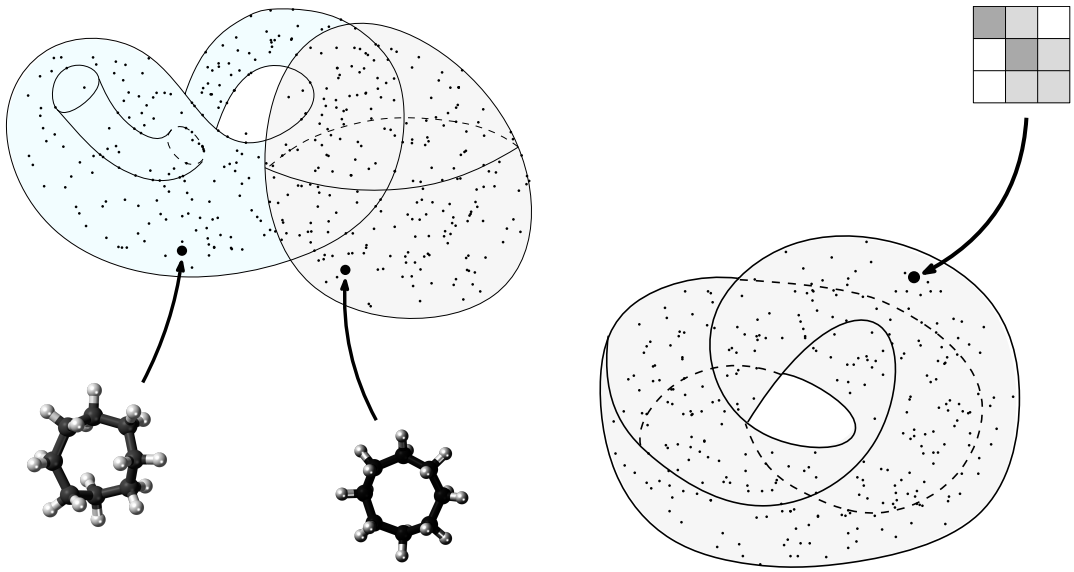
Topological Data Analysis. A fundamental principle of TDA is that the input point cloud, seen as a whole, forms a shape, whose topology is worth understanding.

From a practical perspective, the consequences of this principle are twofold. First, by looking at the intrinsic structure of the data, we often end up with an object of small dimension. This makes it possible to design algorithms whose complexities are no longer limited by the dimension of the ambient space, which may be too large in practice. In other words, one can hope to escape what is known as the curse of dimensionality. Second, this topological viewpoint illuminates data analysis from

a different angle than the usual methods. Rather than applying rigid models to the dataset, we preserve its inherent complexity, and seek to understand it through topological invariants. This opens the door to new insights and discoveries.

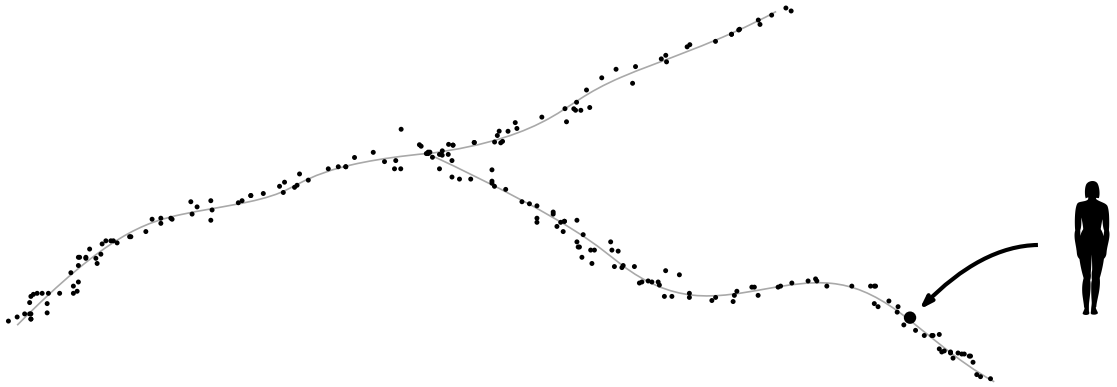
As motivating examples, let us cite a couple datasets where interesting topology appears. The first one comes from chemistry [MTCW10]. The cyclo-octane molecule C_8H_{16} admits several stable configurations, i.e., several spatial arrangements of its atoms. The configuration of such a molecule can be represented by 72 variables—the 3D coordinates of each of its 24 atoms—, or equivalently, by a point in $\mathbb{R}^{72}$. By analyzing many of these molecules, the authors obtain a point cloud in $\mathbb{R}^{72}$. In this large dimensional space, it turns out that the point cloud lies on an object of much smaller dimension, namely, the union of a sphere and a Klein bottle, intersecting in two rings. These two components correspond to distinct spatial arrangements of the molecule: crown conformation in the sphere, and boat-chair conformation in the Klein bottle. The behavior of molecules lying in the intersection is still an open question.

A second example comes from image processing [CIDSZ08]. From a large collection of natural images, the authors extract 3×3 patches. Since it consists of 9 pixels, each of these patches can be seen as a 9-dimensional vector, and the whole set as a point cloud in $\mathbb{R}^9$. It appears that this dataset concentrates near an object that has the homology of a Klein bottle. In a second step, the authors show that a significant part of the points (60%) are well approximated by an embedding of the Klein bottle in $\mathbb{R}^9$. This discovery has led to the construction of Klein-bottle-based image analysis methods [PC14, CG20].



We give a last example, taken from biomedicine [NLC11]. Tissues of patients infected with breast cancer has been analyzed, resulting in 262 genomic variables per patients. Gathering these data yields a point cloud in $\mathbb{R}^{262}$. In a different context from the two previous examples, the analysis here consists in reducing the dimension of the dataset, while not changing its topology too much. More precisely, one is looking for its 1-dimensional structure, known as the Reeb graph. This is performed in practice with the so-called MAPPER algorithm [SMC07]. The result is a graph, which turned out to be composed of three distinct branches. Taking advantage

of this structure, the authors discovered an unexpected subset of patients: they exhibit a 100% survival, and no metastasis. They correspond to a unique molecular signature, that yields to the designation of a new type of breast cancer (*c-MYB*$^+$).



Since its emergence in the 2000s, TDA has been used extensively, on a wide variety of datasets. In addition to the previous examples, we may cite: biomedicine [HGK12, WOC14, ARC14], genetics [DAE$^+$08, ER14], physics [GHI$^+$15, SHCP18], cosmology [Sou11, SPK11, PEVdW$^+$17] and image and signal processing [PD07, PZ16]. In each of these works, the idea is to extract topological and geometric information from the dataset, in order to design new features for analysis and classification tasks, or to discover insightful properties that would lead to a better understanding of the problem.

The work of this thesis is motivated by expanding the mathematical foundations of these new applications. We aim at obtaining mathematical results and concrete algorithms to estimate topological quantities, for which general algorithms have not been proposed, or in contexts that have not yet been studied.

Topological and geometric inference. To appreciate the large range of methods that make up TDA, let us look at where recent research has led us. A lot of work has been carried out with the common ambition to estimate topological or geometric properties of datasets; they form a coherent ensemble that we call *topological and geometric inference*. Most of these works fall within the same framework: the input data is a point cloud, that is, a finite subset of a Euclidean space. This point cloud is seen as a sample of a regular geometric object, typically a compact submanifold, or more generally a compact set with positive reach. In what follows, the point cloud is called X, and the underlying geometric object $\mathcal{M}$. In this context, we aim to estimate properties of $\mathcal{M}$ from the mere observation of X. These properties are distinguished in two classes: topological and geometric. On the one hand, a *topological property* is defined as an invariant of the homotopy class of $\mathcal{M}$. Let us give an idea of the variety of topological properties studied so far:

- the homotopy type of $\mathcal{M}$ [CCSL09, KSC$^+$20],
- its dimension [Cam03, KRW16, LMR17],
- a triangulation of it [Fre02, BF04, BGO09, BG14, BDG18, BDG$^+$, BW],
- its orientability [SW11],
- its singular homology groups [NSW08, CCSL09],

- its cohomology operations [GDR03, Aub11],
- its characteristic classes [Jos04, Gai05, Aub11].

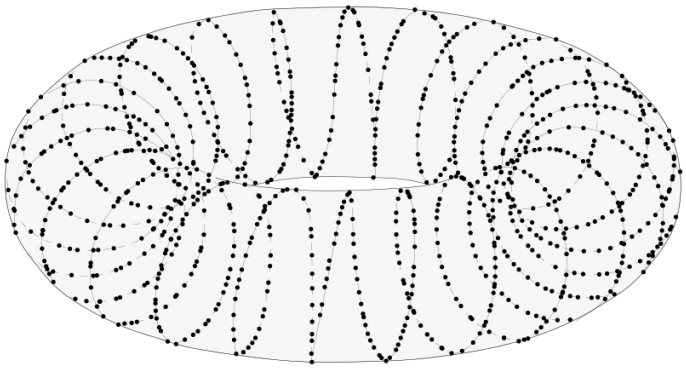
On the other hand, a *geometric property* depends on the way $\mathcal{M}$ is embedded in the Euclidean space. Recent works include the estimation of:

- its tangent spaces [TVF13, AL19],
- its geodesic distance [MLX+08, AB18, MSWW19, LD19],
- its curvature and fundamental forms [CCSL+17, AL19, BLM19b],
- its reach [AKC+19].

This thesis fits into the continuation of these works, particularly that of topological inference. We focus on the estimation of homotopy types, homology groups, Stiefel-Whitney characteristic classes and tangent spaces.

Stability and consistency. An estimation problem, such as those listed above, is usually answered by proposing an algorithm that, starting from the observation X, returns an estimator of a topological or geometric property of $\mathcal{M}$. A natural question is then to evaluate the quality of the estimator. In this thesis, we will evaluate it from two perspectives. The first quality of estimation is that of *consistency*. If we are observing $\mathcal{M}$ itself instead of X, how close is our estimator to the property to be estimated? If they coincide, the estimator is said to be consistent. Otherwise, it is biased, and the bias is to be quantified in order to attest the relevance of the estimator. Besides, the second quality of estimation is called *stability*. How close are the estimators computed from X and from $\mathcal{M}$? A stability result is typically written as a bound on the distance between these estimators, based on a bound on the distance between X and $\mathcal{M}$. Such stability is of critical importance in practice: it means that small measurement errors on X are tolerated, and will not affect the resulting estimation too much.

The problem of the scale. When answering to the estimation problems mentioned above, one faces the problem that topological and geometric features of the dataset strongly depend on the scale at which they are considered. For instance, the following point cloud can be seen as a sample of a curve, as well as a sample of the torus.



To illustrate this issue on a particular inference problem, say that we wish to estimate the homology groups of $\mathcal{M}$ from the observation X—a significant

proportion of this thesis being dedicated to this problem. Notice that the point cloud X, as it is given in practice, is a discrete topological space. Hence it does not contain any interesting topology in itself. However, if it is sampled close enough to $\mathcal{M}$, there exists a construction that allows to recover the homotopy type of $\mathcal{M}$ from X, hence its homology groups as well.

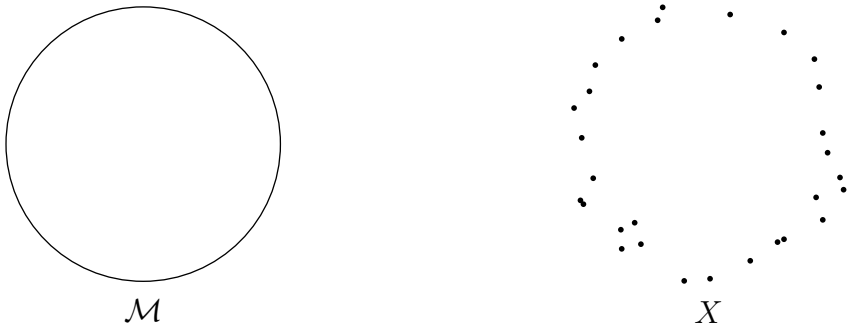
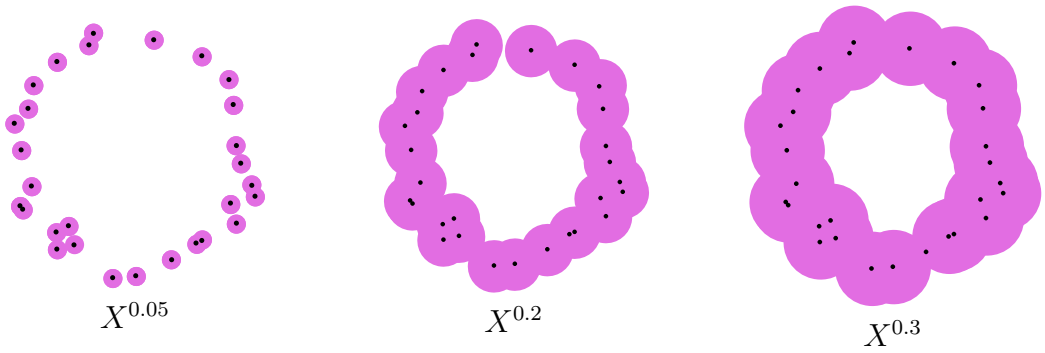


Figure I.17: The underlying space and the observation.

This construction consists in thickening X. For every $t \geq 0$, the t-thickening of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X.



Observe that the last figure is a thickening that matches the homotopy type of $\mathcal{M}$. In order to estimate topological properties associated to $\mathcal{M}$, we would like to select these t's such that X^t and $\mathcal{M}$ are homotopy equivalent. In other words, we would like to know at which scale X is to be seen. If certain geometric quantities are known, such as the reach of $\mathcal{M}$ and the density of X, then there exist procedures to select the thickenings X^t that recover the homotopy type of $\mathcal{M}$ [NSW08, CCSL09]. If no such properties are known a priori, this is where persistent homology theory comes in.

Persistent homology. Let us illustrate its use, in the context of estimating the i^{th} singular homology group $H_i(\mathcal{M})$. In the previous framework, we ended up with a collection of thickenings of X, some of which had the same homotopy type as $\mathcal{M}$. Instead of selecting these thickenings, the idea of persistent homology is to look at them all at once, and then to retrieve the homology groups of $\mathcal{M}$ from this collection.

The family of all thickenings is called the Čech filtration of X, and is denoted $V[X] = (X^t)_{t \geq 0}$. It is an increasing sequence of subsets

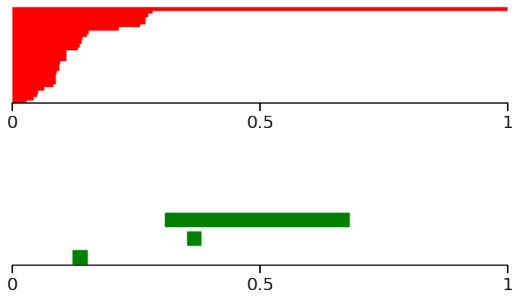
$$\ldots \subset X^{t_1} \subset X^{t_2} \subset X^{t_3} \subset \ldots$$

Applying the i^{th} homology functor yields a diagram of vector spaces

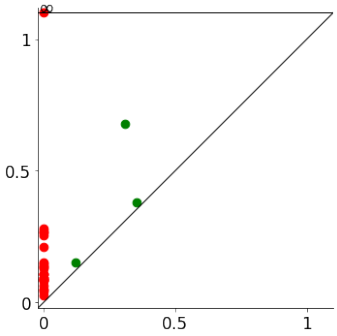
$$\ldots \longrightarrow H_i(X^{t_1}) \longrightarrow H_i(X^{t_2}) \longrightarrow H_i(X^{t_3}) \longrightarrow \ldots$$

Their collection forms a *persistence module*, denoted $\mathbb{V}[X] = (H_i(X^t))_{t\geq 0}$. This is the main object of persistent homology theory. It gathers the homology of X at every scale.

There has been a significant amount of work regarding the algebraic structure of the persistence modules. In some cases, for instance when X is finite, the persistence module $\mathbb{V}[X]$ can be written as a sum of indecomposable persistence modules. In the finite case, this result is a consequence of Gabriel's theorem about decomposition of quivers, and it was later extended to more general persistence modules [BCB20]. In general, a decomposition into indecomposable modules is unique, and results in a complete description of the persistence module in a convenient form, known as the *persistence barcode*. These barcodes summarize the evolution of the homology groups throughout the filtration. They can be pictured in two forms: persistence barcodes or persistence diagrams.



A persistence barcode



A persistence diagram

On a persistence barcode, one reads bars, that correspond to cycles of $H_i(X^t)$ at various values of t. The larger a bar is, the more the corresponding cycle persists in the filtration.

The previous figure shows the persistence barcode of the Čech filtration of a sample X of $\mathcal{M}$, where $\mathcal{M}$ the unit circle in $\mathbb{R}^2$, as in Figure I.17. The persistence module corresponding to the 0^{th} homology is drawn in red, and to the 1^{st} homology in green. One identifies two salient bars: a red one and a green one. They correspond to persisting cycles, which suggest that the underlying object $\mathcal{M}$ has a 0^{th} homology group of dimension 1, as well as a 1^{st} homology group of dimension 1. This is consistent with the homology of the circle. In other words, the persistence barcode built from X gives an idea of the homology of $\mathcal{M}$. More formally, in this thesis, we will build persistence diagrams, and show how to read the homology groups from them.

Contributions. This manuscript presents results on topological inference from point clouds, based on persistent homology theory. It consists in three distinct works, though interconnected. We give a brief description of them in this paragraph, and an extended presentation in the three following subsections. First of all, we consider datasets X corrupted by anomalous points. In this case, the Čech filtration

of X is not suited to estimate the homology of $\mathcal{M}$. Instead, we introduce the DTM-filtrations, and show that the homology of $\mathcal{M}$ can be read from their persistence diagrams. The stability of these filtrations involves the Wasserstein distance—a measure-theoretic distance that allows anomalous points in the dataset. Our second work focuses on datasets that lie close to an immersed manifold. When the underlying object $\mathcal{M}$ is not a submanifold, but an immersion of a manifold $\mathcal{M}_0$, the Čech filtration does not allow to estimate the homology of $\mathcal{M}_0$. We propose a way to adapt this method. The last work takes place in the context where the underlying object $\mathcal{M}$ is endowed with a vector bundle. In this case, its Stiefel-Whitney classes are defined, but there exists no general algorithm to compute them from a discrete approximation. We develop a persistent-theoretic framework for Stiefel-Whitney classes, and prove that they can be estimated from a sample X of the vector bundle.

With a view towards the practical applications of these methods, we provide a `Jupyter notebook` for each of them. They contain implementations of our methods, in `Python` language. They can be found following these links:

- DTM-based filtrations (Chapter III): `https://github.com/raphaeltinarrage/DTM-Filtrations/blob/master/Demo.ipynb`
- Topological inference for immersed manifolds (Chapter IV): `https://github.com/raphaeltinarrage/ImmersedManifolds/blob/master/Demo.ipynb`
- Persistent Stiefel-Whitney classes (Chapter V): `https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Demo.ipynb`

Moreover, all these functions are gathered in the package `velour`, available at `https://github.com/raphaeltinarrage/velour`.

I.2.2 Presentation of Chapter III: DTM-based filtrations

The problem of anomalous points. Let us get back to our initial framework: $\mathcal{M} \subset \mathbb{R}^n$ is a subset whose homology is to be estimated, based on the observation of a point cloud X. The procedure of persistent homology consists in building a filtration from X—that is, an increasing sequence of subsets—and then converting it into a persistence module by applying the homology functor. This persistence module is finally summarized in a barcode, from which we read topological features of X.

Among the many filtrations available to the user, the most common filtration is the Čech filtration $V[X]$, defined as the collection of thickenings of X. The corresponding i^{th} homology persistence module is denoted $\mathbb{V}[X]$. In practice, one can also use the Vietoris-Rips filtration, a simplicial variant, that is easier to compute. Their advantage is that they produce persistence diagrams that are robust to small variations of the input dataset in Hausdorff distance. This is known as the stability theorem [CSEH07, CDSO14]: if $\mathbb{V}[X]$ and $\mathbb{V}[\mathcal{M}]$ denote the persistence modules associated to the Čech filtrations of X and $\mathcal{M}$, then

$$\mathrm{d_i}\left(\mathbb{V}[X], \mathbb{V}[\mathcal{M}]\right) \leq \mathrm{d_H}\left(X, \mathcal{M}\right),$$

where $\mathrm{d_i}(\cdot,\cdot)$ is the interleaving distance between persistence modules, and $\mathrm{d_H}(\cdot,\cdot)$ the Hausdorff distance between subsets of $\mathbb{R}^n$. Besides this stability property, we have a consistency result: the persistence module $\mathbb{V}[\mathcal{M}]$ contains information about the homology of $\mathcal{M}$, provided that $\mathcal{M}$ is a smooth submanifold or a set with positive reach [CCSL09]. As a consequence, if $\mathrm{d_H}(X,\mathcal{M})$ is small, the usual procedure of persistent homology allows to recover the homology of $\mathcal{M}$ from the observation X.

However, if the set X contains anomalous points, then the Hausdorff distance $\mathrm{d_H}(X,\mathcal{M})$ may be large, and the stability theorem would not deliver a relevant bound. The homology of $\mathcal{M}$ cannot be inferred from the Čech filtration $V[X]$ anymore. In other words, these usual filtrations are sensitive to the presence of anomalous points in the dataset X, which makes them unsuitable for this context. An important observation is that the Hausdorff distance $\mathrm{d_H}(X,\mathcal{M})$ is no longer a relevant measure of proximity.

$\mathcal{M}$ X

In order to overcome this issue, we propose to reduce the importance of anomalous points in the filtration. We aim to obtain a filtration that behaves as if the anomalous points were not present. To this end, we will quantify the degree of anomalousness of the points, via a local density estimation. This idea requires to see the point cloud X as a measure.

A measure-theoretic point of view. Let us see the subsets X and $\mathcal{M}$ as probability measures, denoted μ and ν. For instance, if X is finite, μ may be the empirical measure on X, and if $\mathcal{M}$ is a compact submanifold, ν may be the uniform measure on $\mathcal{M}$. The *Wasserstein distance* $\mathrm{W}_2(\mu,\nu)$ between these measures allows to quantify the proximity between X and $\mathcal{M}$ while tolerating anomalous points. Is it possible to build filtrations $W[\mu]$ and $W[\nu]$ such that we have a stability of the form

$$\mathrm{d_i}(W[\mu],W[\nu]) \leq \text{constant} \cdot \mathrm{W}_2(\mu,\nu)? \tag{I.3}$$

Various works have been carried out in this direction in recent years. Among them, the filtration defined by the sublevel sets of the *distance-to-measure* (DTM), introduced in [CCSM11], and some of its variants [PWZ15], allow to capture topological information of the underlying space $\mathcal{M}$.

Unfortunately, from a practical perspective, the exact computation of the sublevel sets filtration of the DTM boils down to the computation of a k^{th} order Voronoï diagram, hence its persistent homology turns out to be far too expensive in most cases. To address this problem, [GMM13] introduces a variant of the DTM, the *witnessed k-distance*, whose persistence is easier to compute, and proves that the

witnessed k-distance approximates the DTM persistence up to an additive constant. There is also in [BL19] a study of another approximation of the DTM, called the *k-PDTM*. In [BCOS16, Buc14], a weighted version of the Vietoris-Rips complex filtration is introduced to approximate the persistence of the DTM function, and several stability and approximation results, comparable to the ones of [GMM13], are established.

We propose an alternative solution to this problem, with the intention of combining both aspects of the methods presented above: stability of the DTM, and computability of its variants. To this end, we introduce the *DTM-filtrations*. These filtrations are halfway between the Čech filtration and the DTM sublevel sets filtration: they are computable in practice, and they are robust against anomalous points.

DTM-filtrations. In order to build the DTM-filtrations, we start by generalizing the notion of *weighted Čech filtration*. We remind the reader that the usual Čech filtration of X at time t is defined as the following subset of $\mathbb{R}^n$:

$$V^t[X] = \bigcup_{x \in X} \overline{\mathcal{B}}\left(x, t\right),$$

where $\overline{\mathcal{B}}\left(x, t\right)$ is the closed ball of center x and radius t. In order to reduce the importance of an anomalous point x in X, we propose to postpone the apparition of the ball $\overline{\mathcal{B}}\left(x, t\right)$ in the filtration. This is achieved by using the modified ball $\overline{\mathcal{B}}_f(x, t)$. It is defined as

$$\overline{\mathcal{B}}_f(x, t) = \begin{cases} \emptyset & \text{if } t < f(x), \\ \overline{\mathcal{B}}\left(x, \left(t^p - f(x)^p\right)^{\frac{1}{p}}\right) & \text{otherwise.} \end{cases}$$

This ball depends on two parameters: a real number $f(x) \geq 0$, that control the delay of apparition of the ball in the filtration, and a parameter $p \in [0, +\infty]$, that controls its growing profile. Once we have chosen a map $f\colon X \to \mathbb{R}^+$, called the *weight function*, and a parameter $p \in [0, +\infty]$, we can define the weighted Čech filtration. It is denoted $V[X, f, p]$, and is defined as the collection of subsets

$$V^t[X, f, p] = \bigcup_{x \in X} \overline{\mathcal{B}}_f(x, t).$$

When $p = 2$, this construction already appeared in [Buc14, BCOS16]. Using classical results, we show that these filtrations are stable with respect to perturbations of X in the Hausdorff metric and perturbations of f with respect to the sup norm (Propositions III.4 and III.5).

Now, for a general function f, these stability results are not suited to deal with data containing anomalous points. We have to choose a weight function f that assigns greater values to them, so that their balls $\overline{\mathcal{B}}_f(x, t)$ appear late in the filtration. A function that fulfills this criterion is the DTM. It depends on a measure μ and a parameter $m \in (0, 1)$, and is denoted $\mathrm{d}_{\mu,m}\colon \mathbb{R}^n \to \mathbb{R}^+$. Given a measure μ, and two parameters $m \in (0, 1)$ and $p \in [0, +\infty]$, we define the *DTM-filtration* $W[\mu, m, p]$ as the collection of subsets

$$W^t[\mu, m, p] = \bigcup_{x \in \mathrm{supp}(\mu)} \overline{\mathcal{B}}_{\mathrm{d}_{\mu,m}}(x, t),$$

where $\mathrm{supp}(\mu)$ denotes the support of the measure μ. In other words, the DTM-filtration is the weighted Čech filtration with set $X = \mathrm{supp}(\mu)$, weight function $f = \mathrm{d}_{\mu,m}$ and parameter p. The corresponding persistence module is denoted $\mathbb{W}[\mu, m, p]$. If μ is the empirical measure on a finite set X, we simply denote $W[X, m, p]$ and $\mathbb{W}[X, m, p]$.

As an illustration, we consider in Figure I.18 a set X which is a sample of the unit circle with anomalous points. With the DTM-filtration, we see that the balls appear earlier on the underlying circle than on anomalous points.

Figure I.18: Left: The usual Čech filtration $V^t[X]$ at time $t = 0.3$. **Right:** The DTM-filtration $W^t[X, m, p]$ with parameters $m = 0.1$ and $p = 1$.

As a counterpart to the simplicity of their construction, the DTM-filtrations do not inherit a stability result as described by Equation (I.3). However, another formulation of their stability is obtained: the closeness between two DTM-filtrations $W[\mu, m, p]$ and $W[\nu, m, p]$ relies on the existence of intermediate measures which are both close to μ and ν in the Wasserstein metric.

Theorem III.22. Let μ and ν be two probability measures on $\mathbb{R}^n$, $m \in (0, 1)$ and $p \in [1, +\infty]$. For every probability measures μ', ν' such that $\mathrm{supp}(\mu') \subset \mathrm{supp}(\mu)$ and $\mathrm{supp}(\nu') \subset \mathrm{supp}(\nu)$, we have

$$\mathrm{d_i}\left(\mathbb{W}[\mu, m, p], \mathbb{W}[\nu, m, p]\right) \leq m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu, \mu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu', \nu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\nu', \nu\right) + c(\mu', m, p) + c(\nu', m, p).$$

The terms $c(\mu', m, p)$ and $c(\nu', m, p)$ are quantities that reflect the regularity of the measures μ' and ν' (defined in Equations (III.2) and (III.3)). For the theorem to give a relevant bound, these measures are to be chosen as clean versions of μ and ν, that is, with the anomalous points being removed. When $p = 1$, one obtains a more explicit stability result by assuming that the initial measure ν is (a, d)-standard (introduced in Subsection III.2.6).

Corollary III.24. If the measure ν is (a, d)-standard, then

$$\mathrm{d_i}\left(W[\mu, m, 1], W[\nu, m, 1]\right) \leq \text{constant} \cdot \left(\left(\frac{\mathrm{W}_2\left(\mu, \nu\right)}{m}\right)^{\frac{1}{2}} + m^{\frac{1}{d}}\right).$$

These results show that the DTM-filtration built on μ is a reliable estimator of the DTM-filtration built on ν, provided that the Wasserstein distance $\mathrm{W}_2\left(\mu, \nu\right)$ and the

parameter m are small.

Besides, we show a consistency result: the DTM-filtrations are close to the sublevel sets filtration of the DTM.

Corollary III.25. Let V be the sublevel sets filtration of the DTM $\mathrm{d}_{\nu,m}$. If the measure ν is (a,d)-standard, then

$$d_i(V,\ W[\nu,m,1]) \leq \text{constant} \cdot m^{\frac{1}{d}}.$$

By combining these stability and consistency results, and by choosing a parameter m of order $\mathrm{W}_2(\mu,\nu)^{\frac{d+1}{2}}$, we obtain

$$d_i(V,\ W[\mu,m,1]) \leq \text{constant} \cdot \mathrm{W}_2(\mu,\nu)^{\frac{d+1}{2d}}.$$

Hence the DTM-filtrations can be used as robust estimators of the homology of the space underlying the observations. This is illustrated by Figure I.19, which represents the persistence diagrams of the filtrations depicted in Figure I.18 (a Čech filtration and a DTM-filtration). Only the second diagram clearly exhibits the homology of a circle. Indeed, among the two points away from the diagonal, the red point indicates a 0^{th} homology group of dimension 1, and the green one indicates a 1^{st} homology group of dimension 1 as well.

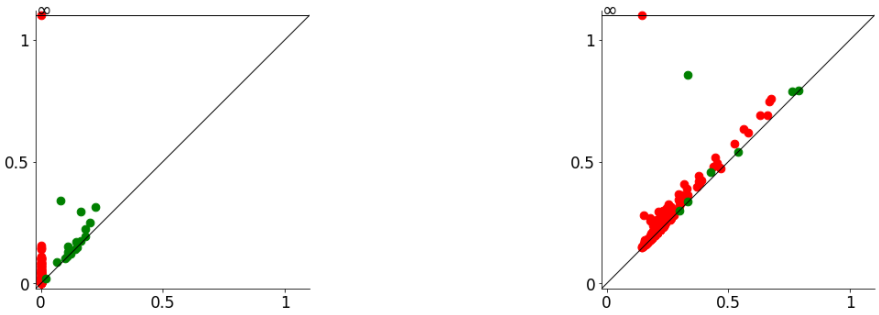


Figure I.19: Left: Persistence diagram of the usual Čech filtration $V[X]$. **Right:** Persistence diagram of the DTM-filtration $W[X,m,1]$.

Practical motivations. It is worth mentioning that the DTM-filtrations were first experimented in the setting of an industrial research project whose goal was to address an anomaly detection problem from inertial sensor data in bridge and building monitoring [Lab18]. In this problem, the input data comes as time series measuring the acceleration of devices attached to the monitored bridge or building. Using sliding windows and time-delay embedding, these times series are converted into a series of fixed size point clouds in $\mathbb{R}^n$. Filtrations are then built on top of these point clouds and their persistence is computed, giving rise to a time-dependent sequence of persistence diagrams, that are then used to detect anomalies or specific features occurring along the time [SDB16, Ume17]. In this practical setting it turned out that the DTM-filtrations reveal to be not only more resilient to noise but also able to better highlight topological features in the data than the standard Vietoris-Rips filtrations.

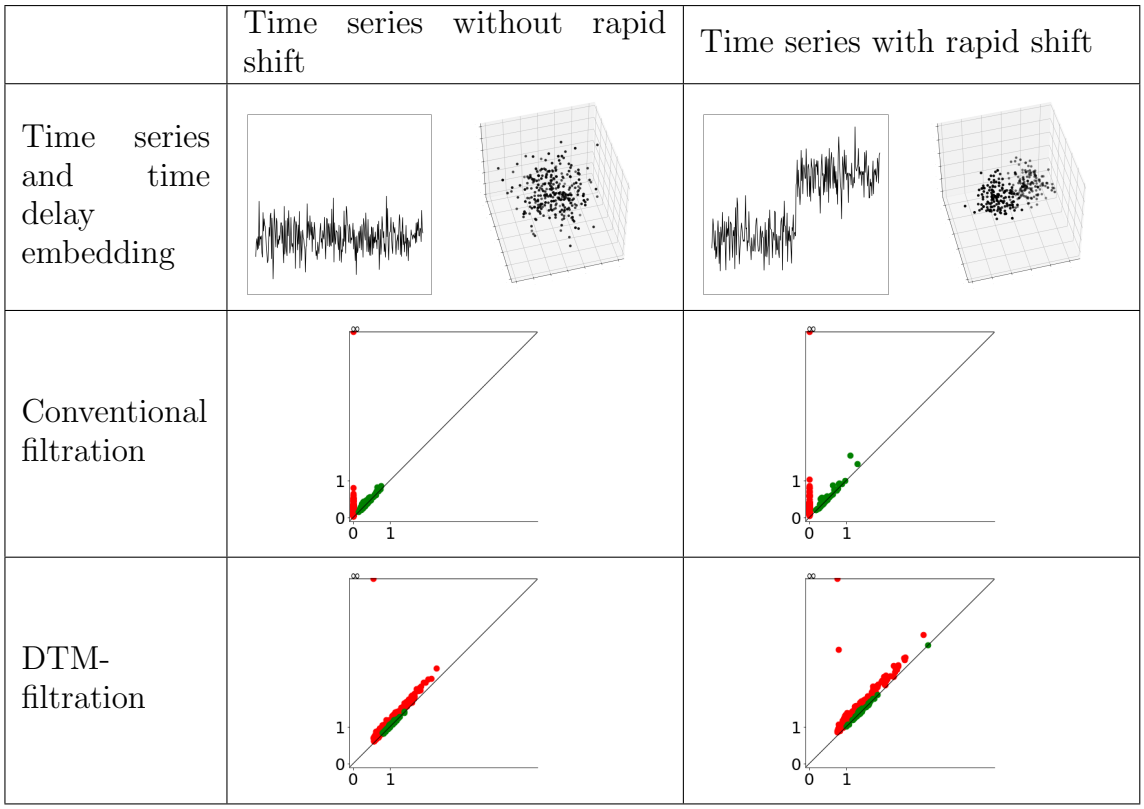


Figure I.20: Comparison of the Vietoris-Rips filtration and the DTM-filtration.

Figure I.20 is a synthetic example comparing Vietoris-Rips filtration to DTM-filtration. The first row represents two time series with very different behavior and their embedding into $\mathbb{R}^3$. Here a series $(x_1, x_2, \ldots, x_n)$ is converted in the 3D point cloud $\{(x_1, x_2, x_3), (x_2, x_3, x_4), \ldots, (x_{n-2}, x_{n-1}, x_n)\}$. The second row shows the persistence diagrams of the Vietoris-Rips filtration built on top of the two point clouds (red and green points represent respectively the 0-dimensional and 1-dimensional diagrams). One observes that the diagrams do not clearly detect the different behavior of the time series. The third row shows the persistence diagrams of the DTM-filtration built on top of the two point clouds. A red point clearly appears away from the diagonal in the second diagram, which is a consequence of the rapid shift occurring in the second time series.

I.2.3 Presentation of Chapter IV: Topological inference for immersed manifolds

Topological inference from immersed manifolds. We know that if the object $\mathcal{M}$ is a smooth submanifold, and if the observation X is sufficiently close in Hausdorff distance, then the persistent homology of X allows to estimate the homology groups of $\mathcal{M}$. In this context, one uses the Čech filtration of X. However, in practice, the smooth submanifold assumption may be too restrictive. There has been some work aimed at reducing the regularity of $\mathcal{M}$. For instance, if we only assume that its reach is positive, or even its μ-reach, then it has been shown that the Čech filtration of X still allows to recover the homology of $\mathcal{M}$ [CCSL09].

Here, we will study a different case of regularity: $\mathcal{M}$ is an immersed manifold, not embedded. In this framework, there exists an abstract manifold $\mathcal{M}_0$, immersed in the Euclidean space via a map $u\colon \mathcal{M}_0 \to \mathbb{R}^n$, whose image is $\mathcal{M}$. As before, the observation X is a subset of $\mathbb{R}^n$, that we suppose close to $\mathcal{M}$. Our goal is to infer the homology of $\mathcal{M}_0$ from X.

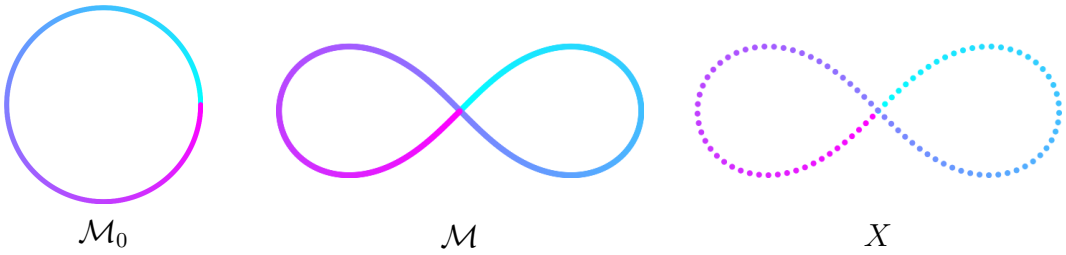


Figure I.21: Left: The abstract manifold $\mathcal{M}_0$, a circle. **Middle:** The immersion $\mathcal{M} \subset \mathbb{R}^2$, known as the lemniscate of Bernoulli. **Right:** The observation X.

Being an immersion, $\mathcal{M}$ may self-intersect, and the sets $\mathcal{M}_0$ and $\mathcal{M}$ may have different homotopy types. The Čech filtration of $\mathcal{M}$, or X, would reveal the homology of $\mathcal{M}$, not that of $\mathcal{M}_0$. Consequently, the usual approach based on the Čech filtration no longer applies here, and new methods must be developed.

Among the works that involve immersed manifolds, let us cite [ACLZ17], which is set in the context where $\mathcal{M}$ is a union of intersecting submanifolds. Hence $\mathcal{M}$ is not a submanifold itself, but it is an immersed manifold. The authors propose an algorithm to classify the different components of $\mathcal{M}$, that is, the connected components of $\mathcal{M}_0$. This algorithm is based on the estimation of tangent spaces, so as to separate the set $\mathcal{M}$ where it self-intersects; this is a point of view that we also adopt.

Lifting immersed manifolds. In order to estimate the homology of a manifold from an immersion of it, we propose to estimate its tangent bundle, seen as a subset of another Euclidean space. As we will see, in the process of estimating this tangent bundle, we will make errors, which will result in anomalous points. This issue will be solved by using the DTM-filtrations, described before. We therefore consider a measure-theoretic framework from now on. Let us describe the method.

Let $\mathcal{M}_0$ be a compact $\mathcal{C}^2$-manifold of dimension d, and μ_0 a Radon probability measure on $\mathcal{M}_0$ with full support. Let $u\colon \mathcal{M}_0 \to \mathbb{R}^n$ be an immersion. We assume that the immersion is such that self-intersection points correspond to different tangent spaces. In other words, for every $x_0, y_0 \in \mathcal{M}_0$ such that $x_0 \neq y_0$ and $u(x_0) = u(y_0)$, the tangent spaces $d_{x_0}u(T_{x_0}\mathcal{M}_0)$ and $d_{y_0}u(T_{y_0}\mathcal{M}_0)$ of $\mathcal{M}_0$, seen in $\mathbb{R}^n$, are different. Define the image of the immersion $\mathcal{M} = u(\mathcal{M}_0)$ and the pushforward measure $\mu = u_*\mu_0$. We suppose that we are observing the measure μ, or a close measure ν. Our goal is to infer the singular homology of $\mathcal{M}_0$ (with coefficients in $\mathbb{Z}_2$ for instance) from ν.

To get back to $\mathcal{M}_0$, we proceed as follows: let $\mathrm{M}(\mathbb{R}^n)$ be the vector space of

$n \times n$ matrices, and $\check{u}\colon \mathcal{M}_0 \to \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ the application

$$\check{u}\colon x_0 \longmapsto \left(u(x_0),\ \frac{1}{d+2} p_{T_x\mathcal{M}}\right),$$

where $p_{T_x\mathcal{M}}$ is the matrix of the orthogonal projection on the tangent space $T_x\mathcal{M} \subset \mathbb{R}^n$, written in the canonical basis of $\mathbb{R}^n$. The set $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$ is a submanifold of $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, diffeomorphic to $\mathcal{M}_0$. It is called the *lift* of $\mathcal{M}_0$. The space $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ is called the *lift space*.

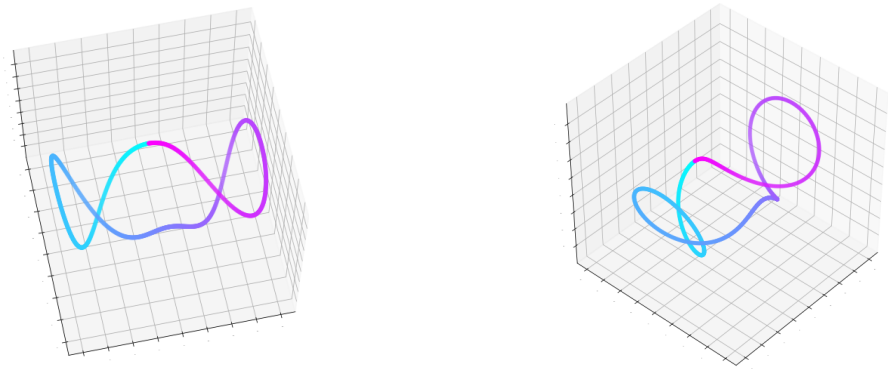


Figure I.22: Two views of the submanifold $\check{\mathcal{M}} \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, projected in a 3-dimensional subspace via PCA. Observe that it does not self-intersect. The initial set $\mathcal{M}$ is represented in Figure I.21.

Suppose that one is able to estimate $\check{\mathcal{M}}$ from ν. Then one could consider the persistent homology of a filtration based on $\check{\mathcal{M}}$—say the Čech filtration of $\check{\mathcal{M}}$ in the ambient space $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$ for instance—and hope to read the singular homology of $\mathcal{M}_0$ in the corresponding persistent barcode.

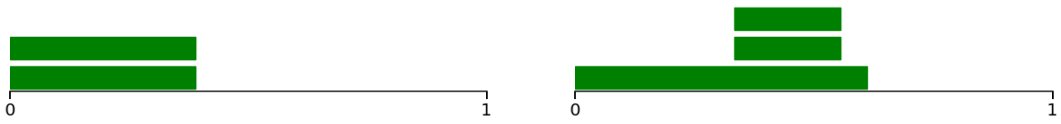


Figure I.23: Left: Persistence barcode of the 1-homology of the Čech filtration of $\mathcal{M}$ in the ambient space $\mathbb{R}^2$. One reads the 1-homology of the lemniscate. **Right:** Persistence barcode of the 1-homology of the Čech filtration of $\check{\mathcal{M}}$ in the lift space $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$. One reads the 1-homology of a circle (at the beginning of the filtration).

Unfortunately, we won't be able to give a good estimation of $\check{\mathcal{M}}$. This is because the tangent spaces $p_{T_x\mathcal{M}}$, that we compute via local covariance matrices, won't be estimated correctly if x is too close to a self-intersection of $\mathcal{M}$. Thus, instead of estimating the lifted submanifold $\check{\mathcal{M}}$, we propose to estimate the *exact lifted measure* $\check{\mu}_0$, defined as $\check{\mu}_0 = \check{u}_*\mu_0$. It is a measure on the lift space $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, with support $\check{\mathcal{M}}$. Using measure-based filtrations—such as the DTM-filtrations—one can also hope to recover the singular homology of $\mathcal{M}_0$.

It is worth noting that $\check{\mathcal{M}}$ can be naturally seen as a submanifold of $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^n)$, where $\mathcal{G}_d(\mathbb{R}^n)$ denotes the Grassmannian of d-dimensional linear subspaces of $\mathbb{R}^n$. From this point of view, $\check{\mu}_0$ can be seen as a measure on $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^n)$, i.e., a varifold.

However, for computational reasons, we choose to work in the matrix space $\mathrm{M}(\mathbb{R}^n)$ instead of $\mathcal{G}_d(\mathbb{R}^n)$.

Here is an alternative definition of $\check{\mu}_0$: for any test function $\phi\colon \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n) \to \mathbb{R}$,

$$\int \phi(x, A) \mathrm{d}\check{\mu}_0(x, A) = \int_{\mathcal{M}_0} \phi\left(u(x_0), \frac{1}{d+2} p_{T_x\mathcal{M}}\right) \mathrm{d}\mu_0(x_0).$$

Getting back to the observed measure ν, we propose to estimate $\check{\mu}_0$ with the *lifted measure* $\check{\nu}$, defined as follows: for any test function $\phi\colon \mathbb{R}^n \times \mathrm{M}(E) \to \mathbb{R}$,

$$\int \phi(x, A) \mathrm{d}\check{\nu}(x, A) = \int_{\mathcal{M}} \phi\Big(x, \overline{\Sigma}_\nu(x)\Big) \mathrm{d}\nu(x),$$

where $\overline{\Sigma}_\nu(x)$ is normalized local covariance matrix (defined in Section IV.3). It depends on a parameter $r > 0$. We prove that $\overline{\Sigma}_\nu(x)$ can be used to estimate the tangent spaces $\frac{1}{d+2} p_{T_x\mathcal{M}}$ of $\mathcal{M}$. However, this estimation is biased next to the self-intersection of $\mathcal{M}$, as shown in Figure I.24. In order to quantify the quality of this approximation, we introduce a new geometric quantity.

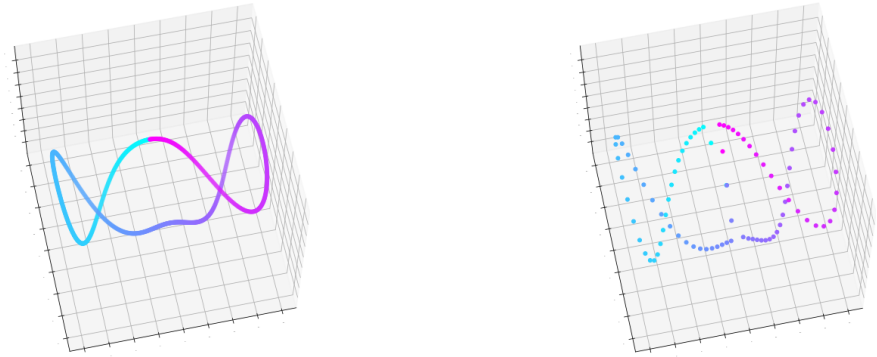


Figure I.24: Left: The sets $\mathrm{supp}(\mu) = \mathcal{M}$ and $\mathrm{supp}(\check{\mu}_0) = \check{\mathcal{M}}$, where μ is the uniform measure on $\mathcal{M}$ (see Figure I.21). **Right:** The sets $\mathrm{supp}(\nu)$ and $\mathrm{supp}(\check{\nu})$, where ν is the empirical measure on X. Parameters $\gamma = 2$ and $r = 0.1$.

Normal reach. Since the usual notion of reach is no longer relevant in the case of immersed manifolds, we introduce the normal reach—a function that indicates locally the extent to which the immersed manifold can be seen as a submanifold (see Subsection IV.2.2). It is denoted $\lambda\colon \mathcal{M} \to \mathbb{R}^+$. Our results are also expressed in terms of ρ, a measurement of curvature of $\mathcal{M}_0$.

Theorem IV.11. Let $\mathcal{M}$ be an immersed $\mathcal{C}^2$-manifold. Let $x \in \mathcal{M}$ and $r < \min\left\{\frac{1}{4\rho}, \lambda(x)\right\}$. Then $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ is a set of reach at least $\frac{1-2\rho r}{\rho}$.

When $\mathcal{M}$ is an embedded manifold, we connect the normal reach to the usual notion of reach.

Proposition IV.10. Let $\mathcal{M}$ be an embedded $\mathcal{C}^2$-manifold. We have

$$\mathrm{reach}\,(\mathcal{M}) = \min\left\{\frac{1}{\rho_*}, \frac{1}{2}\lambda_*\right\},$$

where ρ_* is the supremum of the operator norms of the second fundamental forms of $\mathcal{M}_0$, and $\lambda_* = \inf_{x\in\mathcal{M}} \lambda(x)$ is the infimum of the normal reach.

The normal reach allows to quantify the quality of approximation of the exact lifted measure $\check{\mu}_0$ with the lifted measure $\check{\nu}$. We prove a global estimation result, of the following form: $\check{\mu}_0$ and $\check{\nu}$ are close in the Wasserstein metric, as long as μ and ν are. We use a modified version of the Wasserstein distance in the lift space, denoted $\mathrm{W}_{p,\gamma}(\cdot,\cdot)$, which depends on a parameter $\gamma > 0$ (defined in Subsection IV.3.1). This parameter is designed to balance the importance given to the spatial information ($\mathbb{R}^n$-component) and the tangent space information ($\mathrm{M}(\mathbb{R}^n)$-component) in $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$. Moreover, our results rely on the technical Hypotheses 1, 2, 3 and 4, that we describe in Subsection IV.1.1.

Theorem IV.33. Let $\mathcal{M}$ be an immersed manifold that satisfies Hypotheses 1, 2 and 3. Let $\gamma, r > 0$. If $\mathrm{W}_p(\mu,\nu)$ and r are small enough, then

$$\mathrm{W}_{p,\gamma}(\check{\nu},\check{\mu}_0) \leq \text{constant}\cdot\gamma\cdot\left(\mu(\lambda^r)^{\frac{1}{p}} + r + \left(\frac{\mathrm{W}_p(\mu,\nu)}{r^{d+1}}\right)^{\frac{1}{2}}\right) + 2\mathrm{W}_p(\mu,\nu).$$

The quantity $\mu(\lambda^r)$ refers to $\mu(\lambda^{-1}([0,r]))$, that is, the measure of points of $\mathcal{M}$ with small normal reach. By adding an assumption about this quantity, we obtain a simpler result.

Corollary IV.35. In addition, if the immersion $\mathcal{M}$ satisfies Hypothesis 4, then

$$\mathrm{W}_{p,\gamma}(\check{\nu},\check{\mu}_0) \leq \big(1 + \text{constant}\cdot\gamma\big) r^{\frac{1}{p}}.$$

As a consequence of this result, we are able to estimate the homotopy type of the manifold $\mathcal{M}_0$ (Corollary IV.38). By using DTM-filtrations, we also recover the homology of $\mathcal{M}_0$ (Corollary IV.42).

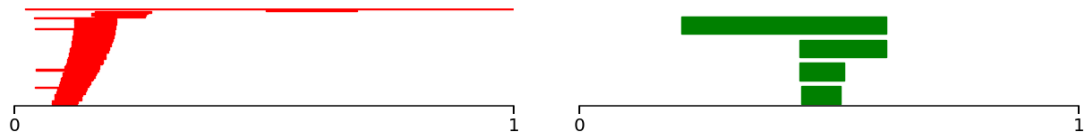


Figure I.25: Persistence barcodes of the 0-homology **(left)** and 1-homology **(right)** of the DTM-filtration of the lifted measure $\check{\nu}$. Observe that the homology of the circle is salient on these barcodes (one large red bar and one large green bar). Parameters $\gamma = 2$, $r = 0.1$ and $m = 0.01$.

Tangent space estimation. Our method relies on the estimation of the tangent spaces of $\mathcal{M}$ from the observation of X, via local covariance matrices. We study them by restricting the measures μ and ν to balls $\overline{\mathcal{B}}(x,r)$. The corresponding probability measures are denoted $\overline{\mu_x}$ and $\overline{\nu_x}$. We show that the restricted probability measures inherit the initial Wasserstein distance (Equation (IV.25)). Local stability of measures have also been studied in [MMM18, MSW19] and tangent space estimation via local covariance matrices in [ACLZ17], and our result improve the one of [MSW19].

As a consequence of the stability, and using the normal reach, we prove that $\overline{\Sigma}_\mu(x)$ is a consistent estimator of the tangent spaces $\frac{1}{d+2}p_{T_x\mathcal{M}}$ of $\mathcal{M}$, and that it is stable with respect to μ.

Proposition IV.24 and Equation (IV.26). Let $\mathcal{M}$ be an immersed manifold and μ a measure that satisfy Hypotheses 2 and 3. Let $x \in \mathcal{M}$ and $r < \min\left\{\frac{1}{2\rho}, \lambda(x)\right\}$. We have

$$\left\|\overline{\Sigma}_\mu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} \leq \text{constant} \cdot r.$$

Moreover, if ν is any other probability measure, then for $\mathrm{W}_1(\mu,\nu)$ small enough, we have

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq \text{constant} \cdot \left(\frac{\mathrm{W}_1(\mu,\nu)}{r^{d+1}}\right)^{\frac{1}{2}}.$$

By choosing a radius r of order $\mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}$, we obtain

$$\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} \leq \text{constant} \cdot \mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}.$$

In other words, the local covariance matrices of ν estimate the tangent spaces of $\mathcal{M}$ at speed $\mathrm{W}_1(\mu,\nu)^{\frac{1}{d+3}}$. This bound is not as tight as state of the art methods of tangent space estimation [AL19]. However, our result is quite general, as it only assume that the measure ν is close to μ in Wasserstein distance.

I.2.4 Presentation of Chapter V: Persistent Stiefel-Whitney classes

Other topological invariants. As before, let $\mathcal{M}$ and X be subsets of $\mathbb{R}^n$, where X is seen as a sample of $\mathcal{M}$. So far, we have focused on estimating the homology groups of $\mathcal{M}$ from X. However, there exists in algebraic topology many other invariants associated to $\mathcal{M}$, which are interesting to estimate. Some examples are given by the cohomology ring $H^*(\mathcal{M})$ of $\mathcal{M}$, or the cohomology operations in $H^*(\mathcal{M})$. Their estimation has been studied in the context of persistence theory. In [Yar10], the author propose an algorithm to compute the cup-product, and in [Aub11] an algorithm to compute the Steenrod squares.

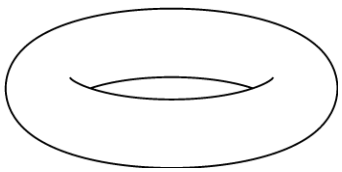
The Stiefel-Whitney classes are another example of invariant. They are associated to any topological space endowed with a vector bundle. We consider the following inference problem: $\mathcal{M}$ is endowed with a vector bundle, and we aim to estimate its Stiefel-Whitney classes from the observation X. Our approach is based on a persistent-theoretic framework: we define persistent Stiefel-Whitney classes, and give consistency and stability results.

To our knowledge, the problem of estimating Stiefel-Whitney classes has not been adressed in the context of a point cloud observation. Still, in [Aub11], the author propose an algorithm to compute the Stiefel-Whitney classes, in the particular case of the tangent bundle of a Euler mod-2 space (that is, a simplicial complex that triangulates a manifold).

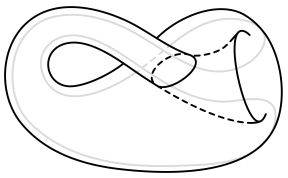
Stiefel-Whitney classes as a refinement of cohomology. In practice, the Stiefel-Whitney classes provide more topological information than the cohomolgy groups alone. To see this, let $\mathcal{M}_0$ and $\mathcal{M}'_0$ denote the torus and the Klein bottle. Only one of them is orientable, hence these two manifolds are not homeomorphic. Let $\mathbb{Z}_2$ be the field with two elements. Observe that the cohomology groups of $\mathcal{M}_0$ and $\mathcal{M}'_0$ over $\mathbb{Z}_2$ are equal:

$$\begin{aligned} H^0(\mathcal{M}_0) &= H^0(\mathcal{M}'_0) = \mathbb{Z}_2, \\ H^1(\mathcal{M}_0) &= H^1(\mathcal{M}'_0) = \mathbb{Z}_2 \times \mathbb{Z}_2, \\ H^2(\mathcal{M}_0) &= H^2(\mathcal{M}'_0) = \mathbb{Z}_2. \end{aligned}$$

Therefore, the cohomology groups alone do not permit to differentiate the manifolds $\mathcal{M}_0$ and $\mathcal{M}'_0$. To do so, several refinements from algebraic topology may be used. Here, we will study the Stiefel-Whitney classes. If we equip $\mathcal{M}_0$ and $\mathcal{M}'_0$ with their tangent bundles, their first Stiefel-Whitney classes are distinct: only one of them is zero. Hence we are able to differentiate these two manifolds.



$H^*(\mathcal{M}_0) = \mathbb{Z}_2[x,y]/\langle x^2, y^2 \rangle$
$w_1(\tau_{\mathcal{M}_0}) = 0$



$H^*(\mathcal{M}'_0) = \mathbb{Z}_2[x,y]/\langle x^3, y^2, x^2y \rangle$
$w_1(\tau_{\mathcal{M}'_0}) = x$

Figure I.26: The cohomology rings of $\mathcal{M}_0$ and $\mathcal{M}'_0$ over $\mathbb{Z}_2$, and the first Stiefel-Whitney classes of their respective tangent bundles $\tau_{\mathcal{M}_0}$ and $\tau_{\mathcal{M}'_0}$.

More generally, Stiefel-Whitney classes have been widely used in differential topology, for instance in the problem of immersing manifolds in low-dimensional spaces, or in cobordism problems [MS16]. Our work is motivated by introducing this tool in persistent homology theory.

Persistent Stiefel-Whitney classes. In general, if X is a topological space endowed with a vector bundle ξ of dimension d, there exists a collection of cohomology classes $w_1(\xi), ..., w_d(\xi)$, the Stiefel-Whitney classes, such that $w_i(\xi)$ is an element of the cohomology group $H^i(X)$ over $\mathbb{Z}_2$ for $i \in [1, d]$. In order to define Stiefel-Whitney classes in a persistent-theoretic framework, we will use a convenient definition of vector bundles: defining a vector bundle over a compact space X is equivalent to defining a continuous map $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ for m large enough, where $\mathcal{G}_d(\mathbb{R}^m)$ is the Grassmann manifold of d-planes in $\mathbb{R}^m$. Such a map is called a *classifying map* for ξ. It is closely related to the Gauss map of submanifolds of $\mathbb{R}^3$, as explained in Figure I.27.

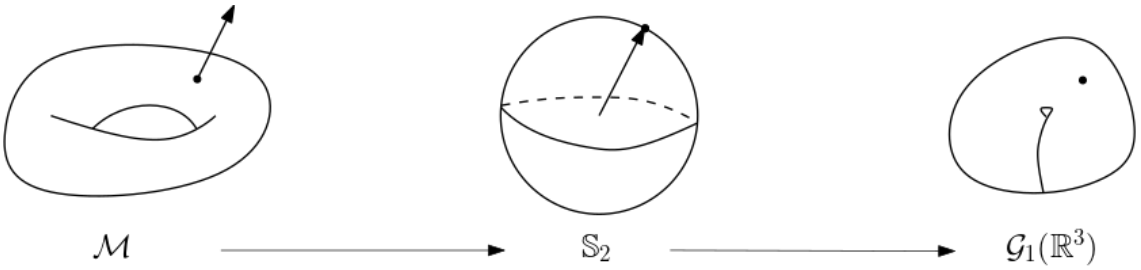


Figure I.27: If $\mathcal{M}$ is an orientable 2-submanifold of $\mathbb{R}^3$, the Gauss map $g\colon \mathcal{M} \to \mathbb{S}_2$ maps every $x \in \mathcal{M}$ to a normal vector of $\mathcal{M}$ at x. By post-composing this map with the usual quotient map $\mathbb{S}_2 \to \mathcal{G}_1(\mathbb{R}^3)$, we obtain a classifying map $p\colon \mathcal{M} \to \mathcal{G}_1(\mathbb{R}^3)$ for the normal bundle of $\mathcal{M}$.

Given a classifying map $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ of a vector bundle ξ, the Stiefel-Whitney classes $w_1(\xi), ..., w_d(\xi)$ can be defined by pushing forward some particular classes of the Grassmannian via the induced map in cohomology $p^*\colon H^*(X) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^m))$. If w_i denotes the i^{th} Stiefel-Whitney class of the Grassmannian, then the i^{th} Stiefel-Whitney class of the vector bundle ξ is

$$w_i(\xi) = p^*(w_i). \tag{I.4}$$

In order to translate these considerations in a persistent-theoretic setting, suppose that we are given a dataset of the form (X, p), where X is a finite subset of $\mathbb{R}^n$, and p is a map $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$. Denote by $(X^t)_{t\geq 0}$ the Čech filtration of X, that is, the collection of the t-thickenings X^t of X in the ambient space $\mathbb{R}^n$. In order to define some *persistent* Stiefel-Whitney classes, one would try to extend the map $p\colon X \to \mathcal{G}_d(\mathbb{R}^m)$ to $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$. However, we did not find any interesting way to extend this map. To overcome this issue, we propose to look at the dataset in a different way. Transform the vector bundle (X, p) into a subset of $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$, via

$$\check{X} = \{(x, p(x)), x \in X\}.$$

The Grassmann manifold $\mathcal{G}_d(\mathbb{R}^m)$ can be naturally embedded in $\mathrm{M}(\mathbb{R}^m)$, the space of $m \times m$ matrices. From this viewpoint, $\check{X}$ can be seen as a subset of $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Let $(\check{X}^t)_{t\geq 0}$ denotes the Čech filtration of $\check{X}$ in the ambient space $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. A natural map $p^t\colon \check{X}^t \to \mathcal{G}_d(\mathbb{R}^m)$ can be defined: map a point $(x, A) \in \check{X}^t$ to the projection of A on $\mathcal{G}_d(\mathbb{R}^m)$, seen as a subset of $\mathrm{M}(\mathbb{R}^m)$:

$$p^t\colon (x, A) \in \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m) \longmapsto \mathrm{proj}\,(A, \mathcal{G}_d(\mathbb{R}^m)).$$

The projection is well-defined if A does not belong to the medial axis of $\mathcal{G}_d(\mathbb{R}^m)$. We show that this condition can be verified in practice (Lemma V.3). The Čech filtration of $\check{X}$, endowed with the extended projection maps $(p^t\colon \check{X}^t \to \mathcal{G}_d(\mathbb{R}^m))_t$, is called the *Čech bundle filtration*. Now we can define the i^{th} persistent Stiefel-Whitney class as the collection of classes $w_i(X) = (w_i^t(X))_t$, where $w_i^t(X)$ is the push-forward

$$w_i^t(X) = (p^t)^*(w_i),$$

and where w_i is the i^{th} Stiefel-Whitney class of the Grassmann manifold (compare with Equation (I.4)). We summarize the information given by a persistent Stiefel-Whitney class in a diagram, that we call a *lifebar*.

Stability and consistency. To illustrate our results, consider the embedding of the torus $u\colon \mathcal{M}_0 \to \mathcal{M} \subset \mathbb{R}^3$ depicted in Figure I.28. Denote by $p_{T_x\mathcal{M}}$ the projection matrix on the tangent space of $\mathcal{M}$ at x. The set $\check{\mathcal{M}} = \{(x, p_{T_x\mathcal{M}}), x \in \mathcal{M}\}$ is a subset of $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$.

Figure I.28: The submanifold $\mathcal{M} \subset \mathbb{R}^3$, and the submanifold $\check{\mathcal{M}} \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3) \simeq \mathbb{R}^{12}$ projected in a 3-dimensional subspace via PCA.

The lifebar of the first persistent Stiefel-Whitney class of this torus is depicted in Figure I.29. The bar is hatched, which means that the class is zero all along the filtration. This is coherent with the actual first Stiefel-Whitney class of the normal bundle of the torus, which is zero.

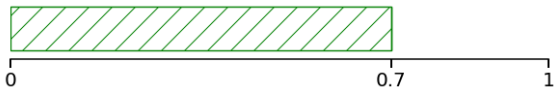


Figure I.29: Lifebar of the first persistent Stiefel-Whitney class of $\check{\mathcal{M}}$. It is only defined on the interval $\left[0, \frac{\sqrt{2}}{2}\right)$ (see Definition V.4).

To continue, consider the immersion of the Klein bottle $u'\colon \mathcal{M}'_0 \to \mathcal{M}' \subset \mathbb{R}^3$ depicted in Figure I.30. For $x_0 \in \mathcal{M}'_0$, denote by $p_{T_{x_0}\mathcal{M}}$ the projection matrix on the tangent space of $\mathcal{M}'_0$ at x_0, seen in $\mathbb{R}^3$. The set $\check{\mathcal{M}}' = \left\{\left(u(x_0), p_{T_{x_0}\mathcal{M}}\right), x_0 \in \mathcal{M}'_0\right\}$ is a subset of $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$. Note that $\check{\mathcal{M}}'$ is a submanifold (diffeomorphic to the Klein bottle), while $\mathcal{M}'$ is not.

Figure I.30: The set $\mathcal{M}' \subset \mathbb{R}^3$, and the submanifold $\check{\mathcal{M}}' \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3) \simeq \mathbb{R}^{12}$ projected in a 3-dimensional subspace via PCA.

Just as before, we can define persistent Stiefel-Whitney classes over the Čech filtration of $\check{\mathcal{M}}'$. Figure I.31 represents the lifebar of the first Stiefel-Whitney class of this filtration. The bar is filled, which means that the class is nonzero all along the

filtration. This is coherent with the first Stiefel-Whitney class of the normal bundle of the Klein bottle.

Figure I.31: Lifebar of the first persistent Stiefel-Whitney class of $\check{\mathcal{M}}'$.

The construction we propose is defined for any subset $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. In particular, it can be applied to finite samples of $\check{\mathcal{M}}$ and $\check{\mathcal{M}}'$. We prove that this construction is stable, a result reminiscent of the usual stability theorem of persistent homology.

> **Corollary V.8.** Consider two subsets $X, Y \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ such that $\mathrm{d_H}(X, Y) \leq \epsilon$. Then there exists an ϵ-interleaving between the persistence comodules of their Čech vector bundle filtrations such that the persistent Stiefel-Whitney classes are sent onto persistent Stiefel-Whitney classes.

We also show that the persistent Stiefel-Whitney classes are consistent estimators of Stiefel-Whitney classes.

> **Corollary V.11.** Let $\mathcal{M}_0 \to \mathcal{M} \subset \mathbb{R}^n$ be an immersion, $p\colon \mathcal{M}_0 \to \mathcal{G}_d(\mathbb{R}^m)$ a vector bundle, and $\check{\mathcal{M}} \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ the corresponding lifted manifold. Let $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ be any subset such that $\mathrm{d_H}(X, \mathcal{M}) \leq \epsilon$. Then for every $t \in [4\epsilon, \mathrm{reach}\left(\check{\mathcal{M}}\right) - 3\epsilon)$, the composition of inclusions $\mathcal{M}_0 \hookrightarrow \check{\mathcal{M}} \hookrightarrow X^t$ induces an isomorphism $H^*(\mathcal{M}_0) \leftarrow H^*(X^t)$ which sends the i^{th} persistent Stiefel-Whitney class $w_i^t(X)$ of the Čech bundle filtration of X to the i^{th} Stiefel-Whitney class of $(\mathcal{M}_0, p)$.

As an illustration, Figure I.32 represents the lifebars of the first persistent Stiefel-Whitney classes of samples X and X' of $\check{\mathcal{M}}$ and $\check{\mathcal{M}}'$. Observe that they are close to the original ones.

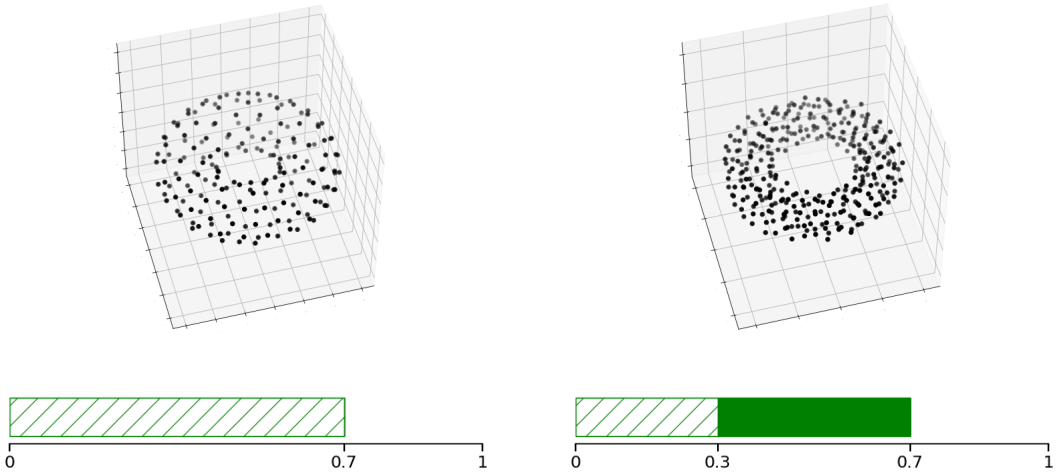


Figure I.32: Left: a sample of $\check{\mathcal{M}} \subset \mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$, seen in $\mathbb{R}^3$, and the lifebar of its first persistent Stiefel-Whitney class. **Right:** same for $\check{\mathcal{M}}'$.

An algorithm. We propose a concrete algorithm to compute the persistent Stiefel-Whitney classes. This algorithm is based on several ingredients, including the triangulation of projective spaces, and the simplicial approximation method.

The simplicial approximation, widely used in theory, can be applied only if the simplicial complex is refined enough, a property that is attested by the *star condition*. However, this condition cannot be verified in practice. We circumvent this problem by introducing the *weak star condition*, a variant that only depends on the combinatorial structure of the simplicial complex. When the simplicial complex is fine enough, the star condition and the weak star condition turn out to be equivalent notions (Proposition V.20).

II Background

II.1 Notations 57
II.2 Background on differential geometry 60
II.2.1 Basic notions of differential geometry 60
II.2.2 Basic notions of simplicial topology 61
II.2.3 Basic notions of vector bundle theory 65
II.2.4 Basic notions of Stiefel-Whitney classes 68
II.2.5 Basic notions of Riemannian geometry 69
II.3 Background on Euclidean geometry of compact sets 71
II.3.1 Basic notions of topology 71
II.3.2 Thickenings and tubular neighborhoods 72
II.3.3 Reach 75
II.3.4 Weak feature size and μ-reach 76
II.4 Background on persistent homology 78
II.4.1 Basic notions of singular and simplicial homology 78
II.4.2 Persistence modules 81
II.4.3 Decomposition of persistence modules 85
II.4.4 Stability of persistence modules 87
II.4.5 Persistent cohomology theory 88
II.5 Background on measure theory 90
II.5.1 Basic notions of measure theory 90
II.5.2 Distance-to-measure 91
II.5.3 Varifolds 92
II.6 Homology inference with Čech filtrations 94
II.6.1 Consistency 94
II.6.2 Stability 96

II.1 Notations

We start by gathering the notations that will be used throughout this manuscript.

General notations.

- $n, d, k > 0$ are integers.
- $\mathbb{R}$ denotes the real numbers and $\mathbb{R}^+ = [0, +\infty)$ the nonnegative real numbers,
- If $x, y \in \mathbb{R}$, $x \wedge y$ is the minimum of x and y.

- I denotes a set, $\mathrm{card}(I)$ its cardinal if it is finite, and I^c its complement.
- If f is a map with values in $\mathbb{R}$ and $t \in \mathbb{R}$, f^t denotes the sublevel set $f^t = f^{-1}((-\infty, t])$.
- $\mathbb{R}^n$ and $\mathbb{R}^m$ denote the Euclidean spaces of dimension n and m, E denotes a Euclidean space.
- $\mathrm{M}(\mathbb{R}^m)$ is the vector space of $m \times m$ matrices, $\mathcal{G}_d(\mathbb{R}^m)$ the Grassmannian of d-subspaces of $\mathbb{R}^m$, and $\mathbb{S}_k \subset \mathbb{R}^{k+1}$ the unit k-sphere.
- For $x, y \in E$, $x \perp y$ denotes the orthogonality of x and y. If $X \subset E$ is a subset, $X^\perp$ denotes its orthogonal.
- If $x, y \in E$, $x \otimes y = x^t y \in \mathrm{M}(E)$ is the outer product, and $x^{\otimes 2} = x \otimes x$.
- If $X \subset E$ is any subset and $t \geq 0$, X^t denotes its t-thickening, and $X^t_<$ its open t-thickening (see Subsection II.3.2).
- $\|\cdot\|$ denotes the Euclidean norm on E and $\langle \cdot, \cdot \rangle$ the corresponding inner product, $\|\cdot\|_{\mathrm{F}}$ the Frobenius norm on $\mathrm{M}(E)$, $\|\cdot\|_\gamma$ the norm on $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ defined as $\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_{\mathrm{F}}^2$ where $\gamma > 0$ is a parameter (see Equation (IV.21) in Subsection IV.23 or Equation (V.1) in Subsection V.1.2).
- If T is a subspace of E, p_T denotes the orthogonal projection matrix on T.

Measure-theoretic notations.

- $\mathrm{W}_p(\cdot, \cdot)$ denotes the p-Wasserstein distance between measures on E (see Subsection II.5.1), and $\mathrm{W}_{p,\gamma}(\cdot, \cdot)$ the (p, γ)-Wasserstein distance between measures on $E \times \mathrm{M}(E)$ (see Equation (IV.22) in Subsection IV.3.1).
- $\mathcal{H}^d$ denotes the d-dimensional Hausdorff measure on E or on a subspace $T \subset E$.
- If μ is a measure of positive finite mass, $|\mu|$ denotes its mass, $\overline{\mu} = \frac{1}{|\mu|}\mu$ is the associated probability measure, and $\check{\mu}$ denotes the associated lifted measure (see Definition IV.23).
- If μ is a probability measure, $\mathrm{d}_{\mu,m}$, or simply d_μ, denote the corresponding DTM with parameter m (see Subsection II.5.2)

Geometric notations.

- $\mathcal{B}(x, r)$ and $\overline{\mathcal{B}}(x, r)$ denote the open and closed balls of E, and $\partial\mathcal{B}(x, r)$ the sphere. V_d and S_{d-1} denote the quantities $\mathcal{H}^d(\mathcal{B}(0, 1))$ and $\mathcal{H}^{d-1}(\partial\mathcal{B}(0, 1))$ (note that $S_{d-1} = dV_d$).
- If T is a subspace of E, $\mathcal{B}_T(x, r)$ and $\overline{\mathcal{B}}_T(x, r)$ denote the open and closed balls of T for the Euclidean distance.
- $\mathcal{M}_0$ denotes a Riemannian manifold, and $\mathcal{B}_{\mathcal{M}_0}(x, r)$ and $\overline{\mathcal{B}}_{\mathcal{M}_0}(x, r)$ denote the open and closed geodesic balls. For all $x_0, y_0 \in \mathcal{M}_0$, $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ denotes the geodesic distance between x_0 and y_0. The second fundamental form at $x_0 \in \mathcal{M}_0$ is denoted II_{x_0} (see Subsection II.2.5).

- If A is a subset of E, then $\mathrm{med}\,(A)$ denotes its medial axis and $\mathrm{reach}\,(A)$ its reach. The function distance to A is denoted $\mathrm{dist}\,(\cdot, A)$ or $\mathrm{d}_A\,(\cdot)$. The projection on A is denoted $\mathrm{proj}\,(\cdot, A)$ or $\mathrm{proj}_A\,(\cdot)$ (see Subsection II.3.3). The weak feature size of A is denoted $\mathrm{wfs}\,(A)$, and its μ-reach is denoted $\mathrm{reach}_\mu\,(A)$ (see Subsection II.3.4).
- $\mathrm{d_H}\,(\cdot, \cdot)$ denotes the Hausdorff pseudo-distance between two sets of E (see Subsection II.3.1).
- If K is a simplicial complex, $K^{(i)}$ denotes its i-skeleton. If $v \in K^{(0)}$ is a vertex, $\mathrm{St}(v)$ and $\overline{\mathrm{St}}(v)$ denote the open and closed star. The topological realization of K is denoted $|K|$, and the topological realization of a simplex $\sigma \in K$ is $|\sigma|$. The face map is denoted $\mathcal{F}_K \colon |K| \to K$ (see Subsection II.2.2).
- If $f \colon K \to L$ is a simplicial map, $|f| : |K| \to |L|$ denotes its topological realization. The i^{th} barycentric subdivision of the simplicial complex K is denoted $\mathrm{sub}^i(K)$ (see Subsection II.2.2).

Persistent homology notations.

- $\mathbb{X}$ denotes a set filtration, with $\mathbb{X} = (X^t)_{t\in T}$. $\mathbb{V}[\mathbb{X}]$ denotes the corresponding persistent homology module. If X is a subset of E, then $V[X] = (V^t[X])_{t\geq 0}$, or $\mathbb{X} = (X^t)_{t\in T}$, denote the Čech set filtration of X, and $\mathbb{V}[X] = (H_i(X^t))_{t\geq 0}$ the corresponding i^{th} homology persistence module (see Subsection II.4.2).
- $\mathcal{U}$ denotes a cover of a topological space, and $\mathcal{N}(\mathcal{U})$ its nerve (see Subsection II.4.1). $\mathbb{S} = (S^t)_{t\in T}$ denotes a simplicial filtration. If X is a subset of E, $\mathcal{V}[X] = (\mathcal{V}^t[X])_{t\geq 0}$ denotes its Čech simplicial filtration, and $\mathrm{Rips}(\mathcal{V}[X]) = (\mathrm{Rips}(\mathcal{V}^t[X]))_{t\geq 0}$ its Vietoris-Rips filtration (see Subsection II.4.2).
- $(\mathbb{X}, \mathbb{p})$ denotes a vector bundle filtration, with $\mathbb{X}$ a set filtration, and $\mathbb{p} = (p^t)_{t\in T}$ a family of maps $p^t \colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$ (see Definition V.2). If X is a subset of $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$, then $(V[X], \mathbb{p})$ or $(\mathbb{X}, \mathbb{p})$ denotes the Čech bundle filtration associated to X.
- If X is a topological space (resp. a simplicial complex), $H_i(X)$ denotes its i^{th} singular (resp. simplicial) homology group. If $f \colon X \to Y$ is a continuous map, $f_* \colon H_i(X) \to H_i(Y)$ is the map induced in homology (see Subsection II.4.1).
- If X is a topological space, $H^*(X)$ denotes its cohomology ring, and $H^i(X)$ the i^{th} cohomology group. The cup product of two elements $x, y \in H^*(X)$ is denoted $x \smile y$. If $f \colon X \to Y$ is a continuous map, $f^* \colon H^*(X) \leftarrow H^*(Y)$ is the map induced in cohomology (see Subsection II.4.1).
- $(\mathbb{V}, \mathbb{v})$ denotes a persistence module (resp. persistence comodule) over T, with $\mathbb{V} = (V^t)_{t\in T}$ a family of vector spaces, and $\mathbb{v} = (v_s^t \colon X^s \to X^t)_{s\leq t\in T}$ (resp. $\mathbb{v} = (v_s^t \colon X^s \leftarrow X^t)_{s\leq t\in T}$) a family of linear maps (see Subsections II.4.2 and II.4.5).
- If ξ is a vector bundle, $w_i(\xi)$ denotes its i^{th} Stiefel-Whitney class (see Subsection II.2.4). If $(\mathbb{X}, \mathbb{p})$ is a vector bundle filtration, $w_i(\mathbb{p})$ denotes the i^{th} persistent Stiefel-Whitney class, with $w_i(\mathbb{p}) = (w_i^t(\mathbb{p}))_{t\in T}$ (see Definition V.2).

II.2 Background on differential geometry

II.2.1 Basic notions of differential geometry

This presentation follows [Mun16]. We assume that the notions of topological spaces and differentiability of maps between Euclidean spaces are known. Let $d \geq 1$ be an integer, $k \geq 0$ an integer or $k = \infty$, and let $\mathbb{R}^d$ denotes the Euclidean space of dimension d. In this subsection, we define manifolds of constant dimension.

Manifolds. Let $\mathcal{M}$ be a topological space, that we suppose Hausdorff-separated and second-countable. A *chart* on $\mathcal{M}$ is a homeomorphism $\phi_\alpha \colon U_\alpha \to V_\alpha$ where U_α and V_α are respectively open sets of $\mathcal{M}$ and $\mathbb{R}^d$. Given two charts ϕ_α and ϕ_β, the corresponding *transition map* is the composition $\phi_\beta \circ \phi_\alpha^{-1}$, with domain $\phi_\alpha(U_\alpha \cap U_\beta)$ and codomain $\phi_\beta(U_\alpha \cap U_\beta)$, which are both subsets of $\mathbb{R}^d$. A *$\mathcal{C}^k$-atlas* on $\mathcal{M}$ is a collection of charts $\{\phi_\alpha \colon U_\alpha \to V_\alpha, \alpha \in A\}$ such that

1. $\bigcup_{\alpha \in A} U_\alpha = \mathcal{M}$,
2. each transition map $\phi_\beta \circ \phi_\alpha^{-1}$ is of class $\mathcal{C}^k$.

A $\mathcal{C}^k$-atlas is a *maximal atlas* if it is maximal for the inclusion between $\mathcal{C}^k$-atlases. Endowed with a $\mathcal{C}^k$-maximal atlas, $\mathcal{M}$ is called a $\mathcal{C}^k$-manifold. We also say that $\mathcal{M}$ is endowed with a $\mathcal{C}^k$-structure.

The data of a $\mathcal{C}^k$-maximal atlas on $\mathcal{M}$ is equivalent to the data of an equivalence class of $\mathcal{C}^k$-atlases under the following equivalence relation: two $\mathcal{C}^k$-atlases are *$\mathcal{C}^k$-compatible* if their union is still a $\mathcal{C}^k$-atlas.

Two $\mathcal{C}^k$-structures $\{\phi_\alpha, \alpha \in A\}$ and $\{\psi_\beta, \beta \in B\}$ are said *essentially equivalent*, or *$\mathcal{C}^k$-diffeomorphic*, if there exists a continuous map $f \colon \mathcal{M} \to \mathcal{M}$ such that the compositions $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ are all $\mathcal{C}^k$. An example of essentially equivalent but not compatible $\mathcal{C}^k$-structures, with $k \geq 1$, is given by $\mathcal{M} = \mathbb{R}$ and the atlases $\{x \mapsto x\}$ and $\{x \mapsto x^3\}$.

When $k = 0$, $\mathcal{M}$ is called a *topological manifold*, and the condition 2 of the definition is superfluous. A topological space admits at most one maximal $\mathcal{C}^0$-atlas, which is given by the collection of all continuous maps $\phi_\alpha \colon U_\alpha \to V_\alpha$. Therefore, all atlases on $\mathcal{M}$ are $\mathcal{C}^0$-compatible, hence $\mathcal{C}^0$-essentially equivalent.

When $k = 1$, $\mathcal{M}$ is called a *differentiable manifold*. In dimension $d \leq 3$, every topological manifold admits essentially a unique $\mathcal{C}^1$-structure. In dimension $d \geq 5$, every compact topological manifold admits essentially at most a finite number of $\mathcal{C}^1$-structures. The topological manifold $\mathbb{R}^4$ admits uncountably many $\mathcal{C}^1$-structures that are not essentially equivalent. There exists a topological manifold that admits no $\mathcal{C}^1$-structure.

When $k = \infty$, $\mathcal{M}$ is called a *smooth manifold*. Each $\mathcal{C}^l$-manifold, with $l \geq 1$, admits a $\mathcal{C}^\infty$-structure, which is essentially unique.

Immersions and embeddings. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{C}^k$-manifolds, and $f \colon \mathcal{M} \to \mathcal{N}$ a continuous map. If $\phi_\alpha \colon U_\alpha \to V_\alpha$ and $\psi_\beta \colon U_\beta \to V_\beta$ are charts of $\mathcal{M}$ and $\mathcal{N}$, the *expression* of f in these charts is the composition $\psi_\beta \circ f \circ \phi_\alpha^{-1}$, with domain $\phi_\alpha(U_\alpha \cap f^{-1}(U_\beta))$ and codomain $\psi_\beta(f(U_\alpha) \cap U_\beta)$. These are subsets of $\mathbb{R}^d$. The map f is said *differentiable* if its expression is differentiable in any charts. More

generally, f is said of *class* $\mathcal{C}^k$ if its expression is of class $\mathcal{C}^k$ in any charts. If x is a point of $\mathcal{M}$ and $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ an expression of f such that x belongs to its domain, then the rank of the differential $d_{\phi_\alpha^{-1}(x)}(\psi_\beta \circ f \circ \phi_\alpha^{-1})$ does not depend on the charts, and is called the *rank* of f at x. However, the differential $d_{\phi_\alpha^{-1}(x)}(\psi_\beta \circ f \circ \phi_\alpha^{-1})$, seen as a linear endomorphism of $\mathbb{R}^n$, does depend on the charts.

We can define a canonical differential $d_x f$ as follows. Define the *tangent space* of $\mathcal{M}$ at x as the quotient set of differentiable curves $\gamma\colon (-1,1) \to \mathcal{M}$, with $\gamma(0) = x$, under the following relation: γ_1 is equivalent to γ_2 if for any chart $\phi\colon U \to \mathbb{R}^d$ of $\mathcal{M}$ with $x \in U$ we have $(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$. The tangent space $T_x\mathcal{M}$ can be canonically given a d-dimensional vector space structure. Now, define the differential $d_x f\colon T_x\mathcal{M} \to T_{f(x)}\mathcal{N}$ as follows: for every $u \in T_x\mathcal{M}$ and $\gamma\colon (-1,1) \to \mathcal{M}$ that defines u, let $d_x f(u)$ be the equivalence class of $f \circ \gamma$. The map $d_x f\colon T_x\mathcal{M} \to T_{f(x)}\mathcal{N}$ is linear.

A differentiable map is called an *immersion* if $d_x f$ is injective for every $x \in \mathcal{M}$. We then say that $f(\mathcal{M})$ is an *immersed manifold* of $\mathcal{N}$. The map f is called an *embedding* if, additionaly, f is an homeomorphism onto its image $f(\mathcal{M})$, where $f(\mathcal{M}) \subset \mathcal{N}$ is endowed with the subspace topology. We then say that $f(\mathcal{M})$ is an *embedded manifold* of $\mathcal{N}$. As a consequence of the inverse function theorem, an immersion is a local embedding, that is, for every $x \in \mathcal{M}$, there exists a neighborhood U such that the restriction $f_{|U}$ is an embedding. When $\mathcal{M}$ is compact, an injective immersion is an embedding.

An embedded manifold $f(\mathcal{M}) \subset \mathcal{N}$, with f of class $\mathcal{C}^k$, can be given a natural $\mathcal{C}^k$-structure, via the atlas $\{\phi_\alpha \circ f, \alpha \in A\}$, where $\{\phi_\alpha, \alpha \in A\}$ is an atlas for $\mathcal{N}$.

By Whitney's embedding theorem, every d-dimensional $\mathcal{C}^k$-manifold, with $k \geq 1$, can be embedded in the Euclidean space $\mathbb{R}^{2d}$ via a $\mathcal{C}^k$-embedding.

Submanifolds. A *$\mathcal{C}^k$-submanifold* of dimension d of the Euclidean space is a subset S of $\mathbb{R}^n$ such that for every point $x \in S$ there exists a neighborhood $U \subset \mathbb{R}^n$ of x, an open set $V \subset \mathbb{R}^d$ and an immersion $\phi : V \to U$ of class C^k such that ϕ is a homeomorphism onto $S \cap U$.

Seen as a topological space for the induced topology, the maps ϕ give S a $\mathcal{C}^k$-structure, and the inclusion $S \to \mathbb{R}^n$ is a $\mathcal{C}^k$-embedding. Conversely, the image of a $\mathcal{C}^k$-embedding of a manifold is a $\mathcal{C}^k$-submanifold. Hence the notions of $\mathcal{C}^k$-embedded manifolds and $\mathcal{C}^k$-submanifolds are equivalent.

The tangent spaces of a submanifold S can be defined as follows: for every $x \in S$,

$$T_x S = \{0\} \cup \left\{ v \in \mathbb{R}^n, \forall \epsilon > 0, \exists y \in S \text{ s.t. } y \neq x, \|y - x\| < \epsilon, \left\| \frac{v}{\|v\|} - \frac{y - x}{\|y - x\|} \right\| < \epsilon \right\}.$$

We also define its normal spaces: for every $x \in S$, the *normal space* of S at x is the subspace $N_x S = (T_x S)^\perp$.

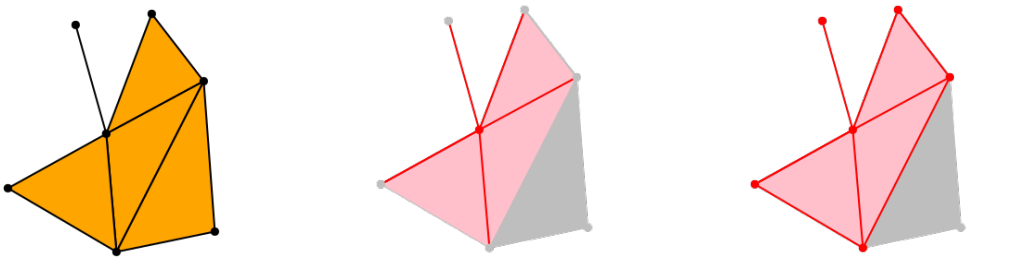
II.2.2 Basic notions of simplicial topology

We start by defining the simplicial complexes and their topology. We then describe the technique of simplicial approximation, based on [Hat02, Section 2.C].

(Combinatorial) simplicial complexes. A *simplicial complex* is a set K such that there exists a set V, the set of *vertices*, with $K \subseteq \mathcal{P}(V) \setminus \{\emptyset\}$, and such that K satisfies the following condition: for every $\sigma \in K$ and every subset $\nu \subseteq \sigma$, ν is in K. The elements of K are called *faces* or *simplices* of the simplicial complex K.

For every simplex $\sigma \in K$, we define its dimension $\dim(\sigma) = \mathrm{card}(\sigma) - 1$. The *dimension* of K, denoted $\dim(K)$, is the maximal dimension of its simplices. For every $i \geq 0$, the *i-skeleton* $K^{(i)}$ is defined as the subset of K consisting of simplices of dimension at most i. Note that $K^{(0)}$ corresponds to the underlying vertex set V, and that $K^{(1)}$ is a graph.

Given a simplex $\sigma \in K$, its *(open) star* $\mathrm{St}(\sigma)$ is the set of all the simplices $\nu \in K$ that contain σ. The open star is not a simplicial complex in general. We also define its *closed star* $\overline{\mathrm{St}}(\sigma)$ as the smallest simplicial subcomplex of K which contains $\mathrm{St}(\sigma)$.



K $\quad$ $\mathrm{St}(v)$ in red and pink $\quad$ $\overline{\mathrm{St}}(v)$ in red and pink

Figure II.1: Open and closed star of a vertex of K.

Given a graph G, the corresponding *clique complex* is the simplicial complex whose simplices are the sets of vertices of the cliques of G. We say that a simplicial complex K is a *flag complex* if it is the clique complex of its 1-skeleton $K^{(1)}$.

Topological realizations. For every $p \geq 0$, the *standard p-simplex* Δ^p is the topological space defined as the convex hull of the canonical basis vectors $e_1, ..., e_{p+1}$ of $\mathbb{R}^{p+1}$, endowed with the subspace topology. We now describe the construction of the *topological realization* of the simplicial complex K, denoted $|K|$. It is a particular case of the construction CW-complexes, as in [Hat02, Appendix].

1. Start with the discrete topological space $\left|K^{(0)}\right|$ consisting of the vertices of K.
2. Inductively, form the p-skeleton $\left|K^{(p)}\right|$ from $\left|K^{(p-1)}\right|$ by attaching p-dimensional simplices to $\left|K^{(p-1)}\right|$. More precisely, for each $\sigma \in K$ of dimension p, take a copy of the standard p-simplex Δ^p. Denote this simplex by $\Delta\sigma$. Label its vertices with the elements of σ. Whenever $\tau \subset \sigma \in K$, identify $\Delta\tau$ with a subset of $\Delta\sigma$, via the face inclusion which sends the elements of τ to the corresponding elements of σ. Give $\left|K^{(p)}\right|$ the quotient topology.
3. Endow $|K| = \bigcup_{p\geq 0} \left|K^{(p)}\right|$ with the weak topology: a set $A \subset |K|$ is open if and only if $A \cap \left|K^{(p)}\right|$ is open in $\left|K^{(p)}\right|$ for each $p \geq 0$.

Alternatively, the topology on $|K|$ can be described as follows: a subset $A \subset |K|$ is open (or closed) if and only if for every $\sigma \in K$, the set $A \cap \Delta\sigma$ is open (or closed) in $\Delta\sigma$. Note that condition 3 is superfluous when K is finite dimensional.

If $\sigma = [v]$ is a vertex of K, we shall denote by $|\sigma|$ the singleton $\{v\}$, seen as a subset of $|K|$. If σ is a face of K of dimension at least 1, we shall denote by $|\sigma|$ the open subset of $|K|$ which corresponds to the interior of the face $\Delta\sigma \subset |K|$. We denote by $\overline{|\sigma|}$ its closure in $|K|$. Observe that if $\overline{\sigma}$ denotes the smallest simplicial subcomplex of K that contains σ, then $|\overline{\sigma}| = \Delta\sigma = \overline{|\sigma|}$. The following set is a partition of $|K|$:

$$\{|\sigma|\,, \sigma \in K\}\,.$$

This allows to define the *face map* of K. It is the unique map $\mathcal{F}_K \colon |K| \to K$ that satisfies $x \in |\mathcal{F}_K(x)|$ for every $x \in |K|$.

If L is a subset of K, we define its topological realization as $|L| = \bigcup_{\sigma \in L} |\sigma|$. For every simplex $\sigma \in K$, the topological realization of its open star, $|\mathrm{St}(\sigma)|$, is open in $|K|$. Besides, the topological realization of its closed star, $\left|\overline{\mathrm{St}}(\sigma)\right|$, is equal to $\overline{|\mathrm{St}(\sigma)|}$, hence is closed.

If σ is a face of K of dimension at least 1, the subset $|\sigma|$ of $|K|$ is canonically homeomorphic to the interior of the standard p-simplex Δ^p, where $p = \dim(\sigma)$. This allows to define on $|K|$ the barycentric coordinates: for every face $\sigma = [v_0, ..., v_p] \in K$, the points $x \in |\sigma|$ can be written as

$$x = \sum_{i=0}^{p} \lambda_i v_i$$

with $\lambda_0, ..., \lambda_p > 0$ and $\sum_{i=0}^{p} \lambda_i = 1$.

Triangulations. Let X be a topological space. A *triangulation* of X consists of a simplicial complex K together with a homeomorphism $h \colon X \to |K|$.

Let us point out that, when $X = \mathcal{M}$ is a topological manifold, a finer notion of triangulation exists, that we won't consider in this document. Namely, a *piecewise linear structure* on $\mathcal{M}$ is an atlas consisting of charts such that their transition maps are piecewise linear maps. Every smooth manifold admits an essentially unique compatible piecewise linear structure. This does not hold for topological manifolds.

Simplicial maps. A *simplicial map* between simplicial complexes K and L is a map between topological realizations $g \colon |K| \to |L|$ which sends vertices on vertices and is linear on every simplices. In other words, for every $\sigma = [v_0, ..., v_p] \in K$, the map g restricted to $|\sigma| \subset |K|$ can be written in barycentric coordinates as

$$\sum_{i=0}^{p} \lambda_i v_i \;\longmapsto\; \sum_{i=0}^{p} \lambda_i g(v_i). \tag{II.1}$$

A simplicial map $g \colon |K| \to |L|$ is uniquely determined by its restriction to the vertex sets $g_{|K^{(0)}} \colon K^{(0)} \to L^{(0)}$. Reciprocally, let $f \colon K^{(0)} \to L^{(0)}$ be a map between vertex sets which satisfies the following condition:

$$\forall \sigma \in K, f(\sigma) \in L. \tag{II.2}$$

Then f induces a simplicial map via barycentric coordinates, denoted $|f| \colon |K| \to |L|$. In the rest of this dissertation, a simplicial map shall either refer to a map $g \colon |K| \to |L|$ which satisfies Equation (II.1), to a map $f \colon K^{(0)} \to L^{(0)}$ which satisfies Equation (II.2), or to the induced map $f \colon K \to L$.

Simplicial approximation. Let $g\colon |K| \to |L|$ be any continuous map. The problem of *simplicial approximation* consists in finding a simplicial map $f\colon K \to L$ with topological realization $|f| : |K| \to |L|$ homotopy equivalent to g (see Subsection II.3.1 for a definition of homotopy equivalence). A way to solve this problem is to consider the following property: we say that the map g satisfies the *star condition* if for every vertex v of K, there exists a vertex w of L such that

$$g\left(\left|\overline{\mathrm{St}}(v)\right|\right) \subseteq |\mathrm{St}(w)|\,.$$

If this is the case, let $f\colon K^{(0)} \to L^{(0)}$ be any map between vertex sets such that for every vertex v of K, we have $g\left(\left|\overline{\mathrm{St}}(v)\right|\right) \subseteq |\mathrm{St}(f(v))|$. Equivalently, f satisfies

$$g\left(\overline{\mathrm{St}}(v)\right) \subseteq \mathrm{St}(f(v)).$$

Such a map is called a *simplicial approximation to* g. One shows that it is a simplicial map, and that its topological realization $|f|$ is homotopic to g [Hat02, Theorem 2C.1].

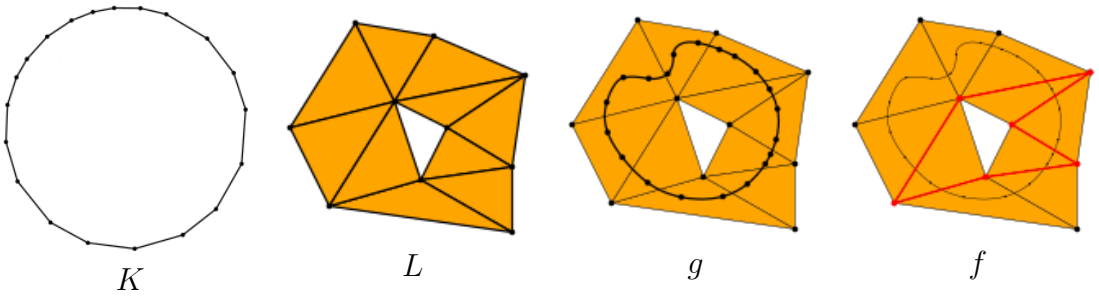


Figure II.2: The map $f\colon K \to L$ (in red) is a simplicial approximation to g.

In general, a map g may not satisfy the star condition. However, there is always a way to subdivise the simplicial complex K in order to obtain an induced map which does. We describe this construction in the following paragraph.

Barycentric subdivisions. Let us describe briefly the process of barycentric subdivision of a simplicial complex. A more extensive description can be found in [Hat02, Proof of Proposition 2.21]. Let Δ^p denote the standard p-simplex, with vertices denoted $v_0, ..., v_p$. The *barycentric subdivision* of Δ^p consists in decomposing Δ^p into $(p+1)!$ simplices of dimension p. It is a simplicial complex, whose vertex set corresponds to the points $\sum_{i=0}^{p} \lambda_i v_i$ for which some λ_i are zero and the other ones are equal. Equivalently, one can see this new set of vertices as a the power set of the set of vertices of Δ^p.

More generally, if K is a simplicial complex, its barycentric subdivision $\mathrm{sub}(K)$ is the simplicial complex obtained by subdivising each of its faces. The set of vertices of $\mathrm{sub}(K)$ can be seen as a subset of the power set of the set of vertices of K.

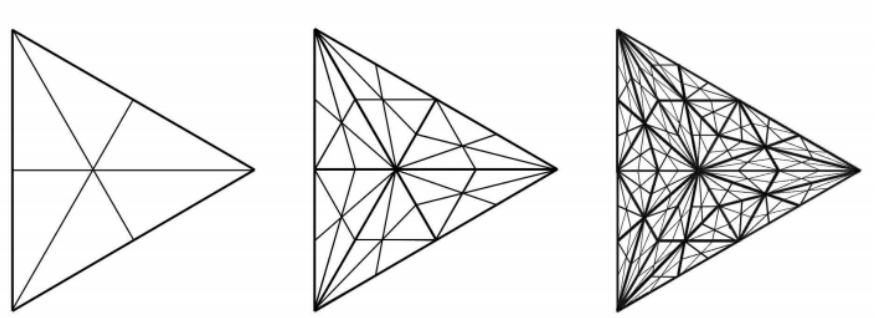


Figure II.3: The first three barycentric subdivisions of a 2-simplex.

If $g\colon |K| \to |L|$ is any map, there exists a canonical extended map $|\mathrm{sub}(K)| \to |L|$, still denoted g.

Observe that subdivising K shrinks its faces. More precisely, if Δ^p denotes the standard p-simplex, with D its diameter, then the faces of the barycentric subdivision of Δ^p are of diameter at most $\frac{p}{p+1}D$. Therefore one can repeat the subdivision to obtain arbitrarily small faces. Applying n times the barycentric subdivision procedure to K shall be denoted $\mathrm{sub}^n(K)$.

Theorem II.1 ([Hat02, Theorem 2C.1]). *Consider two simplicial complexes K, L with K finite, and let $g\colon |K| \to |L|$ be a continuous map. Then there exists $n \geq 0$ such that $g\colon |\mathrm{sub}^n(K)| \to |L|$ satisfies the star condition.*

As a consequence, such a map $g\colon |\mathrm{sub}^n(K)| \to |L|$ admits a simplicial approximation. This is known as the simplicial approximation theorem. As an illustration, Figure II.4 represents a map $g\colon |K| \to |L|$ which does not satisfies the star condition, but whose first barycentric subdivision does.

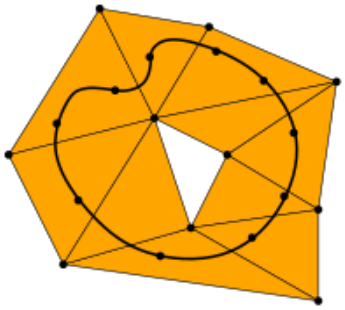


Figure II.4: The map $g\colon |K| \to |L|$ does not satisfy the star condition, but its first barycentric subdivision does (see Figure II.2).

II.2.3 Basic notions of vector bundle theory

This subsection follows the presentation of [MS16, Chapters 2 to 5]. Let X be a topological space and $d \geq 1$ an integer.

Vector bundles. A *vector bundle* ξ of dimension d over X consists of a topological space $A = A(\xi)$, the *total space*, a continuous map $\pi = \pi(\xi)\colon A \to X$, the *projection map*, and for every $x \in X$, a structure of d-dimensional vector space on the *fiber* $\pi^{-1}(\{x\})$. Moreover, ξ must satisfy the local triviality condition: for every $x \in X$,

there exists a neighborhood $U \subseteq X$ of x and a homeomorphism $h\colon U \times \mathbb{R}^d \to \pi^{-1}(U)$ such that for every $y \in U$, the map $z \mapsto h(y,z)$ defines an isomorphism between the vector spaces $\mathbb{R}^d$ and $\pi^{-1}(\{y\})$.

$$\begin{array}{c} A(\xi) \\ \downarrow{\scriptstyle \pi} \\ X \end{array} \qquad\qquad \begin{array}{ccc} \pi^{-1}(U) & \xleftarrow{\ h\ } & U \times \mathbb{R}^d \\ {\scriptstyle \pi}\downarrow & \swarrow_{p_1} & \\ U & & \end{array}$$

In this subsection, the fibers $\pi^{-1}(\{x\})$ are denoted $F_x(\xi)$.

Isomorphisms of vector bundles. An *isomorphism between vector bundles* ξ and η with common base space X is a homeomorphism $f\colon A(\xi) \to A(\eta)$ which sends each fiber $F_x(\xi)$ isomorphically into $F_{f(x)}(\eta)$. We obtain a commutative diagram

$$\begin{array}{ccc} A(\xi) & \xrightarrow{\quad f \quad} & A(\eta) \\ & {\scriptstyle \pi(\xi)}\searrow \quad \swarrow{\scriptstyle \pi(\eta)} & \\ & X & \end{array}$$

The *trivial bundle* of dimension d over X, denoted $\epsilon = \epsilon^d_X$, is defined with the total space $A(\epsilon) = X \times \mathbb{R}^d$, with the projection map π being the projection on the first coordinate, and where each fiber is endowed with the usual vector space structure of $\mathbb{R}^d$. A vector bundle ξ over X is said trivial if it is isomorphic to ϵ.

Operations on vector bundles. If ξ, η are two vector bundles on X, we define their *Whitney sum* $\xi \oplus \eta$ by

$$A(\xi \oplus \eta) = \{(x, a, b), x \in X, a \in F_x(\xi), b \in F_x(\eta)\},$$

where the projection map is given by the projection on the first coordinate, and where the vector space structures are the product structures. If η is a vector bundle on Y and $g\colon X \to Y$ a continuous map, the *pullback bundle* $g^*\xi$ is the vector bundle on X defined by the total space

$$A(g^*\xi) = \{(x, a), x \in X, a \in F_{g(x)}(\xi)\},$$

and where the projection map is given by the projection on the first coordinate.

Bundle maps. A *bundle map* between two vector bundles ξ and η with base spaces X and Y is a continuous map $f\colon A(\xi) \to A(\eta)$ which sends each fiber $F_x(\xi)$ isomorphically into another fiber $F_{x'}(\eta)$. If such a map exists, there exists a unique map $\overline{f}$ which makes the following diagram commute:

$$\begin{array}{ccc} A(\xi) & \xrightarrow{\quad f \quad} & A(\eta) \\ {\scriptstyle \pi(\xi)}\downarrow & & \downarrow{\scriptstyle \pi(\eta)} \\ X & \xrightarrow[\quad \overline{f} \quad]{} & Y \end{array}$$

In this case, ξ is isomorphic to the pullback bundle $\overline{f}^*\eta$ [MS16, Lemma 3.1]. We say that the map $\overline{f}$ *covers* f.

Universal bundles. Let $0 < d \leq m$. The Grassmann manifold $\mathcal{G}_d(\mathbb{R}^m)$ is a set which consists of all d-dimensional linear subspaces of $\mathbb{R}^m$. It can be given a smooth manifold structure. When $d = 1$, $\mathcal{G}_1(\mathbb{R}^m)$ corresponds to the real projective space $\mathbb{P}_n(\mathbb{R})$. On $\mathcal{G}_d(\mathbb{R}^m)$, there exists a canonical vector bundle of dimension d, denoted γ_d^m. It consists in the total space

$$A(\gamma_d^m) = \{(V, v), V \in \mathcal{G}_d(\mathbb{R}^m), v \in V\} \subset \mathcal{G}_d(\mathbb{R}^m) \times \mathbb{R}^m,$$

with the projection map on the first coordinate, and the linear structure inherited from $\mathbb{R}^m$.

> **Lemma II.2 ([MS16, Lemma 5.3]).** *Let ξ be vector bundle of dimension d over a compact space X. Then for m large enough, there exists a bundle map from ξ to γ_d^m.*

If such a bundle map $f\colon \xi \to \gamma_d^m$ exists, then ξ is isomorphic to the pullback $\overline{f}^* \gamma_d^m$, where $\overline{f}$ denotes the map that f covers.

In order to avoid mentioning m, it is convenient to consider the infinite Grassmannian. The infinite Grassmann manifold $\mathcal{G}_d(\mathbb{R}^\infty)$ is the set of all d-dimensional linear subspaces of $\mathbb{R}^\infty$, where $\mathbb{R}^\infty$ is the vector space of series with a finite number of nonzero terms. The infinite Grassmannian is topologized as the direct limit of the sequence $\mathcal{G}_d(\mathbb{R}^d) \subset \mathcal{G}_d(\mathbb{R}^{d+1}) \subset \mathcal{G}_d(\mathbb{R}^{d+2}) \subset \cdots$. Just as before, there exists on $\mathcal{G}_d(\mathbb{R}^\infty)$ a canonical bundle γ_d^∞. It is called a *universal bundle*, for the following reason:

> **Lemma II.3 ([MS16, Lemma 5.3]).** *if ξ is vector bundle of dimension d over a paracompact space X, then there exists a bundle map from $\xi \to \gamma_d^\infty$.*

If we denote such a bundle map $f_\xi\colon A(\xi) \to A(\gamma_d^\infty)$, then the underlying map $\overline{f}_\xi : X \to \mathcal{G}_d(\mathbb{R}^\infty)$ is called a *classifying map* for ξ. As before, ξ is isomorphic to the pullback $(\overline{f}_\xi)^* \gamma_d^\infty$.

If f is a bundle map given by Lemma II.2, then the following composition is a classifying map for ξ, as in Lemma II.3:

$$X \xrightarrow{\overline{f}} \mathcal{G}_d(\mathbb{R}^m) \hookrightarrow \mathcal{G}_d(\mathbb{R}^\infty).$$

A correspondence. Let ξ, η be bundles over X. If the classifying maps $\overline{f_\xi}$ and $\overline{f_\eta}$ are homotopic, one shows that the bundles ξ and η are isomorphic. The following theorem states that the converse is also true.

> **Theorem II.4 ([MS16, Corollary 5.10]).** *Let X be a paracompact space. There exists a bijection between the vector bundles over X (up to isomorphism) and the continuous maps $X \to \mathcal{G}_d(\mathbb{R}^\infty)$ (up to homotopy). It is given by $\xi \mapsto \overline{f}_\xi$, where $\overline{f}_\xi$ denotes a classifying map for ξ.*

This result leads to the following convention:

In the rest of this manuscript, we shall consider that vector bundles are given as a continuous maps $X \to \mathcal{G}_d(\mathbb{R}^m)$ or $X \to \mathcal{G}_d(\mathbb{R}^\infty)$.

II.2.4 Basic notions of Stiefel-Whitney classes

We still follow the presentation of [MS16]. The Stiefel-Whitney classes are a particular instance of the theory of characteristic classes, with coefficient group being $\mathbb{Z}_2$. We first define them axiomatically, and then describe their construction. Basic notions of cohomology are presented in Subsection II.4.1.

Axioms for Stiefel-Whitney classes. To each vector bundle ξ over a paracompact base space X, one associates a sequence of cohomology classes

$$w_i(\xi) \in H^i(X, \mathbb{Z}_2), \qquad i \in \mathbb{N},$$

called the *Stiefel-Whitney classes of* ξ. These classes satisfy:

- **Axiom 1:** w_0 is equal to $1 \in H^0(X, \mathbb{Z}_2)$, and if ξ is of dimension d, then $w_i(\xi) = 0$ for $i > d$.
- **Axiom 2:** if $f\colon \xi \to \eta$ is a bundle map, then $w_i(\xi) = \overline{f}^* w_i(\eta)$, where $\overline{f}^*$ is the map in cohomology induced by the underlying map $\overline{f}\colon X \to Y$ between base spaces.
- **Axiom 3:** if ξ, η are bundles over the same base space X, then for all $k \in \mathbb{N}$, $w_k(\xi \oplus \eta) = \sum_{i=0}^{k} w_i(\xi) \smile w_{k-i}(\eta)$, where $\smile$ denotes the cup product.
- **Axiom 4:** if γ_1^1 denotes the universal bundle of the projective line $\mathcal{G}_1(\mathbb{R}^2)$, then $w_1(\gamma_1^1) \neq 0$.

The Stiefel-Whitney classes are invariants of vector bundles, and carry topological information. For instance, the following lemma shows that the first Stiefel-Whitney class detects orientability.

Proposition II.5 ([MS16, Lemma 11.6 and Problem 12-A]). *Let X be a compact manifold and τ its tangent bundle. Then X is orientable if and only if $w_1(\tau) = 0$.*

Construction of the Stiefel-Whitney classes. The cohomology rings of the Grassmann manifolds admit a simple description: $H^*(\mathcal{G}_d(\mathbb{R}^\infty), \mathbb{Z}_2)$ is the free abelian ring generated by d elements $w_1, ..., w_d$. As a graded algebra, the degree of these elements are $|w_1| = 1, ..., |w_d| = d$ [MS16, Theorem 7.1]. Hence we can write

$$H^*(\mathcal{G}_d(\mathbb{R}^\infty), \mathbb{Z}_2) \simeq \mathbb{Z}_2[w_1, ..., w_d].$$

In particular, the infinite projective $\mathbb{P}_\infty = \mathcal{G}_1(\mathbb{R}^\infty)$ space has cohomology $H^*(\mathbb{P}_\infty, \mathbb{Z}_2) = \mathbb{Z}_2[w_1]$, the polynomial ring.

The generators $w_1, ..., w_d$ can be seen as the Stiefel-Whitney classes of the universal bundle γ_d^∞ on $\mathcal{G}_d(\mathbb{R}^\infty)$. Now, for any vector bundle ξ, define

$$w_i(\xi) = \overline{f}_\xi^*(w_i),$$

where $\overline{f}_\xi\colon X \to \mathcal{G}_d(\mathbb{R}^\infty)$ is a classifying map for ξ (as in Theorem II.4), and $\overline{f}_\xi^*\colon H^*(X) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^\infty))$ the induced map in cohomology. This construction yields the Stiefel-Whitney classes.

Theorem II.6 ([MS16, Theorem 7.3]). *Defined this way, the classes satisfy the four axioms. And they are unique.*

II.2.5 Basic notions of Riemannian geometry

This subsection is based on [dC92]. We present Riemannian manifolds, their geodesics and second fundamental forms.

Riemannian manifold. Let $\mathcal{M}$ be a $\mathcal{C}^1$-manifold. A *Riemannian structure* on $\mathcal{M}$ is the data of an inner product $\langle \cdot, \cdot \rangle_x$ on each tangent space $T_x\mathcal{M}$, which satisfies the following condition: for every system of coordinates $\phi\colon U \subset \mathbb{R}^d \to \mathcal{M}$, for every $i, j \in [1, d]$, the map $x \mapsto \langle d_x\phi(e_i), d_x\phi(e_j)\rangle_{\phi(x)}$ is differentiable on U, where $(e_i)_{i\in[1,d]}$ denotes the canonical basis of $\mathbb{R}^d$. For every $x \in \mathcal{M}$ and $v \in T_x\mathcal{M}$, we denote $\|v\|_x = \langle v, v\rangle_x$.

Any differentiable manifold admits a Riemannian structure. If $\mathcal{M}$ and $\mathcal{N}$ are two Riemannian manifolds, an *isometry* is a $\mathcal{C}^1$-diffeomorphism $f\colon \mathcal{M} \to \mathcal{N}$ such that $\langle u, v\rangle_x = \langle d_x f(u), d_x f(v)\rangle_{f(x)}$ for every $x \in \mathcal{M}$ and $u, v \in T_x\mathcal{M}$.

Geodesics. Geodesics can be defined from a metric point of view, or from a differential-equation point of view. We start with the first one.

Let I denotes the segment $[a, b] \subset \mathbb{R}$, and let $\gamma\colon I \to \mathcal{M}$ be a differentiable curve. Denote by $\dot\gamma$ the derivative of γ. The *length* of c is defined as

$$\ell(\gamma) = \int_a^b \|\dot\gamma(t)\|_{\gamma(t)}\,\mathrm{d}t.$$

This quantity does not depend on the parametrization of γ. Indeed, if $\phi\colon [a', b'] \to [a, b]$ is a $\mathcal{C}^1$-diffeomorphism, then integrating by substitution yields

$$\begin{aligned}\ell(\gamma\circ\phi) = \int_{a'}^{b'} \left\|\dot\phi(t)\cdot\dot\gamma\circ\phi(t)\right\|_x \mathrm{d}t &= \int_{a'}^{b'} \|\dot\gamma\circ\phi(t)\|_x\,|\dot\phi(t)|\mathrm{d}t\\ &= \int_a^b \|\dot\gamma(t)\|_x\,\mathrm{d}t = \ell(\gamma).\end{aligned}$$

More generally, one defines the length of piecewise differentiable curves. Now, given two points $x, y \in \mathcal{M}$, the *geodesic distance* between x and y is

$$\mathrm{d}_{\mathcal{M}}(x, y) = \inf\left\{\ell(\gamma),\ \gamma\colon I \to \mathcal{M} \text{ piecewise differentiable}, \left|\begin{array}{l}\gamma(a) = x\\ \gamma(b) = y\end{array}\right.\right\}.$$

The map $\mathrm{d}_{\mathcal{M}}(\cdot,\cdot)$ is a distance on $\mathcal{M}$, hence $\mathcal{M}$ can be seen as a metric space. The topology induced by this distance coincides with the initial topology on $\mathcal{M}$. Now, a *unit speed length-minimizing curve* is a continuous curve $\gamma\colon I \to \mathcal{M}$ such that every $t_0 \in I$ admits a neighborhood $J \subset I$ such that for every $s, t \in J$, we have

$$\mathrm{d}_{\mathcal{M}}(\gamma(s), \gamma(t)) = |s - t|.$$

We remark that this is a local property, hence a length-minimizing curve is to be seen as a *locally* length-minimizing curve. We give a partial statement of Hopf-Rinow theorem, which states that $\mathcal{M}$ is path-connected by length-minimizing curves.

Theorem II.7 ([dC92, Theorem 7.2.8]). *Suppose that $\mathcal{M}$ is complete (as a topological space). Then for every $x, y \in \mathcal{M}$, there exists a length-minimizing curve γ joining x to y and such that $\ell(\gamma) = \mathrm{d}_{\mathcal{M}}(x, y)$.*

We now consider the differential-equation point of view. Let $\gamma\colon I \to \mathcal{M}$ be a differentiable curve. One defines the *covariant derivative along γ associated to the Levi-Civita connection on $\mathcal{M}$*, as in [dC92, Proposition 2.2.2]. It is an operator, denoted D_t, that acts on the vector fields along γ. Now, a *geodesic* is defined as a differentiable curve $\gamma\colon I \to \mathcal{M}$ such that

$$D_t\dot{\gamma} = 0$$

at each $t \in I$. The following computation shows that the norm of $\dot{\gamma}$ is constant:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \dot{\gamma}(t), \dot{\gamma}(t)\right\rangle = 2\left\langle D_t\dot{\gamma}, \dot{\gamma}\right\rangle = 0.$$

If this constant is 1, we say that γ is a *unit speed geodesic*, or *an arc-length parametrized geodesic*. A geodesic is a solution of a second-order differential equation on $\mathcal{M}$. By considering the curve $t \mapsto (\gamma(t), \dot{\gamma}(t))$, we can see geodesics as solutions of a first-order differential equation on $T\mathcal{M}$, the tangent bundle of $\mathcal{M}$. This differential equation defines the *geodesic flow* on $T\mathcal{M}$.

Proposition II.8 ([dC92, From proposition 3.2.7]). *For every $x \in \mathcal{M}$, there exists a neighborhood V of x in $T_x\mathcal{M}$ such that for every $v \in V$, there exists a unique geodesic $\gamma\colon (-2, 2) \to \mathcal{M}$ with $\gamma(0) = x$ and $\dot{\gamma}(0) = v$.*

Let U be the union of these V's for all $x \in \mathcal{M}$. We can define a map $\exp\colon U \to \mathcal{M}$ by

$$\exp\colon (x, v) \in U \subset T\mathcal{M} \longmapsto \gamma(1),$$

where γ is the geodesic given by the previous proposition. This map is called the *exponential map* of $\mathcal{M}$. It is differentiable. As a consequence of Hopf-Rinow theorem [dC92, Theorem 7.2.8], exp is defined on the whole tangent bundle $T\mathcal{M}$, provided that $\mathcal{M}$ is a complete topological space.

Sometimes, we will fix $x \in \mathcal{M}$, and see the exponential map as $\exp_x\colon U \subset T_x\mathcal{M} \to \mathcal{M}$. According to [dC92, Theorem 3.2.9], for every $x \in \mathcal{M}$, there exists $\epsilon > 0$ such that $\exp_x$ restricted to $\mathcal{B}_{T_x\mathcal{M}}(0, \epsilon)$ is a $\mathcal{C}^1$-diffeomorphism onto its image, where $\mathcal{B}_{T_x\mathcal{M}}(0, \epsilon)$ denotes the open ball of radius ϵ and center 0 of $T_x\mathcal{M}$. We will invoke more regularity results of this kind in Chapter IV.

We can now connect the two notions of geodesic we defined:

Proposition II.9 ([dC92, Proposition 3.3.6 and Corollary 3.3.9]). *The unit-speed length-minimizing curves are unit-speed geodesics, and conversely.*

Immersed manifolds. If $f\colon \mathcal{M} \to \mathbb{R}^n$ is an immersion of a $\mathcal{C}^1$-manifold, then $\mathcal{M}$ is naturally endowed with a Riemannian structure by pulling back the inner product of $\mathbb{R}^n$. This makes f an isometry.

On such an immersed manifold, one defines the *(generalized) Gauss map* $G\colon \mathcal{M} \to \mathcal{G}_d(\mathbb{R}^n)$, where $\mathcal{G}_d(\mathbb{R}^n)$ denotes the Grassmannian of d-planes of $\mathbb{R}^n$, as

$$G\colon x \mapsto d_x(f)(T_x\mathcal{M}).$$

Here $d_x(f)$ denotes the differential of f at x, hence $d_x(f)(T_x\mathcal{M})$ represents the tangent space of $T_x\mathcal{M}$ seen in $\mathbb{R}^n$. If f is a $\mathcal{C}^k$-embedding, wit $k \geq 1$, then G is of class $\mathcal{C}^{k-1}$.

Second fundamental form. We still consider that the Riemannian manifold $\mathcal{M}$ is immersed in $\mathbb{R}^n$. For every $x \in \mathcal{M}$, one defines the *second fundamental form* of $\mathcal{M}$ at x, as in [dC92, Subsection 6.2]. It is a symmetric bilinear form

$$\mathrm{II}_x\colon\ T_x\mathcal{M} \times T_x\mathcal{M} \longrightarrow N_x\mathcal{M},$$

where $N_x\mathcal{M}$ denotes the normal space of $\mathcal{M}$ at x. The second fundamental form is closely related to the curvature of $\mathcal{M}$, for instance via Gauss theorem [dC92, Theorem 6.2.5].

Let $x \in \mathcal{M}$, $v \in T_x\mathcal{M}$ a unit vector, and consider an unit-speed geodesic $\gamma\colon I \to \mathcal{M}$ such that $\gamma(0) = x$ and $\dot\gamma(0) = v$. The following relation can be found in [NSW08, Section 6] or [BLW19, Section 3]:

$$\mathrm{II}_x(v, v) = \ddot\gamma(0).$$

In particular, any bound on the operator norm $\|\mathrm{II}_x\|_{\mathrm{op}}$ of II_x implies a bound on $\|\ddot\gamma(0)\|$. We will use this relation in Chapter IV.

II.3 Background on Euclidean geometry of compact sets

II.3.1 Basic notions of topology

Homotopies and retractions. Let X, Y be two topological spaces, and $f, g\colon X \to Y$ two continuous maps. A *homotopy* between f and g is a continuous map $F\colon X \times [0,1] \to Y$ such that $F(\cdot, 0) = f$ and $F(\cdot, 1) = g$. The maps f and g are said *homotopic*. The spaces X and Y are *homotopy equivalent* if there exist continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f\colon X \to X$ is homotopic to the identity on X and $f \circ g\colon Y \to Y$ is homotopic to the identity on Y.

Let us define three notions of retraction. Let X be a topological space and A a subset. A *retraction* of X onto A is a map $r\colon X \to A$ such that r restricted to A is the identity. A *deformation retraction* of X onto A is a homotopy $F\colon X \times [0,1] \to X$ between the identity on A and a retraction of X onto A. A *strong deformation retraction* of X onto A is a deformation retraction $F\colon X \times [0,1] \to X$ such that for every $t \in [0,1]$, $F(\cdot, t)$ is the identity on A.

Some retractions do not come from a deformation retraction, and some deformation retractions are not strong deformation retractions. If there exists a deformation retraction from X onto A, then X and A are homotopy equivalent.

Hausdorff distance. Let X be any subset of $\mathbb{R}^n$ endowed with a norm $\|\cdot\|$. The function *distance to* X is the map $\mathrm{dist}\,(\cdot, X) : y \in E \mapsto \inf\{\|y - x\|, x \in X\}$. A projection of $y \in \mathbb{R}^n$ on X is a point $x \in X$ which attains this infimum. Such a point exists when X is compact. If Y is another subset of $\mathbb{R}^n$, we define the *non-symmetric Hausdorff distance* by

$$\mathrm{d_H}\,(X;Y) = \sup\{\mathrm{dist}\,(y, X), y \in Y\},$$

and their *Hausdorff distance* $\mathrm{d_H}\,(X, Y)$ by

$$\mathrm{d_H}\,(X, Y) = \max\{\mathrm{d_H}\,(X;Y), \mathrm{d_H}\,(Y;X)\}.$$

If X and Y are compact, their Hausdorff distance is finite. The application $\mathrm{d_H}\,(\cdot,\cdot)$ is a distance between compact subsets of $\mathbb{R}^n$.

The Hausdorff distance is connected to the sup norm of functions. For any subset $X \subset \mathbb{R}^n$, denote by d_X the function distance to X. Now, if Y is any other subset of $\mathbb{R}^n$, we see directly that

$$\mathrm{d_H}\,(X, Y) = \|\mathrm{d}_X - \mathrm{d}_Y\|_\infty.$$

II.3.2 Thickenings and tubular neighborhoods

Let X be a subset of the usual Euclidean space $(\mathbb{R}^n, \|\cdot\|)$, that we suppose close to some embedded $\mathcal{C}^k$-manifold $\mathcal{M} \subset \mathbb{R}^n$ of dimension d, with $k \geq 0$. In this subsection, we consider that their closeness is quantified via their Hausdorff distance $\mathrm{d_H}\,(X, \mathcal{M})$.

Thickenings. The primary construction of persistent homology is to consider the thickenings of X. For every $t \geq 0$, the *t-thickening* of X is the subset

$$X^t = \{x \in \mathbb{R}^n, \mathrm{dist}\,(x, X) \leq t\}.$$

When X is closed, the thickenings can be described as

$$X^t = \{y \in \mathbb{R}^n, \exists x \in X, \|x - y\| \leq t\}.$$

In this case, the 0-thickening X^0 is equal to X itself.

Let $\mathrm{d_H}\,(X, \mathcal{M})$ be the Hausdorff distance between X and $\mathcal{M}$. The thickening construction inherits the initial Hausdorff distance between the sets. Namely, for every $t \geq 0$, the Hausdorff distance between the thickenings is bounded by

$$\mathrm{d_H}\,(X^t, \mathcal{M}^t) \leq \mathrm{d_H}\,(X, \mathcal{M})\,.$$

As a consequence, the thickenings are close to each other as long as the initial sets are. Note that the reverse inequality $\mathrm{d_H}\,(X^t, \mathcal{M}^t) \geq \mathrm{d_H}\,(X, \mathcal{M})$ does not hold in general.

We propose to look at the thickenings with two different points of view. From an algorithmic perspective, the thickenings X^t of a finite set X is a finite union of balls:

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}\,(x, t)\,.$$

As we will see in Subsection II.4.1, the nerve theorem implies that such a subset has computable topology. That is, we can build naturally over X a simplicial complex that has the homotopy type of X^t. On the other hand, from a differentiable viewpoint, the thickenings $\mathcal{M}^t$ of a submanifold $\mathcal{M}$ can be seen as a tubular neighborhoods. We adopt this point of view in the following paragraph.

Tubular neighborhoods. We start in the $\mathcal{C}^\infty$ setting, as in [Hir12, Chapter 4] and [Spi70, Chapter 9]. Let $\xi = (\pi, A, \mathcal{M})$ be a vector bundle over $\mathcal{M}$, where p is the projection map, and A the total space. We say that this vector bundle defines a $\mathcal{C}^\infty$*-tubular neighborhood* if it is endowed with an $\mathcal{C}^\infty$-embedding $f\colon A \to \mathbb{R}^n$ such that

- $f_{|\mathcal{M}} = \mathrm{id}$, where $\mathcal{M}$ is identified with the 0-section of A,
- $f(A)$ is an open neighborhood of $\mathcal{M}$ in $\mathbb{R}^n$.

The image of the tubular neighborhood, $f(A)$, is a subset of $\mathbb{R}^n$ that deform retracts onto $\mathcal{M}$. For instance, a natural retraction $r\colon f(A) \to \mathcal{M}$ is given by

$$r\colon v \in F_x(\xi) \mapsto x,$$

where $F_x(\xi)$ denotes the fiber over $x \in \mathcal{M}$. The retraction r is smoothly homotopic to the identity via $H\colon f(A) \times [0,1] \to f(A)$ defined as

$$H\colon (v,t) \in F_x(\xi) \times [0,1] \mapsto f(x, tv).$$

Theorem II.10 (Tubular neighborhood theorem). *Any smooth embedded manifold admits a $\mathcal{C}^\infty$-tubular neighborhood.*

Proof, as in [Hir12, Theorem 5.1]. Let $u\colon \mathcal{M} \to \mathcal{G}_{n-d}(\mathbb{R}^n)$ be a (smooth) field of transverse $(n-d)$-planes to $\mathcal{M}$, that is, for every $x \in \mathcal{M}$, we have $u(x) \oplus T_x\mathcal{M} = \mathbb{R}^n$. For instance, one can choose the normal space $u(x) = N_x\mathcal{M}$ for all $x \in \mathcal{M}$. Define the vector bundle $\xi = (\pi, A, \mathcal{M})$ with $A \subset \mathbb{R}^n \times \mathbb{R}^n$ defined as

$$A = \{(x,v) \in \mathcal{M} \times \mathbb{R}^n, v \in u(x)\},$$

and $\pi\colon E \to \mathcal{M}$ the projection on the first coordinate. Define the map $f : A \to \mathbb{R}^n$ by

$$f(x,v) = x + v.$$

By compacity of $\mathcal{M}$, there exists a neighborhood U of $\mathcal{M}$ in A on which f is an embedding. From this neighborhood U, one builds a tubular neighborhood.

Normal tubular neighborhood. The construction of Theorem II.10 shows that a particular choice for $u(x)$ can be the normal space $N_x\mathcal{M}$. This yields to a *normal tubular neighborhood*, which is the viewpoint of [Lee13, Chapter 6]. It can be shown that U can be chosen small enough so that, for every $x \in X$, the image of the fiber

$f(U \cap F_x(\xi))$ is the set of points whose nearest point of $\mathcal{M}$ is x. This case will be of particular interest for us, since it connects to the notion of thickenings.

For every $t > 0$, define the *open t-thickening* of $\mathcal{M}$ as

$$\mathcal{M}^t_< = \{x \in \mathbb{R}^n, \operatorname{dist}(x, \mathcal{M}) < t\}.$$

Alternatively, we can write

$$\mathcal{M}^t_< = \bigcup_{x \in \mathcal{M}} \mathcal{B}_{N_x\mathcal{M}}(x, t),$$

where $\mathcal{B}_{N_x\mathcal{M}}(x, t)$ is the ball of radius t in the normal space $N_x\mathcal{M} \subset \mathbb{R}^n$. Hence the open thickening $\mathcal{M}^t_<$ can be seen as the image of the tubular neighborhood built in the proof of Theorem II.10. We then obtain the following theorem as a consequence of the tubular neighborhood theorem.

Theorem II.11 (ϵ-normal tubular neighborhood theorem). *If $\mathcal{M}$ is a compact embedded $\mathcal{C}^\infty$-manifold, there exists $\epsilon > 0$ such that $\mathcal{M}^\epsilon_<$ is the image of a normal tubular neighborhood of $\mathcal{M}$.*

This theorem is not constructive, and does not indicate how small the ϵ may be. In our context, we would like to have a quantitative control of ϵ. To do so, we introduce in Subsection II.3.3 the reach of $\mathcal{M}$. It is exactly the supremum of the ϵ's such that Theorem II.11 holds.

Non-smooth manifolds. We now assume that $\mathcal{M}$ is only $\mathcal{C}^1$. Consider the construction of Theorem II.10: $u\colon \mathcal{M} \to \mathcal{G}_{n-d}(\mathbb{R}^n)$ is the field of normal spaces to $\mathcal{M}$, $A = \{(x, v) \in \mathcal{M} \times \mathbb{R}^n, v \in u(x)\} \subset \mathbb{R}^n \times \mathbb{R}^n$ is the corresponding total space, and f is the map

$$\begin{aligned} f\colon A &\to \mathbb{R}^n \\ (x, v) &\mapsto x + v. \end{aligned}$$

Let A_ϵ denote the subset of A defined as $\{(x, v) \in \mathcal{M} \times N_x\mathcal{M}, \|v\| < \epsilon\}$. We look for an ϵ such that $f\colon A_\epsilon \to \mathbb{R}^n$ is an embedding. One shows that such an ϵ exists if the map u satisfies the local Lipschitz property, with respect to the geodesic distance on $\mathcal{M}$ [Hir12, Exercise 6]. Note that the continuity of u is not a sufficient condition, as shown by the $\mathcal{C}^1$-embedding $x \mapsto x^{3/2}$.

This regularity leads to the following definition: an embedded $\mathcal{C}^1$-manifold is said of *class* $\mathcal{C}^{1,1}$ if the map $x \in \mathcal{M} \mapsto T_x\mathcal{M} \in \mathcal{G}_d(\mathbb{R}^n)$ satisfies the local Lipschitz property. We deduce that Theorem II.11 actually holds for $\mathcal{C}^{1,1}$ manifolds. Conversely, if $\mathcal{M}$ is a $\mathcal{C}^0$ manifold such that Theorem II.11 holds, then $\mathcal{M}$ is of class $\mathcal{C}^{1,1}$ [Lyt05, Proposition 1.4].

Non-manifolds. Suppose now that $\mathcal{M}$ is any subset of $\mathbb{R}^n$. The previous considerations raise the following question: for which $t > 0$ does there exist a map $f : \mathcal{M}^t \to \mathcal{M}$ which is a deform retract?

According to the last paragraph, such a t exists when $\mathcal{M}$ is of class $\mathcal{C}^{1,1}$, and the retraction can be chosen as the projection on $\mathcal{M}$. In the rest of this section, we answer to this question more generally:

- For all sets of positive reach, one can retract $\mathcal{M}^t$ on $\mathcal{M}$ via the projection.
- More generally, if $\mathcal{M}$ is a subset with positive μ-reach, $\mathcal{M}^t$ retract on $\mathcal{M}$ via a deformation retraction that may not come from the projection.

In both cases, $\mathcal{M}$ does not have to be a submanifold.

II.3.3 Reach

We present the definition of [Fed59].

Reach. Let X be any subset of $\mathbb{R}^n$. The *medial axis* of X is the subset $\mathrm{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\mathrm{med}(X) = \{y \in \mathbb{R}^n, \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}(y, X)\}.$$

The *reach* of X is

$$\mathrm{reach}(X) = \inf\{\|x - y\|, x \in X, y \in \mathrm{med}(X)\}.$$

Equivalently, the reach of X can be defined as the supremum of $t \geq 0$ such that the thickening X^t does not intersect $\mathrm{med}(X)$.

Suppose that X is closed and that $\mathrm{reach}(X)$ is positive. Then for every $t \in [0, \mathrm{reach}(X))$, the thickening X^t deform retracts onto X. A homotopy is given by a linear deformation on each fiber:

$$\begin{aligned} X^t \times [0,1] &\longrightarrow X^t \\ (x,t) &\longmapsto (1-t)x + t \cdot \mathrm{proj}(x, X). \end{aligned}$$

Regularity imposed by the reach. A useful property of sets with positive reach is the approximation by tangent spaces. For a general set X, we define the tangent cone at $x \in X$ as:

$$\mathrm{Tan}(X,x) = \{0\} \cup \left\{ v \in \mathbb{R}^n, \forall \epsilon > 0, \exists y \in X \text{ s.t. } y \neq x, \|y - x\| < \epsilon, \left\| \frac{v}{\|v\|} - \frac{y-x}{\|y-x\|} \right\| < \epsilon \right\}.$$

Note that if X is a submanifold, we recover the usual notion of tangent space.

The following characterization is fundamental in the study of sets with positive reach.

Theorem II.12 ([Fed59, Theorem 4.18(2)]). *A closed set $X \subset \mathbb{R}^n$ has positive reach τ if and only if for every $x, y \in X$,*

$$\mathrm{dist}(y - x, \mathrm{Tan}(X,x)) \leq \frac{1}{2\tau} \|y - x\|^2.$$

The reach is a quantity that controls both the local and global regularity of the set X. When $X = \mathcal{M}$ is a topological manifold, having a positive reach implies that $\mathcal{M}$ is of regularity $\mathcal{C}^{1,1}$ [Lyt05, Proposition 1.4]. Moreover, it can be shown that $\mathrm{reach}(\mathcal{M})$ is caused either by a bottleneck structure or by high curvature:

Theorem II.13 ([AKC+19, Theorem 3.4]). *A closed submanifold $\mathcal{M}$ with positive reach must satisfy at least one of the following two properties:*

- Global case: *there exist $x, y \in \mathcal{M}$ with $\|x - y\| = 2\mathrm{reach}(\mathcal{M})$ and $\frac{1}{2}(x + y) \in \mathrm{med}(\mathcal{M})$,*
- Local case: *there exists an arc-length parametrized geodesic $\gamma\colon I \to \mathcal{M}$ with $\|\ddot{\gamma}(0)\| = \mathrm{reach}(\mathcal{M})^{-1}$.*

Homotopy type estimation. Let $\mathcal{M}$ be a differential submanifold with positive reach. The following result allows to estimate the homotopy type of $\mathcal{M}$ based on the thickenings of a closed subset X.

Theorem II.14 (Corollary of [CCSL09, Theorem 4.6, case $\mu = 1$]). *Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$. Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that*

$$t \in [4\mathrm{d_H}(X, \mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M})).$$

The following theorem is another form of this result, under the stronger assumption that X is a finite subset of $\mathcal{M}$.

Theorem II.15 ([NSW08, Proposition 3.1]). *Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$. Suppose that $\mathcal{M}$ has positive reach. Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that*

$$t \in \left[2\mathrm{d_H}(X, \mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right).$$

In the next subsection, we present a weaker form of the reach, which still allows to recover $\mathcal{M}$ from the thickenings of X.

II.3.4 Weak feature size and μ-reach

This subsection is based on [BCY18].

Weak feature size. Let X be any compact subset of $\mathbb{R}^n$, and denote by d_X the function distance to X. It is not differentiable in general. However, one can define a *generalized gradient vector field* $\nabla d_X\colon \mathbb{R}^n \to \mathbb{R}^n$ as follows: for every $x \in \mathbb{R}^n$, let $\Gamma(x)$ denote the set of projections of x onto X. It is a compact set. Let $c_X(x)$ be the center of the smallest enclosing ball of $\Gamma(x)$. We define the *generalized gradient* of d_X as

$$\nabla d_X(x) = \frac{x - c_X(x)}{d_X(x)}.$$

It is not continuous in general. However, $\|\nabla d_X\|$ is a lower semi-continuous function, and one is able to define a flow for this gradient field, as in [BCY18, Section 9.2].

A point $x \in \mathbb{R}^n$ is called a *critical point* of d_X if $\nabla d_X(x) = 0$. Equivalently, x is a critical point if it lies in the convex hull of its projections on X. The *weak feature size* of X is defined as

$$\operatorname{wfs}(X) = \inf\{\operatorname{dist}(x, X), x \text{ is a critical point of } d_X\}.$$

The weak feature size and reach of X satisfy the inequality $\operatorname{reach}(X) \leq \operatorname{wfs}(X)$.

The Isotopy Lemma [BCY18, Theorem 9.5] states that for every $s, t \in \mathbb{R}$ such that $0 < s \leq t < \operatorname{wfs}(X)$, the thickening X^t is isotopic to X^s. However, it may happen that X^t is not isotopic X, neither homotopic. An example is given by the Warsaw circle [KSC+20, Figure 4]. Nonetheless, if X has a positive reach, then for every $t \in [0, \operatorname{wfs}(X))$, the thickening X^t deform retracts on X.

μ-reach. Let $\mu \in (0, 1]$. Let X be any subset of $\mathbb{R}^n$. The *μ-medial axis* of X is the subset $\operatorname{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ on which the distance function d_X has a small generalized gradient:

$$\operatorname{med}_\mu(X) = \{y \in \mathbb{R}^n \setminus X, \|\nabla d_X(y)\| < \mu\}.$$

The *μ-reach* of X is

$$\operatorname{reach}_\mu(X) = \inf\{\|x - y\|, x \in X, y \in \operatorname{med}_\mu(X)\}.$$

Equivalently, the μ-reach of X can be defined as the supremum of $t \geq 0$ such that the thickening X^t does not intersect $\operatorname{med}_\mu(X)$. If $\mu = 1$, the μ-reach corresponds to the reach. We have the inequality $\operatorname{reach}(X) \leq \operatorname{reach}_\mu(X) \leq \operatorname{wfs}(X)$. Hence the μ-reach can be seen as quantity that interpolates between the reach and the weak feature size of X.

Homotopy type estimation with the μ-reach. As for the reach, the thickenings of a subset with positive μ-reach deform retract on X. However, the deformation may not define a normal tubular neighborhood.

Theorem II.16 ([KSC+20, Theorem 12]). *Let $\mu \in (0, 1]$. If $X \subset \mathbb{R}^n$ is a subset with positive μ-reach, then for every $t \in [0, \operatorname{reach}_\mu(X)]$, the thickening X^t deform retracts on X.*

The following result is an equivalent of Theorem II.14 for the μ-reach.

Theorem II.17 ([CCSL09, Theorem 4.6]). *Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive μ-reach, and that*

$$\mathrm{d_H}(X, \mathcal{M}) \leq \frac{\mu^2}{5\mu^2 + 12} \cdot \operatorname{reach}_\mu(\mathcal{M}).$$

Then the thickenings X^t and $\mathcal{M}^\eta$ have the same homotopy type, provided that $\eta > 0$ is small enough, and that

$$t \in \left[\frac{4}{\mu^2}\mathrm{d_H}(X, \mathcal{M}), \operatorname{reach}_\mu(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M})\right).$$

In addition, if $\mathcal{M}$ has positive reach, we know that $\mathcal{M}^\eta$ deform retracts on $\mathcal{M}$ for η small enough, therefore the theorem gives that X^t and $\mathcal{M}$ are homotopy equivalent.

II.4 Background on persistent homology

II.4.1 Basic notions of singular and simplicial homology

This subsection is based on [Hat02]. Let G be an abelian group.

Chain complexes and homology. A *chain complex* is a sequence $C = (C_n)_{n\geq 0}$ of abelian G-modules, together with a sequence of homomorphisms $(\partial_n \colon C_n \to C_{n-1})_{n\geq 1}$, the *boundary operators*, such that $\partial_n \circ \partial_{n+1} = 0$ for all $n \geq 1$. For every $n \geq 0$, we define the *n-cycles* $Z_n(C) = \ker(\partial_n)$ and the *n-boundaries* $B_n(C) = \mathrm{im}(\partial_{n+1})$. The relation $Z_n(C) \subseteq B_n(C)$ allows to define the n^{th} homology group $H_n(C) = Z_n(C)/B_n(C)$.

If $C = (C_n)_{n\geq 0}$ and $D = (D_n)_{n\geq 0}$ are two chain complexes, a *chain complex morphism* is a collection of morphisms $\phi = (\phi_n \colon C_n \to D_n)_{n\geq 0}$ such that $\phi_n \circ \partial_{n+1} = \partial_n \circ \phi_{n+1}$. Such a morphism induces a morphism between homology groups, denoted $(\phi_n)_* \colon H_n(C) \to H_n(D)$.

The n^{th} homology is a functor $H_n \colon \mathrm{Chain} \to \mathrm{Ab}$, where Chain is the category of chain complexes, and Ab the category of abelian groups. When $G = R$ is a ring, it can be seen as a functor $H_n \colon \mathrm{Chain} \to R\text{-Mod}$, where R-Mod is the category of R-modules.

If $\phi, \psi \colon C \to D$ are two chain complex morphisms, a *chain homotopy* between ϕ and ψ is a collection of morphisms $(K_n \colon C_n \to D_{n+1})_{n\geq 0}$ such that $\phi_n - \psi_n = \partial_{n+1} \circ K_n + K_{n-1} \circ \partial_n$ for all $n \geq 0$, where K_{-1} is defined as 0. If such a chain homotopy exists, then the induced morphisms $(\phi_n)_*$ and $(\psi_n)_* \colon H_n(C) \to H_n(D)$ are equal for all $n \geq 0$.

Universal coefficient theorem. In this paragraph, we denote by $H_n(C; G)$ the homology groups of the chain complex C with coefficients in G, and $H_n(C; \mathbb{Z})$ with coefficients in $\mathbb{Z}$. The universal coefficient theorem states that there is a short exact sequence

$$0 \to H_n(C; \mathbb{Z}) \otimes G \to H_n(C; G) \to \mathrm{Tor}(H_{n-1}(C; \mathbb{Z}), G) \to 0,$$

where $\otimes$ denotes the tensor product, and Tor is the Tor functor. If $H_{n-1}(C; \mathbb{Z})$ is a free group, then $\mathrm{Tor}(H_{n-1}(C; \mathbb{Z}), G) = 0$, and we deduce that $H_n(C; G) \simeq H_n(C; \mathbb{Z}) \otimes G$.

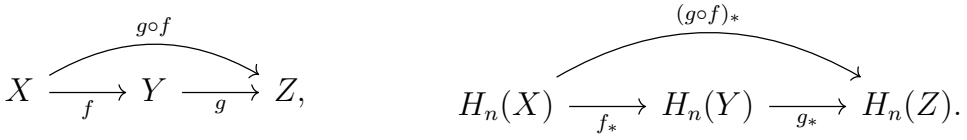
In particular, if $G = \mathbb{Z}_p$ is the finite field with p elements, we obtain a simple description of $H_n(C, \mathbb{Z}_p)$ based on $H_n(C, \mathbb{Z})$ and $H_{n-1}(C, \mathbb{Z})$. Suppose that $H_n(C, \mathbb{Z})$ and $H_{n-1}(C, \mathbb{Z})$ are finitely generated. Then we have an isomorphism $H_n(C, \mathbb{Z}_p) \simeq \mathbb{Z}_p^{k_1+k_2+k_3}$, where

- k_1 is the number of $\mathbb{Z}$ summands in $H_n(C, \mathbb{Z})$,
- k_2 is the number of $\mathbb{Z}_{p^k}$ summands in $H_n(C, \mathbb{Z})$, $k \geq 1$,
- k_3 is the number of $\mathbb{Z}_{p^k}$ summands in $H_{n-1}(C, \mathbb{Z})$, $k \geq 1$.

Singular homology. Let X be a topological space. Recall that, for every $n \geq 0$, the standard n-simplex Δ^n is the topological space defined as the convex hull of the canonical basis vectors $e_1, ..., e_{n+1}$ of $\mathbb{R}^{n+1}$, endowed with the subspace

topology. A *singular n-simplex* is a continuous map $\sigma\colon \Delta^n \to X$. For every $i \in [0,n]$, its i^{th} face is the singular $(n-1)$-simplex defined as $\delta_i\sigma\colon (t_0,...,t_{n-1}) \mapsto \sigma(t_0,...,t_i,0,t_{i+1},...,t_{n-1})$. Let $C_n(X)$ be the free group generated by the singular n-simplices and with coefficients in G. We define the boundary operator $\partial_n\colon C_n(X) \to C_{n-1}(X)$ as $\partial_n(\sigma) = \sum_{i=0}^{n}(-1)^i\delta_i\sigma$. They satisfy the relation $\delta_n \circ \delta_{n+1} = 0$. Hence the family $(C_n(X))_{n\geq 0}$, endowed with $(\partial_n)_{n\geq 0}$, is a chain complex. The corresponding homology groups are called *singular homology groups*.

The n^{th} singular homology is a functor $H_n\colon \text{Top} \to \text{Ab}$, where Top is the category of topological spaces. It associates to every topological space X a group, denoted $H_n(X)$, and to each continuous application $f\colon X \to Y$ a group homomorphism denoted $f_*\colon H_n(X) \to H_n(Y)$. The functoriality property implies that, given two continuous maps $f\colon X \to Y$ and $g\colon Y \to Z$, we have $(g\circ f)_* = g_* \circ f_*$. This property is represented by the two commutative diagrams below.



If $f, g\colon X \to Y$ are homotopic continuous applications between topological spaces, then they induce homotopic applications at the chain level, hence the induced maps $f_*, g_*\colon H_n(X) \to H_n(Y)$ are equal. As a consequence, if X and Y are homotopy equivalent topological spaces, then their cohomology groups are equal.

Simplicial homology. Let K be a simplicial complex. For every $n \geq 0$, let $C_n(K)$ be the free group generated by its simplices of dimension n and with coefficients in G. Each element of $C_n(X)$ can be written as a finite sum $\sum_i \epsilon_i\sigma_i$, where $\epsilon_i \in G$ and $\sigma_i \in K^{(n)}$. For each n-simplex $\sigma = [x_0,...,x_n]$, we define its i^{th} face as $\delta_i\sigma = [x_0,...,x_{i-1},x_{i+1},...,x_n]$, and its boundary as $\partial_n(\sigma) = \sum_{i=0}^{p}(-1)^i\delta_i\sigma$. We define the boundary of any element $\tau = \sum_i \epsilon_i\sigma_i$ of $C_n(X)$ as $\partial_n(\tau) = \sum_i \epsilon_i\partial_n(\sigma_i)$. This defines a morphism $\partial_n\colon C_n(X) \to C_{n-1}(X)$ which satisfies $\partial_n \circ \partial_{n+1}$ for all $n \geq 0$. Hence the family $(C_n(X))_{n\geq 0}$, endowed with $(\partial_n)_{n\geq 0}$, is a chain complex. The corresponding homology groups are called *simplicial homology groups*.

The n^{th} simplicial homology is a functor $H_n\colon \text{Simp} \to \text{Ab}$, where Simp is the category of simplicial complexes. The homology groups are denoted $H_n(K)$, and the morphism induced by a simplicial map $f\colon K \to L$ is denoted $f_*\colon H_n(K) \to H_n(L)$.

We say that two simplicial maps $f, g\colon K \to L$ are *contiguous* if for every simplex $\sigma \in K$, the set $f(\sigma) \cup g(\sigma)$ is a simplex of L. In this case, the induced maps $f_*, g_*\colon H_n(K) \to H_n(L)$ are equal. As a consequence, if $f\colon X \to Y$ and $g\colon Y \to X$ are two simplicial maps such that the compositions $g \circ f$ and $f \circ f$ are contiguous to the identity maps, then the simplicial homology groups of X and Y are equal.

Equivalence between singular and simplicial homology. Let S be a simplicial complex, and $|S|$ its topological realization. If $(C_n(S))_{n\geq 0}$ denotes the (simplicial) chain complex associated to S, and $(C_n(|S|))_{n\geq 0}$ the (singular) chain complex associated to the topological space $|S|$, one has a canonical chain morphism $(\phi_n\colon C_n(S) \to C_n(|S|))$ given by mapping each n-simplex σ to its characteristic map $\sigma\colon \Delta^n \to |S|$. This induces a homomorphism $H_n(S) \to H_n(|S|)$ for all $n \geq 0$, and one shows that it is an isomorphism.

Another bridge between singular and simplicial homology can be built, based on the notion of nerve. Let X be a topological space, and $\mathcal{U} = \{U_i\}_{i\in I}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that $\bigcup_{i\in I} U_i = X$. The *nerve* of $\mathcal{U}$ is the simplicial complex with vertex set $\mathcal{U}$ and whose n-simplices are the sets of $n+1$ distinct subsets $\{U_{i_0}, ..., U_{i_n}\}$ such that $\bigcap_{k=0}^{n} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$. The nerve theorem states that $\mathcal{N}(\mathcal{U})$, seen as a topological space, is homotopy equivalent to X, provided that each U_i is open, and that each non-empty intersection of finitely many sets is contractible. In particular, if $X \subset \mathbb{R}^n$ is a union of open convex sets, then it can be described as a simplicial complex. Another version of the nerve theorem can be found [BCY18, Theorem 2.9]. Suppose that X is a subset of $\mathbb{R}^n$, and $\mathcal{U} = \{U_i\}_{i\in I}$ is a cover of X. Then $\mathcal{N}(\mathcal{U})$ is homotopy equivalent to X, provided that I is finite, and each U_i are closed and convex.

In both these cases, the homotopy equivalence between $\mathcal{N}(\mathcal{U})$ and X implies that the homology groups $H_n(\mathcal{N}(\mathcal{U}))$ and $H_n(X)$ are isomorphic.

Cochain complexes and cohomology. The theory of cohomology consists in applying duality in homology theory. As a result, we obtain a contravariant functor, and the arrows go backwards.

Let $C = (C_n)_{n\geq 0}$ be a chain complex, with boundary operators $(\partial_n \colon C_n \to C_{n-1})_{n\geq 1}$. For every $n \geq 0$, let $C^i = \mathrm{Hom}(C_i, G)$ be the dual group, and $\delta_i \colon C^{i+1} \leftarrow C^i$ the dual homomorphism to ∂_{i+1}. The collection $C^* = (C^i)_{i\geq 0}$, endowed with the operators $\delta = (\delta_i)_{i\geq 0}$ is called a *cochain complex*. We define the *cohomology groups* as $H^i(C^*) = \ker(\delta_i)/\mathrm{im}(\delta_{i-1})$, with convention $\delta_{i-1} = 0$. Note that $H^i(C^*)$ is not equal to $\mathrm{Hom}(H_i(C, G))$.

As before, we can define the singular cohomology of topological spaces, and the simplicial cohomology of simplicial complexes. This yields contravariant functors $H^n \colon \mathrm{Top} \to \mathrm{Ab}$ and $H^n \colon \mathrm{Simp} \to \mathrm{Ab}$.

We now suppose that $G = R$ is a ring. If X is a topological space, the direct sum of its singular cohomology groups, denoted $H^*(X) = \bigoplus_{n\geq 0} H^i(X)$, can be given a graded ring structure. Namely, the *cup product* is an application $\smile \colon H^k(X) \times H^l(X) \to H^{k+l}(X)$ defined for all $k, l \geq 0$. This extends to an associative and distributive map $\smile \colon H^*(X) \times H^*(X) \to H^*(X)$. The same construction can be applied to simplicial cohomology. The singular and simplicial cohomology can be seen as contravariant functors $H^n \colon \mathrm{Top} \to R\text{-Alg}$ and $H^n \colon \mathrm{Simp} \to R\text{-Alg}$, where R-Alg denotes the category of R-algebras (potentially non-abelian).

Universal coefficient theorem for cohomology. In this paragraph, we denote by $H_n(C; G)$ and $H^n(C; G)$ the homology and cohomology groups of the chain complex C with coefficients in G, and $H_n(C; \mathbb{Z})$, $H^n(C; \mathbb{Z})$ with coefficients in $\mathbb{Z}$. The universal coefficient theorem for cohomology states that there is a short exact sequence

$$0 \to \mathrm{Ext}(H_{n-1}(C; \mathbb{Z}), G) \to H^n(C; G) \to \mathrm{Hom}(H_n(C; \mathbb{Z}), G) \to 0,$$

where Ext denotes the Ext functor. The term $\mathrm{Ext}(H_{n-1}(C; \mathbb{Z}), G)$ is zero when $H_{n-1}(C; \mathbb{Z})$ is a free group, in which case we deduce that

$$H^n(C; G) \simeq \mathrm{Hom}(H_n(C; \mathbb{Z}), G).$$

More generally, we can deduce a simple description of $H^n(C, \mathbb{Z})$ based on the homology groups $H_n(C, \mathbb{Z})$ and $H_{n-1}(C, \mathbb{Z})$. Suppose that they are finitely generated. Let T_n and T_{n-1} denote the torsion subgroups of $H_n(C, \mathbb{Z})$ and $H_{n-1}(C, \mathbb{Z})$, i.e., the set of elements a such that $ka = 0$ for some $k \in \mathbb{Z} \setminus \{0\}$. Then we have $H^n(C, \mathbb{Z}) \simeq \big(H_n(C, \mathbb{Z})/T_n\big) \oplus T_{n-1}$.

The universal coefficient theorem actually holds when $\mathbb{Z}$ is replaced by any principal ideal domain R, and where G is a module over R. In particular, if R and G are equal to a finite field k, we obtain the relation $H^n(C; k) \simeq \mathrm{Hom}(H_n(C; k), k)$. If the vector space $H_n(C; k)$ is finite-dimensional, then it is isomorphic to its dual, and we deduce that homology and cohomology are vector spaces of same dimension:

$$H^n(C; k) \simeq H_n(C; k).$$

II.4.2 Persistence modules

A reference for the definitions that follow is [CdSGO16]. Let T be a subset of $\mathbb{R}$, $E = \mathbb{R}^n$ the Euclidean space endowed with a norm, and k a field. Usually, T is $\mathbb{R}^+$.

Persistence modules. A *persistence module* $\mathbb{V}$ over T is a pair $(\mathbb{V}, \mathbb{v})$ where $\mathbb{V} = (V^t)_{t \in T}$ is a family of k-vector spaces, and $\mathbb{v} = (v_s^t \colon V^s \to V^t)_{s \leq t \in T}$ a family of linear maps such that:

- for every $t \in T$, $v_t^t \colon V^t \to V^t$ is the identity map,
- for every $r, s, t \in T$ such that $r \leq s \leq t$, we have $v_s^t \circ v_r^s = v_r^t$.

When the context is clear, we may denote $\mathbb{V}$ instead of $(\mathbb{V}, \mathbb{v})$.

Let us give an alternative definition: a persistence module is a functor

$$\mathbb{V} \colon (T, \leq) \to k\text{-Mod},$$

where $(T, \leq)$ is the category associated to the ordered set T, and k-Mod is the category of k-vector spaces. More precisely, the category $(T, \leq)$ has objects being the elements of T, and has an arrow $x \to y$ for every $x, y \in T$ such that $x \leq y$. This point of view is useful to generalize the notion of persistence modules we present in this subsection. For instance,

- [CZ09] defines a *multi-parameter persistence module* as a functor $(\mathbb{R}^n, \leq) \to k$-Mod, where $\leq$ denotes the usual partial order on $\mathbb{R}^n$,
- [CDS10] defines a *zigzag module* as a functor $Q \to k$-Mod, where Q is a quiver of type A_n,
- [BGO19] defines a *persistence comodule* as a contravariant functor $(\mathbb{R}, \leq) \to k$-Mod. We study this notion more precisely in Subsection II.4.5.
- [BCB20] defines a *(generalized) persistence module* as a functor $\mathcal{C} \to k$-Mod, where $\mathcal{C}$ is any small category.

Interleaving distance. Given $\epsilon \geq 0$, an *ϵ-morphism* between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon \mathbb{V}^t \to \mathbb{W}^{t+\epsilon})_{t \in T}$ such that the following diagram commutes for every $s \leq t \in T$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

If $\epsilon = 0$, ϕ is called a *morphism* of persistence modules. Moreover, if each ϕ_t is an isomorphism, the family ϕ is called an *isomorphism* of persistence modules. If $\epsilon = 0$ and $\mathbb{W} = \mathbb{V}$, the morphism ϕ is called an *endomorphism*. The set of all endomorphisms of $\mathbb{V}$ is a k-algebra and is denoted $\mathrm{End}\,(\mathbb{V})$. The persistence modules, endowed with the morphisms between them, form the *category of persistence modules*.

An *ϵ-interleaving* between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t+\epsilon})_{t \in T}$ and $(\psi_t \colon W^t \to V^{t+\epsilon})_{t \in T}$ such that the following diagrams commute for every $t \in T$:

$$\begin{array}{ccccc} V^t & & \xrightarrow{v_t^{t+2\epsilon}} & & V^{t+2\epsilon} \\ & {\scriptstyle \phi_t} \searrow & & \nearrow {\scriptstyle \psi_{t+\epsilon}} & \\ & & W^{t+\epsilon} & & \end{array} \qquad \begin{array}{ccccc} & & V^{t+\epsilon} & & \\ & {\scriptstyle \psi_t} \nearrow & & \searrow {\scriptstyle \phi_{t+\epsilon}} & \\ W^t & & \xrightarrow{w_t^{t+2\epsilon}} & & W^{t+2\epsilon} \end{array}$$

The *interleaving pseudo-distance* between $\mathbb{V}$ and $\mathbb{W}$ is defined as

$$\mathrm{d_i}\,(\mathbb{V}, \mathbb{W}) = \inf\{\epsilon \geq 0, \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}.$$

In some cases, the proximity between persistence modules is expressed with a function. Let $\eta \colon T \to T$ be a non-decreasing function such that for any $t \in T$, $\eta(t) \geq t$. A η-interleaving between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a pair of families of linear maps $(\phi_t \colon V^t \to W^{\eta(t)})_{t \in T}$ and $(\psi_t \colon W^t \to V^{\eta(t)})_{t \in T}$ such that the following diagrams commute for every $t \in T$:

$$\begin{array}{ccccc} V^t & & \xrightarrow{v_t^{\eta(\eta(t))}} & & V^{\eta(\eta(t))} \\ & {\scriptstyle \psi_t} \searrow & & \nearrow {\scriptstyle \phi_{\eta(t)}} & \\ & & W^{\eta(t)} & & \end{array} \qquad \begin{array}{ccccc} & & V^{\eta(t)} & & \\ & {\scriptstyle \phi_t} \nearrow & & \searrow {\scriptstyle \psi_{\eta(t)}} & \\ W^t & & \xrightarrow[w_t^{\eta(\eta(t))}]{} & & W^{\eta(\eta(t))} \end{array}$$

When η is $t \mapsto t + c$ for some $c > 0$, it is called an *additive c-interleaving* and corresponds with the previous definition. When η is $t \mapsto ct$ for some $c > 1$, it is called a *multiplicative c-interleaving*.

Filtrations of sets and simplicial complexes. A family of subsets $\mathbb{X} = (X^t)_{t \in T}$ of E is a *filtration* if it is non-decreasing for the inclusion, i.e. for any $s, t \in T$, if $s \leq t$ then $X^s \subseteq X^t$. Given $\epsilon \geq 0$, two filtrations $\mathbb{X} = (X^t)_{t \in T}$ and $\mathbb{Y} = (Y^t)_{t \in T}$ of E are *ϵ-interleaved* if, for every $t \in T$, $X^t \subseteq Y^{t+\epsilon}$ and $Y^t \subseteq X^{t+\epsilon}$.

The interleaving pseudo-distance between $\mathbb{X}$ and $\mathbb{Y}$ is defined as the infimum of such ϵ:

$$\mathrm{d_i}(\mathbb{X}, \mathbb{Y}) = \inf\{\epsilon, \mathbb{X} \text{ and } \mathbb{Y} \text{ are } \epsilon\text{-interleaved}\}.$$

Filtrations of simplicial complexes and their interleaving distance are similarly defined: given a simplicial complex S, a *filtration of* S is a non-decreasing family $\mathbb{S} = (S^t)_{t\in T}$ of subcomplexes of S. The interleaving pseudo-distance between two filtrations $(S_1^t)_{t\in T}$ and $(S_2^t)_{t\in T}$ of S is the infimum of the $\epsilon \geq 0$ such that they are ϵ-interleaved, i.e. for any $t \in T$, $S_1^t \subseteq S_2^{t+\epsilon}$ and $S_2^t \subseteq S_1^{t+\epsilon}$.

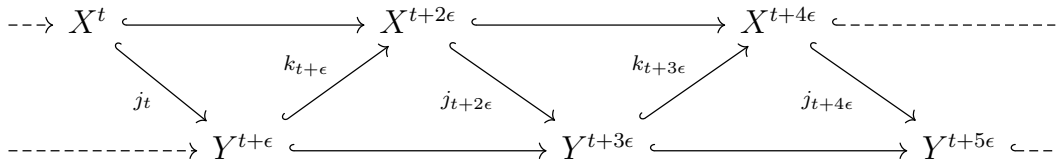
Relation between filtrations and persistence modules. A common procedure to build persistence modules consists in applying the i^{th} homology functor to a filtration. Namely, if $\mathbb{X} = (X^t)_{t\in T}$ is a set filtration, with $(i_s^t \colon X^s \to X^t)_{s\leq t}$ the inclusion maps, then the collection $\mathbb{V}[\mathbb{X}] = (H_i(X^t))_{t\in T}$ is a persistence module, with maps $((i_s^t)_* \colon H_i(X^s) \to H_i(X^t))_{s\leq t}$ being induced by the inclusions. This is pictured by the two following diagrams.

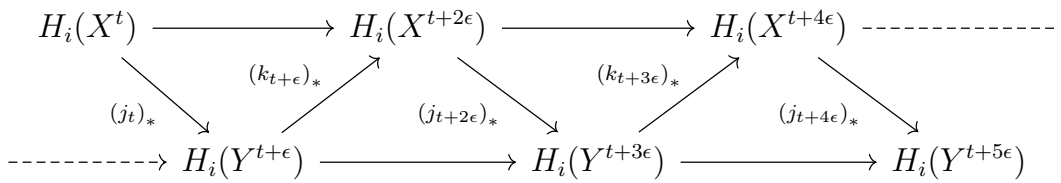
$$\dashrightarrow X^{t_1} \xhookrightarrow{i_{t_1}^{t_2}} X^{t_2} \xhookrightarrow{i_{t_2}^{t_3}} X^{t_3} \xhookrightarrow{i_{t_3}^{t_4}} X^{t_4} \hookleftarrow\dashleftarrow$$

$$\dashrightarrow H_i(X^{t_1}) \xrightarrow{(i_{t_1}^{t_2})_*} H_i(X^{t_2}) \xrightarrow{(i_{t_2}^{t_3})_*} H_i(X^{t_3}) \xrightarrow{(i_{t_3}^{t_4})_*} H_i(X^{t_4}) \dashleftarrow$$

The persistence module $\mathbb{V}[\mathbb{X}]$ is called the persistence module *associated to* the filtration $\mathbb{X}$, or *corresponding to* $\mathbb{X}$.

As a consequence of this construction, if two filtrations $\mathbb{X}$ and $\mathbb{Y}$ are ϵ-interleaved, then their associated persistence modules $\mathbb{V}[\mathbb{X}]$ and $\mathbb{V}[\mathbb{Y}]$ are also ϵ-interleaved, the interleaving homomorphisms being induced by the interleaving inclusion maps. This is pictured by the two following diagrams. The first one represents the interleaving of the filtrations, and the second one of the persistende modules.





Similarly, one can apply the simplicial homology functor to a simplicial filtration to obtain a persistence module, and the previous remark holds as well.

Čech set filtrations. Let X denote any subset of E. The *Čech set filtration* associated to X is the filtration of E defined as the collection of subsets $V[X] = (X^t)_{t\geq 0}$, where X^t denotes the t-thickening of X in E (see Subsection II.3.2). By

applying the i^{th} homology functor to $V[X]$, we obtain a persistence module, that we denote $\mathbb{V}[X]$.

If X is a $\mathcal{C}^{1,1}$-submanifold, we have seen in Subsection II.3.3 that X has positive reach, and that X^t deform retracts on X for every $t \in [0, \operatorname{reach}(X))$. In this case, the corresponding persistence module $\mathbb{V}[X]$ is constant on the interval $[0, \operatorname{reach}(X))$, and is equal to the homology group $H_i(X)$. Moreover, if Y is any other subset of E with Hausdorff distance $\mathrm{d_H}(X, Y) \leq \epsilon$, then the persistence module $\mathbb{V}[Y]$ is constant on the interval $[4\epsilon, \operatorname{reach}(X) - 3\epsilon)$ and is equal to the homology group $H_i(X)$.

We can state similar results for the μ-reach, $\mu \in (0, 1]$. If X is a subset with positive μ-reach, then X^t deform retracts on X for every $t \in [0, \operatorname{reach}_\mu(X)]$. Accordingly, the persistence module $\mathbb{V}[X]$ is constant on the interval $[0, \operatorname{reach}(X))$, and is equal to $H_i(X)$. Moreover, if X has positive reach, and if Y is any other subset of E with Hausdorff distance $\mathrm{d_H}(X, Y) \leq \epsilon$, then $\mathbb{V}[Y]$ is constant on the interval $[4\epsilon/\mu^2, \operatorname{reach}_\mu(\mathcal{M}) - 3\epsilon)$ and is equal to $H_i(X)$, provided that $\epsilon \leq \frac{\mu^2}{5\mu^2+12}\operatorname{reach}_\mu(\mathcal{M})$ (see Theorem II.17).

Čech simplicial filtrations. Let X denote a finite subset of E and $V[X] = (X^t)_{t\geq 0}$ its associated Čech set filtration. For all $t \geq 0$, X^t is a union of closed balls of radius t: $X^t = \bigcup_{x\in X} \overline{\mathcal{B}}(x, t)$. Consider the simplicial filtration $\mathcal{V}[X] = (\mathcal{V}^t[X])_{t\geq 0}$, where $\mathcal{V}^t[X]$ is the nerve of the cover $\mathcal{V}^t$ defined as $\mathcal{V}^t = \{\overline{\mathcal{B}}(x, t), x \in X\}$. It is called the *Čech simplicial filtration* associated to X. The persistent nerve lemma [CO08, Lemma 3.4] connects these two constructions:

Lemma II.18 ([CO08, Lemma 3.4]). *The persistence (singular) homology module associated to $V[X]$ and the persistence (simplicial) homology module associated to $\mathcal{V}[X]$ are isomorphic.*

This result is fundamental, since it allows to study Čech filtrations via their simplicial counterparts, which can be computed in practice.

Vietoris-Rips filtrations. Let X denote a finite subset of E and $V[X] = (\mathcal{V}^t[X])_{t\geq 0}$ the corresponding simplicial Čech filtration. For every $t \geq 0$, let $\operatorname{Rips}(\mathcal{V}^t[X])$ be the flag complex associated to $\mathcal{V}^t[X]$. The collection $\operatorname{Rips}(\mathcal{V}[X]) = (\operatorname{Rips}(\mathcal{V}^t[X]))_{t\geq 0}$ is called the *Vietoris-Rips filtration* associated to X. It does not yield a persistence module which is isomorphic to the persistent homology module of the Čech filtration. Nonetheless, the Vietoris-Rips filtration and the Čech simplicial filtration are c-close in multiplicative interleaving distance, with $c = \sqrt{\frac{2n}{n+1}}$, and where n is the dimension of the ambient space E [BLM+19a, Theorem 3.1]. In other words, for every $t \geq 0$, we have

$$\mathcal{N}(\mathcal{V}^t[X]) \subseteq \operatorname{Rips}(\mathcal{V}^t[X]) \subseteq \mathcal{N}(\mathcal{V}^{ct}[X]).$$

Sublevel-set filtrations. Let $f\colon E \to \mathbb{R}$ be any function. For every $t \in \mathbb{R}$, let $f^t = f^{-1}((-\infty, t])$ be the sublevel set of f. The family $V[f] = (f^t)_{t\in\mathbb{R}}$ is a set filtration, called the *sublevel-set filtration* of f. As before, we will write $\mathbb{V}[f]$ for the corresponding i^{th} persistence module.

If $f, g\colon E \to \mathbb{R}$ are two functions with finite sup norm $\|f - g\|_\infty$, then it is direct to see that the sublevel-set filtrations $V[f]$ and $V[g]$ are ϵ-interleaved, with

$\epsilon = \|f - g\|_\infty$. Moreover, their interleaving distance is exactly ϵ. In particular, we obtain that the corresponding persistence modules $\mathbb{V}[f]$ and $\mathbb{V}[g]$ are also ϵ-interleaved. This can be written as

$$\mathrm{d_i}\left(V[f], V[g]\right) = \|f - g\|_\infty \qquad \text{and} \qquad \mathrm{d_i}\left(\mathbb{V}[f], \mathbb{V}[g]\right) \leq \|f - g\|_\infty.$$

This result is a first instance of the stability of persistence modules, that we investigate in Subsection II.4.4.

It is worth noting that the Čech filtration $V[X]$ of a closed subset $X \subset E$ is equal to a sublevel-set filtration. Namely, the filtration associated to the function distance to X. We recall that the Hausdorff distance $\mathrm{d_H}\left(X, Y\right)$ between two subsets $X, Y \subset E$ is equal to the sup norm of their distance functions (see Subsection II.3.1). We deduce that the Čech filtrations $V[X]$ and $V[Y]$ are $\mathrm{d_H}\left(X, Y\right)$-interleaved, as well as the corresponding persistence modules:

$$\mathrm{d_i}\left(V[X], V[Y]\right) = \mathrm{d_H}\left(X, Y\right) \qquad \text{and} \qquad \mathrm{d_i}\left(\mathbb{V}[X], \mathbb{V}[Y]\right) \leq \mathrm{d_H}\left(X, Y\right).$$

II.4.3 Decomposition of persistence modules

We follow the presentation of [CdSGO16]. As in the previous subsection, the persistence modules we consider here are seen as functors $\mathbb{V}\colon (T, \leq) \to k\text{-Mod}$, where T is a subset of $\mathbb{R}$.

Decomposability. Let $(\mathbb{V}, \mathbb{v})$ and $(\mathbb{W}, \mathbb{w})$ be two persistence modules. Their *sum* is the persistence module $\mathbb{V} \oplus \mathbb{W}$ defined with the vector spaces $(V \oplus W)^t = V^t \oplus W^t$ and the linear maps

$$(v \oplus w)^t_s\colon (x, y) \in (V \oplus W)^s \longmapsto (v^t_s(x), w^t_s(y)) \in (V \oplus W)^t.$$

A persistence module $\mathbb{U}$ is *indecomposable* if for every pair of persistence modules $\mathbb{V}$ and $\mathbb{W}$ such that $\mathbb{U}$ is isomorphic to the sum $\mathbb{V} \oplus \mathbb{W}$, then one of the summands has to be a trivial persistence module, that is, equal to zero for every $t \in T$. Otherwise, $\mathbb{U}$ is said *decomposable*.

If $\mathbb{U} = \mathbb{V} \oplus \mathbb{W}$ is a decomposable persistence module, then the projection on $\mathbb{V}$ and the projection on $\mathbb{W}$ are two endomorphisms of $\mathbb{U}$. They are idempotent elements of $\mathrm{End}\left(\mathbb{U}\right)$.

Interval modules. Let $I \subset T$ be an interval, that is, a non-empty convex set. Intervals have the form $[a, b]$, $(a, b]$, $[a, b)$ or (a, b), with $a, b \in T$ such that $a \leq b$, and potentially $a = -\infty$ or $b = +\infty$. The *interval module* associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v^t_s\colon \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as

$$\mathbb{B}^t[I] \ = \begin{cases} k & \text{if } t \in T, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad v^t_s \ = \begin{cases} \mathrm{id} & \text{if } s, t \in T, \\ 0 & \text{otherwise.} \end{cases}$$

The endomorphism algebra of an interval module is isomorphic to k. In particular, its only idempotents are 0 and 1. According to the last paragraph, we deduce:

Lemma II.19. *The interval modules are indecomposable.*

As a partial converse, we will see that an indecomposable persistence modules is an interval module, provided that it is pointwise finite-dimensional.

Persistence barcodes and persistence diagrams. A persistence module $\mathbb{V}$ *decomposes into interval module* if there exists a set $\{\mathbb{B}_i, i \in \mathcal{I}\}$ of interval modules such that $\mathbb{V}$ is isomorphic to the sum $\bigoplus_{i\in\mathcal{I}} \mathbb{B}_i$. Equivalently, there exists a multiset $\mathcal{I}$ of intervals of T such that

$$\mathbb{V} \simeq \bigoplus_{I\in\mathcal{I}} \mathbb{B}[I].$$

Multiset means that $\mathcal{I}$ may contain several copies of the same interval I. Such a module is said *decomposable into interval modules*, or simply *decomposable* when the context is clear.

The following theorem is a consequence of Krull–Remak–Schmidt–Azumaya's theorem [Azu50, Theorem 1].

Theorem II.20 ([CdSGO16, Theorem 1.3]). *If a persistence module decomposes into interval modules, then the multiset $\mathcal{I}$ of intervals is unique.*

In this case, $\mathcal{I}$ is called the *persistence barcode* of $\mathbb{V}$, or simply *barcode*. It is written $\mathrm{Barcode}\,(\mathbb{V})$.

Let $\mathbb{V}$ be a decomposable persistence module and $\mathrm{Barcode}\,(\mathbb{V})$ its barcode. For every $[a, b]$, $(a, b]$, $[a, b)$ or (a, b) in $\mathrm{Barcode}\,(\mathbb{V})$, with potentially $a = -\infty$ or $b = +\infty$, consider the point (a, b) of $\mathbb{R}^2$. The collection of all such points is a multiset, that we call the *persistence diagram* of $\mathbb{V}$. It is denoted $\mathrm{Diagram}\,(\mathbb{V})$. In the context of [CdSGO16, Subsection 1.6], it is called an undecorated persistence diagram. It is sometimes required to remove the points of the form (a, a) from $\mathrm{Diagram}\,(\mathbb{V})$, though we won't need this distinction in what follows.

Decomposition of pointwise finite-dimensional modules. A persistence module $\mathbb{V}$ is said *pointwise finite dimensional* if every vector space V^t has finite dimension. We have:

Theorem II.21 ([CB15, Theorem 2.1]). *Every pointwise finite-dimensional persistence module decomposes into interval modules.*

An example of a persistence module that does not decompose into interval modules can be found in [CdSGO16, Theorem 1.4 (3)].

This theorem does not hold for generalized definitions of persistence modules, where the notion of interval modules may not even be well-defined. Although, some weaker results exist [CB15, Theorem 1.1].

Decomposition of q-tame modules. A persistence module $\mathbb{V}$ is said to be *q-tame* if for every $s, t \in T$ such that $s < t$, the map v_s^t has finite rank. It is a generalization of pointwise finite-dimensional persistence modules. As an example, the sublevel-set filtrations of functions $f\colon \mathbb{R}^n \to \mathbb{R}$ induces q-tame persistence modules, provided that f is continuous and proper [CdSGO16, Theorem 2.22]. The

q-tameness of a persistence module ensures that we can define a notion of persistence diagram, called a *persistence measure* [CdSGO16, Subsection 2.1].

Another point of view consists in studying persistence modules in the *observable category of persistence modules*, as in [CCBDS14]. In this category, q-tame modules become interval-decomposable, and persistence diagrams are a complete invariant (see also [Oud15, Theorem 3.7]).

II.4.4 Stability of persistence modules

This subsection is based on [Oud15].

Bottleneck distance. We define the bottleneck distance for persistence diagrams, that is, multisets $\{(a_i, b_i), i \in \mathcal{I}\}$ of $\overline{\mathbb{R}}^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$. Given two multisets P and Q, a *partial matching* between them is a subset M of $P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The points $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said *matched* by M. If a point $p \in P$ (resp. $q \in Q$) is not matched by M, we consider that it is matched with its projection $\overline{p}$ (resp. $\overline{q}$) on the diagonal $\Delta = \{(a, a), a \in \mathbb{R}\}$. The *cost* of a matched pair (p, q) (resp. $(p, \overline{p})$, resp. $(\overline{q}, q)$) is the sup norm $\|p - q\|_\infty$ (resp. $\|p - \overline{p}\|_\infty$, resp. $\|\overline{q} - q\|_\infty$). The *cost* of the partial matching M, denoted $\mathrm{cost}(M)$, is the supremum of all such costs. The *bottleck distance* between P and Q is defined as the infimum of costs over all the partial matchings:

$$\mathrm{d_b}(P, Q) = \inf\{\mathrm{cost}(M),\ M \text{ is a partial matching between } P \text{ and } Q\}.$$

If $\mathbb{U}$ and $\mathbb{V}$ are two decomposable persistence modules, we define their *bottleneck distance* as

$$\mathrm{d_b}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathrm{Diagram}(\mathbb{U}), \mathrm{Diagram}(\mathbb{V})).$$

The isometry theorem. At this point, the category of interval-decomposable modules is endowed with two notions of distance: the interleaving distance and the bottleneck distance. The following result is fundamental in persistence theory:

Theorem II.22 ([CdSGO16, Theorem 4.11, Isometry theorem]). *If the persistence modules $\mathbb{U}$ and $\mathbb{V}$ are interval-decomposable, then $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathbb{U}, \mathbb{V})$.*

This result falls into two parts: the *stability theorem*, $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \geq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$, and the *converse stability theorem*, $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \leq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$.

The converse stability theorem is proven directly by building an interleaving between $\mathbb{U}$ and $\mathbb{V}$ from a partial matching between $\mathrm{Diagram}(\mathbb{U})$ and $\mathrm{Diagram}(\mathbb{V})$. One starts by proving it for interval modules, and generalizes it to interval-decomposable modules by taking sums of intervals.

The stability theorem is less simple to prove. One way of tackling the problem consists in using the interpolation lemma:

Lemma II.23 ([CdSGO16, Lemma 3.4]). *If $\mathbb{U}$ and $\mathbb{V}$ are δ-interleaved, then there exists a family of persistence modules $(\mathbb{U}_t)_{t\in[0,\delta]}$ such that $\mathbb{U}_0 = \mathbb{U}$, $\mathbb{U}_\delta = \mathbb{V}$ and $\mathrm{d_i}(\mathbb{U}_s, \mathbb{U}_t) \leq |s-t|$ for every $s, t \in [0,\delta]$.*

The theorem then follows from the box lemma [CdSGO16, Lemma 4.22] and a compacity argument. Another proof of the stability theorem is given in [BL13], which has the advantage of building an explicit partial matching from an interleaving.

The stability theorem generalizes to q-tame modules. Given two q-tame persistence modules $\mathbb{U}$ and $\mathbb{V}$, one defines their interleaving distance $\mathrm{d_i}(\mathbb{U}, \mathbb{V})$ and their bottleneck distance $\mathrm{d_b}(\mathbb{U}, \mathbb{V})$, and we still have $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathbb{U}, \mathbb{V})$ [CdSGO16, Theorem 4.11].

Using the results of Subsection II.4.2, the isometry theorem (more precisely the stability theorem) yields two immediate consequences.

- Let $f, g\colon \mathbb{R}^n \to \mathbb{R}$ be two continuous and proper functions, so that the persistence modules $\mathbb{V}[f]$ and $\mathbb{V}[g]$ associated to their sublevel-set filtrations are q-tame. Then $\mathrm{d_b}(\mathbb{V}[f], \mathbb{V}[g]) \leq \|f-g\|_\infty$.
- Let $X, Y \subset \mathbb{R}^n$ be two bounded subsets, so that the persistence modules $\mathbb{V}[X]$ and ${}^{\mathbb{V}}[Y]$ associated to their Čech filtrations are q-tame. Then $\mathrm{d_b}(\mathbb{V}[X], \mathbb{V}[Y]) \leq \mathrm{d_H}(X, Y)$.

The generalized persistent nerve theorem. We have seen with Lemma II.18 that the persistent homology of a Čech filtration can be computed via its simplicial version. A stronger statement is known as the generalized persistent nerve theorem, and can be seen as a stability result for simplicial filtrations that come from nerves. Let $T = \mathbb{R}^+$. A *cover filtration* is a collection $\mathcal{U} = \{\mathcal{V}_1, ..., \mathcal{V}_n\}$ of simplicial filtrations $\mathcal{V}_i = (V_i^t)_{t\in T}$ of a fixed simplicial complex S. It is an *ϵ-good cover filtration*, with $\epsilon \geq 0$, if for every subset $\sigma \subset [1,n]$, $t \in T$ and $i \geq 0$, the map induced in homology $H_i(\bigcap_{i\in\sigma} V_i^t) \to H_i(\bigcap_{i\in\sigma} V_i^{t+\epsilon})$ is zero.

Theorem II.24 ([CS18, Theorem 7]). *Let $\mathcal{U} = \{\mathcal{V}_1, ..., \mathcal{V}_n\}$ be a finite ϵ-good cover fitration of a simplicial complex S. Consider its nerve filtration $\mathcal{N}(\mathcal{U})$. Let $W = (W^t)_{t\in T}$ be the simplicial filtration defined as $W^t = \bigcup_{i\in[1,n]} V_i^t$. Then the i^{th}-homology persistence modules associated to $\mathcal{N}(\mathcal{U})$ and W are $(i+1)\epsilon$-interleaved.*

II.4.5 Persistent cohomology theory

In this subsection, we write down the definitions of persistence comodules, and their associated pseudo-distances. Compared to the standard definitions of persistent homology, the arrows go backward. Let $T \subseteq [0, +\infty)$ be an interval that contains 0, let E be a Euclidean space, and k a field. Apart from this subsection, we will use the terms *persistence modules* and *persistence comodules* indifferently in the rest of this document.

Persistence comodules. A *persistence comodule* over T is a pair $(\mathbb{V}, \mathbb{v})$ where $\mathbb{V} = (V^t)_{t\in T}$ is a family of k-vector spaces, and $\mathbb{v} = (v_s^t)_{s\leq t\in T}$ is a family of linear maps $v_s^t\colon V^s \leftarrow V^t$ such that:

- for every $t \in T$, $v_t^t \colon V^t \leftarrow V^t$ is the identity map,
- for every $r, s, t \in T$ such that $r \leq s \leq t$, $v_r^s \circ v_s^t = v_r^t$.

When there is no risk of confusion, we may denote a persistence comodule by $\mathbb{V}$ instead of $(\mathbb{V}, \mathbb{v})$. Given $\epsilon \geq 0$, an *ϵ-morphism* between two persistence comodules $(\mathbb{V}, \mathbb{v})$ and $(\mathbb{W}, \mathbb{w})$ is a family of linear maps $(\phi_t \colon V^t \to W^{t-\epsilon})_{t \geq \epsilon}$ such that the following diagram commutes for every $\epsilon \leq s \leq t$:

$$\begin{array}{ccc} V^s & \xleftarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s-\epsilon} & \xleftarrow[w_{s-\epsilon}^{t-\epsilon}]{} & W^{t-\epsilon} \end{array}$$

If $\epsilon = 0$ and each ϕ_t is an isomorphism, the family $(\phi_t)_{t \in T}$ is an *isomorphism* of persistence comodules. An *ϵ-interleaving* between two persistence comodules $(\mathbb{V}, \mathbb{v})$ and $(\mathbb{W}, \mathbb{w})$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t-\epsilon})_{t \geq \epsilon}$ and $(\psi_t \colon W^t \to V^{t-\epsilon})_{t \geq \epsilon}$ such that the following diagrams commute for every $t \geq 2\epsilon$:

$$\begin{array}{ccccc} V^{t-2\epsilon} & & \xleftarrow{v_{t-2\epsilon}^t} & & V^t \\ & {\scriptstyle \psi_{t-\epsilon}} \nwarrow & & \swarrow {\scriptstyle \phi_t} & \\ & & W^{t-\epsilon} & & \end{array} \qquad \begin{array}{ccccc} & & V^{t-\epsilon} & & \\ & {\scriptstyle \phi_{t-\epsilon}} \swarrow & & \nwarrow {\scriptstyle \psi_t} & \\ W^{t-2\epsilon} & & \xleftarrow[w_{t-2\epsilon}^t]{} & & W^t \end{array}$$

The interleaving pseudo-distance between $(\mathbb{V}, \mathbb{v})$ and $(\mathbb{W}, \mathbb{w})$ is defined as

$$\mathrm{d_i}(\mathbb{V}, \mathbb{W}) = \inf\{\epsilon \geq 0,\ \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}.$$

Persistence barcodes. A persistence comodule $(\mathbb{V}, \mathbb{v})$ is said to be *pointwise finite-dimensional* if for every $t \in T$, V^t is finite-dimensional. This implies that we can define a notion of persistence barcode [BCB20, Theorem 1.2]. It comes from the algebraic decomposition of the persistence comodule into interval modules. Moreover, given two pointwise finite-dimensional persistence comodules $\mathbb{V}, \mathbb{W}$ with persistence barcodes $\mathrm{Barcode}(\mathbb{V}), \mathrm{Barcode}(\mathbb{W})$, the so-called isometry theorem states that $\mathrm{d_b}(\mathrm{Barcode}(\mathbb{V}), \mathrm{Barcode}(\mathbb{W})) = \mathrm{d_i}(\mathbb{V}, \mathbb{W})$ where $\mathrm{d_i}(\cdot, \cdot)$ denotes the interleaving distance between persistence comodules, and $\mathrm{d_b}(\cdot, \cdot)$ denotes the bottleneck distance between barcodes.

More generally, the persistence comodule $(\mathbb{V}, \mathbb{v})$ is said to be *q-tame* if for every $s, t \in T$ such that $s < t$, the map v_s^t is of finite rank. The q-tameness of a persistence comodule ensures that we can still define a notion of persistence barcode, even though the comodule may not be decomposable into interval modules. Moreover, the isometry theorem still holds [CdSGO16, Theorem 4.11].

Relation between filtrations and persistence comodules. Applying the singular cohomology functor to a set filtration (defined in Subsection II.4.2) gives rise to a persistence comodule whose linear maps between cohomology groups are induced by the inclusion maps between sets. As a consequence, if two filtrations are ϵ-interleaved, then their associated persistence comodules are also ϵ-interleaved, the interleaving homomorphisms being induced by the interleaving inclusion maps.

As a consequence of the isometry theorem, if the comodules are q-tame, then the bottleneck distance between their persistence barcodes is upperbounded by ϵ.

The same remarks hold when applying the simplicial cohomology functor to simplicial filtrations.

II.5 Background on measure theory

II.5.1 Basic notions of measure theory

We assume that the notions of measures and probability measures are known. Let $E = \mathbb{R}^n$.

Wasserstein distances. We use the definition of [Vil08]. Given two probability measures μ and ν over E, a transport plan between μ and ν is a probability measure π over $E \times E$ whose marginals are μ and ν. Let $p \geq 1$. The *p-Wasserstein distance between μ and ν* is defined as

$$\mathrm{W}_p(\mu, \nu) = \left(\inf_{\pi} \int_{E \times E} \|x - y\|^p \mathrm{d}\pi(x, y) \right)^{\frac{1}{p}},$$

where the infimum is taken over all the transport plans π.

The p-Wasserstein distance is a well-defined distance on the set of measures with finite p^{th} moment. Moreover, it metrizes the weak-convergence on this set. If q is such that $p \leq q$, then an application of Jensen's inequality shows that $\mathrm{W}_p(\mu, \nu) \leq \mathrm{W}_q(\mu, \nu)$.

If $X, Y \subset \mathbb{R}^n$ are finite subsets, we denote by $\mathrm{W}_p(X, Y)$ the Wasserstein distance between their empirical measures.

When $p = 1$, a particular duality formulation occurs, known as the *Kantorovich–Rubinstein duality* [Vil08, Remark 6.5]: for every probability measure μ and ν, we have

$$\mathrm{W}_1(\mu, \nu) = \sup \left\{ \left| \int \phi \cdot \mathrm{d}\mu - \int \phi \cdot \mathrm{d}\nu \right|, \phi \text{ is 1-Lipschitz} \right\}. \tag{II.3}$$

Empirical measures. For every $x \in E$, the Dirac mass on x is denoted δ_x. If $X = \{x_1, ..., x_n\}$ is a finite subset of E, the *empirical measure* associated to X is the measure

$$\mu_X = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}.$$

Let $p \geq 1$. Suppose that X is an i.i.d sequence of a measure μ, and that μ admits a finite q^{th} moment, with $q > p$. Then the Wasserstein distance

$$\mathrm{W}_p(\mu_X, \mu)$$

converges (in mean) to zero when n goes to $+\infty$. Explicit bounds are given in [FG15].

II.5.2 Distance-to-measure

This subsection is based on [CCSM11]. Let $T = \mathbb{R}^+$ and $E = \mathbb{R}^n$ endowed with the standard Euclidean norm.

DTM. Let μ be a probability measure over E, and $m \in [0, 1)$ a parameter. For every $x \in E$, let $\delta_{\mu,m}$ be the function defined on E by

$$\delta_{\mu,m}(x) = \inf \left\{ r \geq 0, \mu\left(\overline{\mathcal{B}}\left(x, r\right)\right) > m \right\}.$$

The *distance to measure (DTM)* associated to μ with parameter m is the function $\mathrm{d}_{\mu,m} \colon E \to \mathbb{R}$ defined as:

$$\mathrm{d}^2_{\mu,m}(x) = \frac{1}{m} \int_0^m \delta^2_{\mu,t}(x) \mathrm{d}t.$$

When m is fixed and there is no risk of confusion, we may write d_μ instead of $\mathrm{d}_{\mu,m}$. We cite two important properties of the DTM:

Proposition II.25 ([CCSM11, Corollary 3.7]). *For every probability measure μ and $m \in [0, 1)$, $\mathrm{d}_{\mu,m}$ is 1-Lipschitz.*

Theorem II.26 ([CCSM11, Theorem 3.5]). *For every probability measures μ, ν and $m \in (0, 1)$, we have $\|\mathrm{d}_{\mu,m} - \mathrm{d}_{\nu,m}\|_\infty \leq m^{-\frac{1}{2}} \mathrm{W}_2\left(\mu, \nu\right)$.*

In practice, the DTM can be computed. If X is a finite subset of E of cardinal n, we denote by μ_X its empirical measure. Assume that $m = \frac{k_0}{n}$, with k_0 an integer. In this case, $\mathrm{d}_{\mu_X,m}$ reformulates as follows: for every $x \in E$,

$$\mathrm{d}^2_{\mu_X,m}(x) = \frac{1}{k_0} \sum_{k=1}^{k_0} \|x - p_k(x)\|^2,$$

where $p_1(x), ..., p_{k_0}(x)$ are a choice of k_0-nearest neighbors of x in X.

In Subsections IV.4.2 and IV.4.3, we will consider the DTM for measures on the vector space $\mathbb{R}^m$ endowed with the norm $\|\cdot\|_\gamma$, which is not the usual Euclidean norm. In this case, the definition of the DTM is the same as previously, but using the norm $\|\cdot\|_\gamma$ instead of $\|\cdot\|$.

Homotopy type estimation with the DTM. The following theorem shows that the sublevel sets $\mathrm{d}^t_{\mu,m}$ of $\mathrm{d}_{\mu,m}$ can be used to estimate the homotopy type of $\mathrm{supp}(\mu)$.

Theorem II.27 ([CCSM11, Corollary 4.11]). *Let $m \in (0, 1)$, μ any probability measure on E, and denote $K = \mathrm{supp}(\mu)$. Suppose that $\mathrm{reach}\left(K\right) = \tau > 0$, and that μ satisfies the following hypothesis for $r < \left(\frac{m}{a}\right)^{\frac{1}{d}}$: $\forall x \in K, \mu(\mathcal{B}(x, r)) \geq \min\{ar^d, 1\}$. Let ν be another measure, and denote $w = \mathrm{W}_2\left(\mu, \nu\right)$. Suppose that $w \leq m^{\frac{1}{2}}\left(\frac{\tau}{9} - \left(\frac{m}{a}\right)^{\frac{1}{d}}\right)$. Define $\epsilon = \left(\frac{m}{a}\right)^{\frac{1}{d}} + m^{-\frac{1}{2}} w$ and choose $t \in [4\epsilon, \tau - 3\epsilon]$. Then $\mathrm{d}^t_{\nu,m}$ and K are homotopic equivalent.*

Figure II.5 is an example of application of this theorem. It represents a point cloud sampled on a mechanical part, with some anomalous points added. Let μ be the empirical measure on this point cloud. The sublevel-set of the DTM $\mathrm{d}^t_{\mu,m}$, for well-chosen parameters m and t, has the homotopy type of the underlying object.

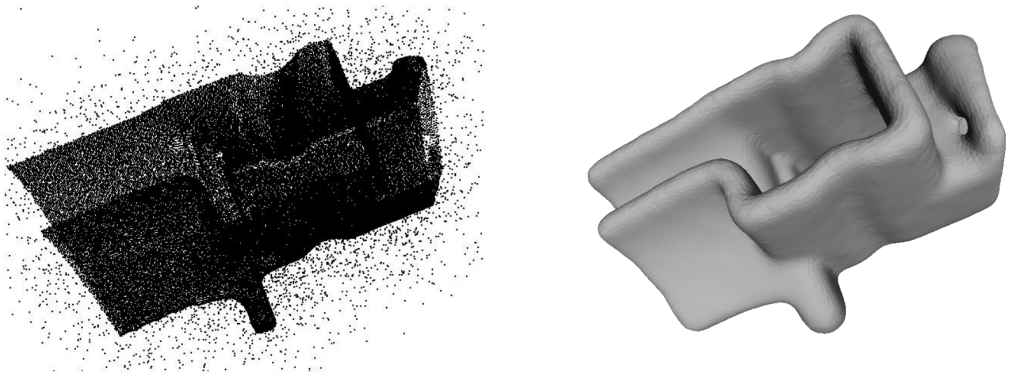


Figure II.5: Reprinted from [CCSM11, Figure 1]. **Left:** A point cloud in $\mathbb{R}^3$. **Right:** A sublevel-set of the DTM.

II.5.3 Varifolds

We follow the presentation of [BLM17, Section 2]. Let $E = \mathbb{R}^n$ be the Euclidean space, d an integer and $\mathcal{G}_d(E)$ the Grassmannian of d-planes of E.

Varifolds. A *d-varifold* is a Radon measure V on $E \times \mathcal{G}_d(E)$, that is, a Borel measure which takes finite values on compacts. The *mass* of V is the measure on E defined as the push-forward $\|V\| = (\mathrm{proj}_1)_* V$, where $\mathrm{proj}_1 \colon E \times \mathcal{G}_d(E) \to E$ is the projection on the first coordinate.

The contributions of V on E and $\mathcal{G}_d(E)$ can be separated via desintegration. Namely, there exists a set of probability measure $\{\nu_x\}_{x\in E}$, defined $\|V\|$-almost surely, such that $V = \|V\| \otimes \{\nu_x\}_x$ [BLM17, Proposition 2.4]. In other words, for every continuous function $\phi \colon E \times \mathcal{G}_d(E) \to \mathbb{R}$ with compact support, we have

$$\int_{(x,T)\in E\times\mathcal{G}_d(E)} \phi(x,T)\cdot \mathrm{d}V(x,T) = \int_{x\in E} \left(\int_{T\in\mathcal{G}_d(E)} \phi(x,T)\cdot \mathrm{d}\nu_x(T) \right) \mathrm{d}\,\|V\|\,(x).$$

Varifolds are endowed with the weak convergence, that is, a sequence of d-varifolds $(V_i)_{i\geq 0}$ converges to V if for every $\phi \colon E \times \mathcal{G}_d(E) \to \mathbb{R}$ with compact support, we have the convergence:

$$\int_{E\times\mathcal{G}_d(E)} \phi(x,T)\cdot \mathrm{d}V_i(x,T) \longrightarrow \int_{E\times\mathcal{G}_d(E)} \phi(x,T)\cdot \mathrm{d}V(x,T).$$

Besides, the varifolds are also endowed with the *flat distance*, or *bounded Lipschitz distance*, defined for every d-varifolds U, V as:

$$\Delta_{1,1}(U,V) = \sup\left\{ \left| \int \phi\cdot \mathrm{d}U - \int \phi\cdot \mathrm{d}V \right|, \|\phi\|_\infty \leq 1, \phi \text{ is 1-Lipschitz} \right\}. \quad \text{(II.4)}$$

If a sequence of d-varifolds $(V_i)_{i\geq 0}$ has masses uniformly bounded and supports included in a fixed compact, then the weak convergence is equivalent to the convergence for $\Delta_{1,1}$ [BLM17, Proposition 2.8]. It is worth noting that the flat distance stands for a natural generalization of the Wasserstein distance W_1, allowing measure with different masses. If U and V are probability measures with supports included in a set of diameter 1, then these two distances are equal (compare Equations (II.4) and (II.3)).

Rectifiable varifolds. A subset $\mathcal{M} \subset E$ is a *d-rectifiable set* if its d-Hausdorff measure $\mathcal{H}^d(\mathcal{M})$ is finite and if there exists countably many Lipschitz functions $f_i \colon \mathbb{R}^d \to E$ such that

$$\mathcal{H}^d\left(\mathcal{M} \setminus \bigcup_{i\geq 0} f_i(\mathbb{R}^d)\right) = 0.$$

Let $\mathcal{H}^d_{\mathcal{M}}$ be the Hausdorff measure restricted to $\mathcal{M}$. Such a rectifiable set comes with a notion of tangent spaces $T_x\mathcal{M}$, defined $\mathcal{H}^d_{\mathcal{M}}(x)$-almost everywhere.

Consider a map $\theta \colon \mathcal{M} \to \mathbb{R}^+$, positive $\mathcal{H}^d_{\mathcal{M}}(x)$-almost everywhere. We can consider the varifold

$$V = \theta \cdot \mathcal{H}^d_{\mathcal{M}} \otimes \{\delta_{T_x\mathcal{M}}\},$$

where $\delta_{T_x\mathcal{M}}$ denotes the Dirac mass on $T_x\mathcal{M}$. In other words, for every continuous function $\phi \colon E \times \mathcal{G}_d(E) \to \mathbb{R}$ with compact support, V is defined as

$$\int_{(x,T)\in E\times\mathcal{G}_d(E)} \phi(x,T) \cdot \mathrm{d}V(x,T) = \int_{x\in E} \phi(x, T_x\mathcal{M}) \cdot \theta(x) \cdot \mathrm{d}\mathcal{H}^d_{\mathcal{M}}(x).$$

Such a varifold is called a *rectifiable d-varifold*. The map θ is called the *multiplicity* of V. If θ is constant and equal to 1, then V is called the *canonical varifold* associated to $\mathcal{M}$.

In particular, a d-dimensional $\mathcal{C}^1$-submanifold $\mathcal{M}$ is d-rectifiable, and the corresponding canonical varifold is $V = \mathcal{H}^d_{\mathcal{M}} \otimes \{\delta_{T_x\mathcal{M}}\}$, where $T_x\mathcal{M}$ denotes here the tangent space of $\mathcal{M}$ at x. More generally, a $\mathcal{C}^1$-immersed manifold is d-rectifiable, and the same construction applies.

Point cloud varifolds. Let X be a finite subset of $E \times \mathcal{G}_d(E) \times \mathbb{R}^+$. The *point cloud varifold* associated to X is

$$V = \sum_{(x,T,m)\in X} m \cdot \delta_x \otimes \delta_T.$$

The following theorem shows that point cloud varifolds allow to approximate rectifiable varifolds.

> **Theorem II.28 ([BLM17, Theorem 6.4]).** *If V is a rectifiable d-varifold, then there exists a sequence of point cloud varifolds $(V_i)_{i\geq 0}$ such that $\Delta_{1,1}(V_i, V) \to 0$.*

The precise statement of the theorem gives explicit bounds of the flat distance $\Delta_{1,1}(V_i, V)$.

In chapter Chapter IV, we prove related results, in a slightly different context. The measures we consider are not measures on $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^n)$, but on $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, where $\mathrm{M}(\mathbb{R}^n)$ is the space of $n \times n$ matrices. However, our goal is to approximate an actual rectifiable d-varifold, and the results are given in Wasserstein distance (see Remark IV.34).

II.6 Homology inference with Čech filtrations

Let X be a subset of $\mathbb{R}^n$. We assume that X is close to an unknown subset $\mathcal{M} \subset \mathbb{R}^n$. In the context of *homology inference*, one wants to estimate the homology groups $H_i(\mathcal{M})$ from X. To do so, one builds a persistence module $\mathbb{V}[X]$ from X—for instance, the persistence module associated to the Čech filtration on X. The persistence module $\mathbb{V}[X]$ is to be seen as an estimator of the corresponding persistence module $\mathbb{V}[\mathcal{M}]$ built from $\mathcal{M}$. Two questions arises:

- **Consistency:** how does the persistence module $\mathbb{V}[\mathcal{M}]$ reveals the homology of $\mathcal{M}$?
- **Stability:** how close are the persistence modules $\mathbb{V}[\mathcal{M}]$ and $\mathbb{V}[X]$?

A consistency result should be of the following form: *there exists an interval $I \subset \mathbb{R}^+$ on which $\mathbb{V}[\mathcal{M}]$ is constant and equal to the homology group $H_i(\mathcal{M})$.* In other words, for every $t \in I$, the vector space $\mathbb{V}^t[\mathcal{M}]$ is equal to $H_i(\mathcal{M})$, and for every $s, t \in I$ with $s \leq t$, the map $v_s^t \colon \mathbb{V}^s[\mathcal{M}] \to \mathbb{V}^t[\mathcal{M}]$ is the identity.

Besides, a stability result should read: *the interleaving distance between the persistence modules $\mathbb{V}[\mathcal{M}]$ and $\mathbb{V}[X]$ is upper bounded by some quantity.* This quantity depends on the proximity between $\mathcal{M}$ and X, which is usually quantified by the Hausdorff distance.

Such consistency and stability results shall be of concern throughout this thesis. In this section, we reformulate well-known results relative to the Čech filtrations, based on Section II.3. In Chapter III, we prove such results for DTM-filtrations, in Chapter IV for lifted filtrations, and in Chapter V for Čech bundle filtrations.

II.6.1 Consistency

Let $\mathcal{M}$ be a closed subset of $\mathbb{R}^n$. Consider its Čech filtration $V[\mathcal{M}] = (\mathcal{M}^t)_{t \geq 0}$, and the corresponding i^{th} homology persistence module $\mathbb{V}[\mathcal{M}] = (\mathbb{V}^t[\mathcal{M}])_{t \geq 0}$. By definition, $\mathbb{V}^t[\mathcal{M}] = H_i(\mathcal{M}^t)$ for all $t \in \mathbb{R}^+$. In this subsection, we describe where the homology group $H_i(\mathcal{M})$ can be read in the persistence module $\mathbb{V}[\mathcal{M}]$.

At the beginning of the filtration—at least for $t = 0$—, one has $\mathbb{V}^t[\mathcal{M}] = H_i(\mathcal{M}^t) = H_i(\mathcal{M})$. A quantity of interest is the largest t such that this is true:

$$t^\circ = \sup\big\{t \geq 0 \ \text{ s.t. } \mathbb{V}[\mathcal{M}] \text{ is constant on } [0, t]\big\}. \tag{II.5}$$

On the other hand, for t large enough, the thickening $\mathcal{M}^t$ has the homotopy type of a point, and $\mathbb{V}^t[\mathcal{M}]$ is trivial. The corresponding quantity is

$$t^\bullet = \inf\big\{t \geq 0 \ \text{ s.t. } \mathbb{V}[\mathcal{M}] \text{ is constant on } [t, +\infty)\big\}.$$

We shall assume that $H_i(\mathcal{M})$ is not trivial, so that we have $t^\circ \leq t^\bullet$. Note that the exponent of $t^\bullet$ refers to the disk, which has trivial homology, and the exponent of t° refers to the circle, which has non-trivial 1-homology.

Based on these quantities, one divides the behaviour of the persistence module $V[\mathcal{M}]$ in three phases:

- on the interval $[0, t^\circ)$, $\mathbb{V}[\mathcal{M}]$ is equal to the homology of $\mathcal{M}$,
- on the interval $(t^\circ, t^\bullet)$, $\mathbb{V}[\mathcal{M}]$ may not be equal to the homology of $\mathcal{M}$,
- on the interval $(t^\bullet, +\infty)$, $\mathbb{V}[\mathcal{M}]$ is trivial.

The more t° and $t^\bullet$ are close, the more the homology of $\mathcal{M}$ appears as a large feature of the barcode of $\mathbb{V}[\mathcal{M}]$. Figures II.6 and II.7 represent the barcodes of the Čech filtrations of two subsets $\mathcal{M}$. The first one is a circle, for which $t^\circ = t^\bullet$. The second one is a curve winding around a torus. One sees that the values t° are $t^\bullet$ are not close anymore. On the interval $(t^\circ, t^\bullet)$, the thickening $\mathcal{M}^t$ shows the homology of a torus, and then the homology of a filled torus (that is, a circle).

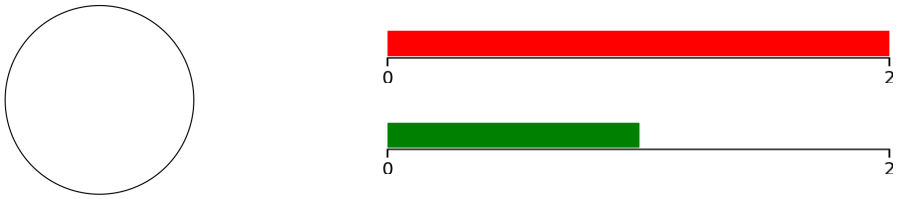


Figure II.6: Example for $t^\circ = t^\bullet$. **Left:** the set $\mathcal{M}$. **Right:** Barcode $(V[\mathcal{M}])$.

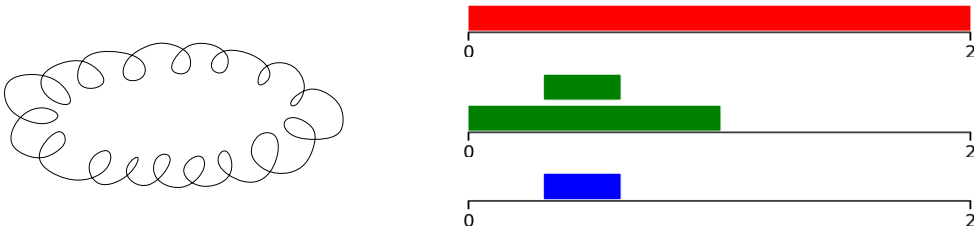


Figure II.7: Example for $t^\circ < t^\bullet$. **Left:** the set $\mathcal{M}$. **Right:** Barcode $(V[\mathcal{M}])$.

Lower bounds on t°. In order to bound t° with geometric quantities related to $\mathcal{M}$, we shall consider

$$t^\circ_{\text{Top}} = \sup\left\{t \geq 0 \text{ s.t. the inclusion } \mathcal{M} \hookrightarrow \mathcal{M}^t \text{ is a homotopy equivalence}\right\}.$$

By definition of the Čech filtration, it is clear that $t^\circ_{\text{Top}} \leq t^\circ$. Now, we have some bounds on the quantity t°_{Top}. Namely, if $\mathcal{M}$ is any closed subset of $\mathbb{R}^n$, then

$$\text{reach}(\mathcal{M}) \leq \text{reach}_\mu(\mathcal{M}) \leq \text{wfs}(\mathcal{M}) \leq t^\circ_{\text{Top}}$$

for any $\mu \in [0, 1)$, where $\text{reach}(\mathcal{M})$ and $\text{reach}_\mu(\mathcal{M})$ denote the reach and μ-reach of $\mathcal{M}$, and $\text{wfs}(\mathcal{M})$ its weak feature size (see Section II.3).

The unit circle $\mathcal{M} = \mathbb{S}_1 \subset \mathbb{R}^2$ is an example for which $\text{reach}(\mathcal{M}) = \text{reach}_\mu(\mathcal{M}) = \text{wfs}(\mathcal{M}) = t^\circ_{\text{Top}} = t^\circ$. However, it may happen that $\text{reach}(\mathcal{M}) < \text{reach}_\mu(\mathcal{M}) < \text{wfs}(\mathcal{M}) < t^\circ_{\text{Top}} < t^\circ$. We will see in Subsection V.3.4 an example for which $t^\circ_{\text{Top}} < t^\circ$.

Upper bounds on $t^\bullet$. Similarly, one studies the quantity $t^\bullet$ via

$$t^\bullet_{\mathrm{Top}} = \inf\left\{t \geq 0 \ \text{ s.t. } \forall u \geq t, \ \text{the inclusion } \mathcal{M}^t \hookrightarrow \mathcal{M}^u \text{ is a homotopy equivalence}\right\}.$$

The inequality $t^\bullet_{\mathrm{Top}} \geq t^\bullet$ holds. Moreover, if $\mathcal{M}$ is any closed subset of $\mathbb{R}^n$, we have

$$\mathrm{diam}\,(\mathcal{M}) \geq \mathrm{mini}\,(\mathcal{M}) \geq t^\bullet_{\mathrm{Top}},$$

where $\mathrm{diam}\,(\mathcal{M})$ and $\mathrm{mini}\,(\mathcal{M})$ denote the diameter of $\mathcal{M}$ and the radius of its minimum enclosing ball.

If $\mathcal{M}$ is not reduced to a point, we even have $\mathrm{diam}\,(\mathcal{M}) > \mathrm{mini}\,(\mathcal{M})$. Note that the following inequality holds: $\mathrm{mini}\,(\mathcal{M}) \geq \frac{1}{2}\mathrm{diam}\,(\mathcal{M})$. However, it may not be true that $\frac{1}{2}\mathrm{diam}\,(\mathcal{M}) \geq t^\bullet_{\mathrm{Top}}$.

As before, the unit circle $\mathcal{M} = \mathbb{S}_1 \subset \mathbb{R}^2$ is an example for which $\mathrm{mini}\,(\mathcal{M}) = t^\bullet_{\mathrm{Top}} = t^\bullet$. Nonetheless, it may happen that $\mathrm{mini}\,(\mathcal{M}) > t^\bullet_{\mathrm{Top}} > t^\bullet$.

II.6.2 Stability

As before, let $\mathcal{M}$ be a subset of $\mathbb{R}^n$. We observe a set $X \subset \mathbb{R}^n$. Suppose that $\mathcal{M}$ and X are at least ϵ-close in Hausdorff distance: $\mathrm{d_H}\,(\mathcal{M}, X) \leq \epsilon$. As we have seen in Subsection II.4.2, the Čech filtrations $(\mathcal{M}^t)_{t\geq 0}$ and $(X^t)_{t\geq 0}$ are ϵ-interleaved. The corresponding persistence modules $V\,[\mathcal{M}]$ and $V\,[X]$ are also ϵ-interleaved. As a consequence of the isometry theorem, their persistence barcodes $\mathrm{Barcode}\,(V\,[\mathcal{M}])$ and $\mathrm{Barcode}\,(V\,[X])$ are ϵ-close in bottleneck distance.

Let $[0, t^\circ)$ be the interval on which $V\,[M]$ is equal to the homology of $\mathcal{M}$ (defined in Equation (II.5)). If ϵ is small compared to t°, then the homology of $\mathcal{M}$ still appears as a large feature of $\mathrm{Barcode}\,(V\,[X])$. Namely, to a bar $[b, d)$ of $\mathrm{Barcode}\,(V\,[\mathcal{M}])$ corresponds a bar $[b', d')$ of $\mathrm{Barcode}\,(V\,[X])$ with $|b - b'| \leq \epsilon$ and $|d - d'| \leq \epsilon$. The situation is represented on Figure II.8.

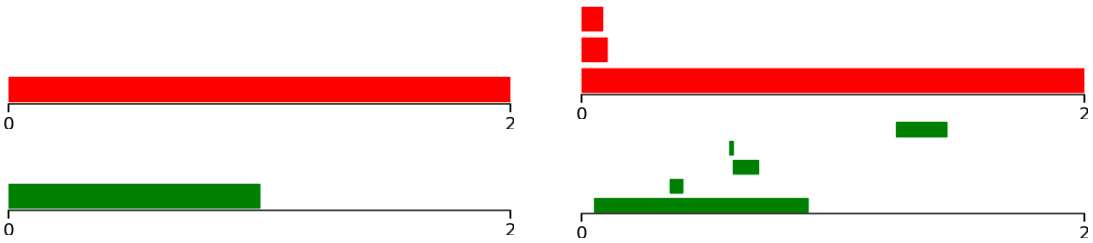


Figure II.8: Two ϵ-close barcodes in bottleneck distance.

One can state a more informative stability result. Suppose that $\mathcal{M}$ has positive reach. On the interval $[4\epsilon, \mathrm{reach}\,(\mathcal{M}) - 3\epsilon)$, the persistence module $V\,[X]$ is constant and equal to the homology of $\mathcal{M}$. This is a direct consequence of Theorem II.14. One identifies this interval on Figure II.9.

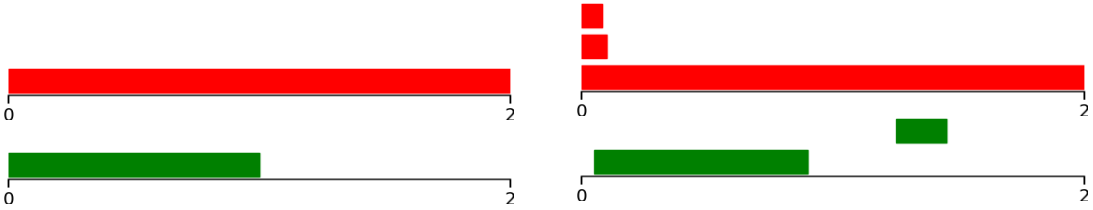


Figure II.9: Two ϵ-close barcodes which are equal on some large interval.

From an application perspective, it is sometimes enough to know that the filtrations $V[X]$ and $V[\mathcal{M}]$ are ϵ-interleaved. For instance, if one were to select the bars with large persistence, then the small bars in Figure II.8 would be ignored. However, if one wants to estimate an interval on which the thickenings $V^t[X]$ have the homotopy type of $\mathcal{M}$, then it is required that these sets are homotopy equivalent on a large interval, hence that the corresponding persistence modules are equal on a large interval as well, as in Figure II.9.

Contributions

III DTM-based filtrations

Abstract. Despite strong stability properties, the persistent homology of filtrations classically used in Topological Data Analysis, such as, e.g. the Čech or Vietoris-Rips filtrations, are very sensitive to the presence of outliers in the data from which they are computed. In this work, we introduce and study a new family of filtrations, the DTM-filtrations, built on top of point clouds in the Euclidean space which are more robust to noise and outliers. The approach adopted in this work relies on the notion of distance-to-measure functions, and extends some previous work on the approximation of such functions.

III.1 Weighted Čech filtrations 102
III.1.1 Definition 102
III.1.2 Stability 104
III.1.3 Weighted Vietoris-Rips filtrations 106
III.2 DTM-filtrations 108
III.2.1 The distance to measure (DTM) 109
III.2.2 DTM-filtrations 109
III.2.3 Stability when $p = 1$ 111
III.2.4 Stability when $p > 1$ 116
III.2.5 Proof of Lemma III.20 and Proposition III.21 120
III.2.6 Consequences under the standard assumption 123
III.3 Conclusion 126
III.A Supplementary results for Section III.1 126
III.B Supplementary results for Section III.2 127

Numerical experiments. A `Python` notebook, containing numerical illustrations, can be found at `https://github.com/raphaeltinarrage/DTM-Filtrations/blob/master/Demo.ipynb`.

Publication. This work has been published as an extended abstract and presented at the Symposium of Computational Geometry, June 2019 [ACG+18]. It has also been published in the proceedings of Abel Symposium, 2018 [ACG+20]. A journal-formatted version is available at `https://arxiv.org/abs/1811.04757`.

Co-authors. This work is a collaboration with Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi and Yuhei Umeda.

Organisation of the chapter. The weighted Čech and Vietoris-Rips filtrations are introduced in Section III.1, where their stability properties are established. The DTM-filtrations are introduced in Section III.2. Their main stability properties are stated in Theorems III.16 and III.22, and their relation with the sublevel set filtration of the DTM-functions is stated in Proposition III.18. For the clarity of the chapter, the proofs of several lemmas have been postponed to the appendices. Please refer to Subsection I.2.2 for an introduction to this chapter.

III.1 Weighted Čech filtrations

In order to define the DTM-filtrations, we go through an intermediate and more general construction, namely the weighted Čech filtrations. It generalizes the usual notion of Čech filtration of a subset of $\mathbb{R}^n$, and shares comparable regularity properties. Basic notions about filtrations and persistence modules can be found in Section II.4.

III.1.1 Definition

In the rest of the chapter, the Euclidean space $E = \mathbb{R}^n$, the index set $T = \mathbb{R}^+$ and a real number $p \geq 1$ or $p = +\infty$ are fixed. Consider $X \subseteq E$ and $f : X \to \mathbb{R}^+$. For every $x \in X$ and $t \in T$, we define

$$r_x(t) = \begin{cases} -\infty & \text{if } t < f(x), \\ \left(t^p - f(x)^p\right)^{\frac{1}{p}} & \text{otherwise.} \end{cases}$$

We denote by $\overline{\mathcal{B}}_f(x,t) = \overline{\mathcal{B}}\left(x, r_x(t)\right)$ the closed Euclidean ball of center x and radius $r_x(t)$. By convention, a Euclidean ball of radius $-\infty$ is the empty set. For $p = \infty$, we also define

$$r_x(t) = \begin{cases} -\infty & \text{if } t < f(x), \\ t & \text{otherwise,} \end{cases}$$

and the balls $\overline{\mathcal{B}}_f(x,t) = \overline{\mathcal{B}}\left(x, r_x(t)\right)$. Some of these radius functions are represented in Figure III.1.

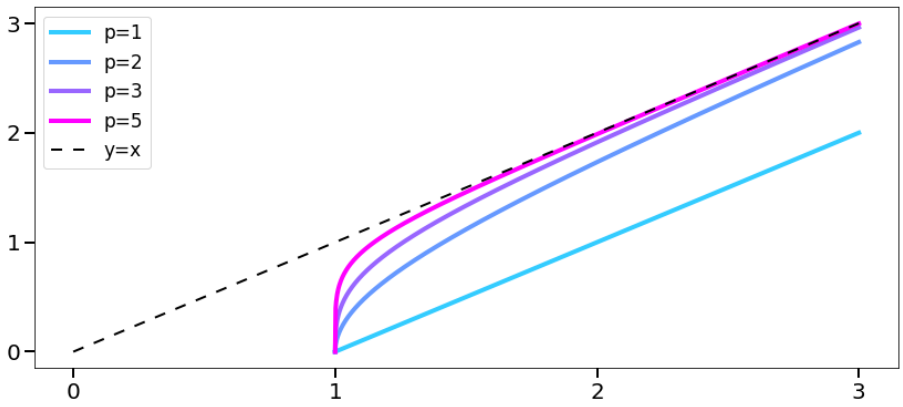


Figure III.1: Graph of $t \mapsto r_x(t)$ for $f(x) = 1$ and several values of p.

Definition III.1. Let $X \subseteq E$ and $f\colon X \to \mathbb{R}^+$. For every $t \in T$, we define the following set:

$$V^t[X,f] = \bigcup_{x\in X} \overline{\mathcal{B}}_f(x,t).$$

The family $V[X,f] = (V^t[X,f])_{t\geq 0}$ is a filtration of E. It is called the *weighted Čech filtration with parameters* (X,f,p). We denote by $\mathbb{V}[X,f]$ its persistence (singular) homology module.

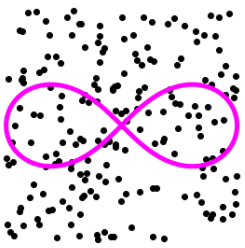
Note that $V[X,f]$ and $\mathbb{V}[X,f]$ depend on the parameter p, that is not made explicit in the notation.

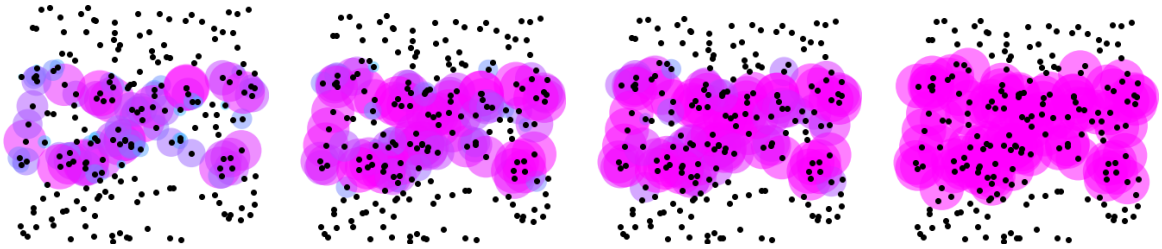
Introduce $\mathcal{V}^t[X,f] = \{\overline{\mathcal{B}}_f(x,t)\}_{x\in X}$. It is a cover of $V^t[X,f]$ by closed Euclidean balls. Let $\mathcal{N}(\mathcal{V}^t[X,f])$ be the nerve of the cover $\mathcal{V}^t[X,f]$. It is a simplicial complex over the vertex set X. The family $\mathcal{N}(\mathcal{V}[X,f]) = (\mathcal{N}(\mathcal{V}^t[X,f]))_{t\geq 0}$ is a filtered simplicial complex. We denote by $\mathbb{V}_{\mathcal{N}}[X,f]$ its persistence (simplicial) homology module. As a consequence of the persistent nerve lemma (Lemma II.18), $\mathbb{V}[X,f]$ and $\mathbb{V}_{\mathcal{N}}[X,f]$ are isomorphic persistence modules.

When $f = 0$, $V[X,f]$ does not depend on $p \geq 1$, and is the Čech set filtration associated to X. We denote it by $V[X,0]$. The corresponding filtered simplicial complex, $\mathcal{N}(\mathcal{V}[X,0])$, is known as the Čech simplicial filtration of X (see Subsection II.4.2).

When $p = 2$, the filtration value of $y \in E$, i.e. the infimum of the t such that $y \in V^t[X,f]$, is called the power distance of y associated to the weighted set (X,f) [BCOS16, Definition 4.1]. The filtration $V[X,f]$ is called the weighted Čech filtration [BCOS16, Definition 5.1].

Example III.2. Consider the point cloud X drawn on the right (black). It is a 200-sample of the uniform distribution on $[-1,1]^2 \subseteq \mathbb{R}^2$. We choose f to be the distance function to the lemniscate of Bernoulli (magenta). Let $t = 0.2$. Figure III.2 represents the sets $V^t[X,f]$ for several values of p. The balls are colored differently according to their radius.





$p = 1$ $p = 2$ $p = 3$ $p = \infty$

Figure III.2: The sets $V^t[X,f]$ for $t = 0.2$ and several values of p.

The following proposition states the regularity of the persistence module $\mathbb{V}[X,f]$.

Proposition III.3. *If $X \subseteq E$ is finite and f is any function, then $\mathbb{V}[X, f]$ is a pointwise finite-dimensional persistence module.*

More generally, if X is a bounded subset of E and f is any function, then $\mathbb{V}[X, f]$ is q-tame.

Proof. First, suppose that X is finite. Then $\mathcal{N}(\mathcal{V}[X, f])$ is a filtration of a finite simplicial complex, and thus $\mathbb{V}_{\mathcal{N}}[X, f]$ is pointwise finite-dimensional. It is also the case for $\mathbb{V}[X, f]$ since it is isomorphic to $\mathbb{V}_{\mathcal{N}}[X, f]$.

Secondly, suppose that X is bounded. Consider the 'filtration value' function:

$$\begin{array}{rccl} t_X & : & E & \longrightarrow & \mathbb{R}^+ \\ & & y & \longmapsto & \inf\left\{t \in \mathbb{R}^+, \exists x \in X, y \in \overline{\mathcal{B}}_f(x,t)\right\} \end{array}$$

For every $y \in E$, $x \in X$ and $t \geq 0$, the assertion $y \in \overline{\mathcal{B}}_f(x,t)$ is equivalent to $\left(\|x-y\|^p + f(x)^p\right)^{\frac{1}{p}} \leq t$. Therefore the function t_X can be written as follows:

$$t_X(y) = \inf\left\{\left(\|x-y\|^p + f(x)^p\right)^{\frac{1}{p}}, x \in X\right\}.$$

It is 1-Lipschitz as it is the infimum of the set of the 1-Lipschitz functions $y \mapsto \left(\|x-y\|^p + f(x)^p\right)^{\frac{1}{p}}$. It is also proper as X is bounded.

Let $\widetilde{V}$ be the filtration of E defined for all $t \geq 0$ by $\widetilde{V}^t = t_X^{-1}(]-\infty, t])$. Let $\widetilde{\mathbb{V}}$ be its persistent homology module. The last two properties of t_X (continuous and proper) imply that $\widetilde{\mathbb{V}}$ is q-tame ([CdSGO16], Theorem 2.22).

Notice that, since X may not be compact, $V^t[X, f]$ may not be equal to $\widetilde{V}^t$. However, it follows from the definition of t_X that $V[X, f]$ and $\widetilde{V}$ are ϵ-interleaved for every $\epsilon > 0$. Therefore $\mathbb{V}[X, f]$ also is q-tame.

III.1.2 Stability

We still consider a subset $X \subseteq E$ and a function $f\colon X \to \mathbb{R}^+$. Using the fact that two ϵ-interleaved filtrations induce ϵ-interleaved persistence modules, the stability results for the filtration $V[X, f]$ of this subsection immediately translate as stability results for the persistence module $\mathbb{V}[X, f]$.

The following proposition states the stability of the filtration $V[X, f]$ with respect to f.

Proposition III.4. *Let $g\colon X \to \mathbb{R}^+$ be a function such that $\sup_{x \in X} |f(x) - g(x)| \leq \epsilon$. Then the filtrations $V[X, f]$ and $V[X, g]$ are ϵ-interleaved.*

Proof. By symmetry, it suffices to show that, for every $t \geq 0$, $V^t[X, f] \subseteq V^{t+\epsilon}[X, g]$.
Let $t \geq 0$. Choose $y \in V^t[X, f]$, and $x \in X$ such that $y \in \overline{\mathcal{B}}_f(x,t)$, i.e. $\left(\|x-y\|^p + f(x)^p\right)^{\frac{1}{p}} \leq t$. Let us prove that $y \in \overline{\mathcal{B}}_g(x, t+\epsilon)$, i.e. $\left(\|x-y\|^p + g(x)^p\right)^{\frac{1}{p}} \leq t + \epsilon$.

From $g(x) \leq f(x)+\epsilon$, we obtain $\left(\|x-y\|^p+g(x)^p\right)^{\frac{1}{p}} \leq \left(\|x-y\|^p+(f(x)+\epsilon)^p\right)^{\frac{1}{p}}$. Now, consider the function $\eta \mapsto (\|x-y\|^p+(f(x)+\eta)^p)^{\frac{1}{p}}$. Its derivative is $\eta \mapsto \left(\frac{f(x)+\eta}{\left(\|x-y\|^p+(f(x)+\eta)^p\right)^{\frac{1}{p}}}\right)^{p-1}$. It is consequently 1-Lipschitz on $\mathbb{R}^+$. The Lipschitz property implies that

$$\left(\|x-y\|^p+\left(f(x)+\epsilon\right)^p\right)^{\frac{1}{p}} \leq \left(\|x-y\|^p+f(x)^p\right)^{\frac{1}{p}}+\epsilon.$$

Hence

$$\begin{aligned}\left(\|x-y\|^p+g(x)^p\right)^{\frac{1}{p}} &\leq \left(\|x-y\|^p+(f(x)+\epsilon)^p\right)^{\frac{1}{p}} \\ &\leq \left(\|x-y\|^p+f(x)^p\right)^{\frac{1}{p}}+\epsilon \leq t+\epsilon.\end{aligned}$$

The following proposition states the stability of $V[X,f]$ with respect to X. It generalizes [BCOS16, Proposition 4.3] (case $p=2$).

Proposition III.5. *Let $Y \subseteq E$ and suppose that $f\colon X \cup Y \to \mathbb{R}^+$ is c-Lipschitz, $c \geq 0$. Suppose that X and Y are compact and that the Hausdorff distance satisfies $\mathrm{d_H}(X,Y) \leq \epsilon$. Then the filtrations $V[X,f]$ and $V[Y,f]$ are k-interleaved with $k = (1+c^p)^{\frac{1}{p}}\epsilon$.*

Proof. It suffices to show that for every $t \geq 0$, $V^t[X,f] \subseteq V^{t+k}[Y,f]$.
Let $t \geq 0$. Choose $z \in V^t[X,f]$, and $x \in X$ such that $z \in \overline{\mathcal{B}}_f(x,t)$, i.e. $\|x-z\| \leq r_x(t)$. From the hypothesis $\mathrm{d_H}(X,Y) \leq \epsilon$, there exists $y \in Y$ such that $\|y-x\| \leq \epsilon$. Let us prove that $z \in \overline{\mathcal{B}}_f(t,t+k)$, i.e. $\|z-y\| \leq r_y(t+k)$.

By the triangle inequality, we have $\|z-y\| \leq \|z-x\|+\|x-y\| \leq r_x(t)+\epsilon$. It is enough to show that $r_x(t)+\epsilon \leq r_y(t+k)$, i.e.

$$\underbrace{\left((t+k)^p-f(y)^p\right)^{\frac{1}{p}}}_{r_y(t+k)} - \underbrace{\left(t^p-f(x)^p\right)^{\frac{1}{p}}}_{r_x(t)} \geq \epsilon.$$

The left-hand side of this expression is decreasing in $f(y)$. Moreover, since f is c-Lipschitz, $f(y)$ is at most $f(x)+c\epsilon$. Therefore:

$$\begin{aligned}&((t+k)^p-f(y)^p)^{\frac{1}{p}} &&- &&(t^p-f(x)^p)^{\frac{1}{p}} \\ \geq\ &((t+k)^p-(f(x)+c\epsilon)^p)^{\frac{1}{p}} &&- &&(t^p-f(x)^p)^{\frac{1}{p}}.\end{aligned}$$

It is now enough to prove that this last expression is not less than ϵ, which is the content of Lemma III.26.

Notice that the bounds in Proposition III.4 and III.5 are tight. In the first case, consider for example $E=\mathbb{R}$, the set $X=\{0\}$, and the functions $f=0$ and $g=\epsilon$. For every $t<\epsilon$, we have $V^t[Y,f]=\emptyset$, while $V^t[X,f] \neq \emptyset$. Regarding the second proposition, consider $E=\mathbb{R}$, $f\colon x \mapsto cx$, $X=\{0\}$ and $Y=\{\epsilon\}$. We have, for every $t \geq 0$, $V^t[X,f] = \overline{\mathcal{B}}(0,t)$ and $V^t[Y,f] = \overline{\mathcal{B}}\left(\epsilon, t^p-(c\epsilon)^p)^{\frac{1}{p}}\right)$. For every

$t < \epsilon(1+c^p)^{\frac{1}{p}}$, we have $(t^p - (c\epsilon)^p)^{\frac{1}{p}} < \epsilon$, hence $0 \notin V^t[Y,f]$. In comparison, we have $\forall t \geq 0$, $0 \in V^t[X,f]$.

When considering data with outliers, the observed set X may be very distant from the underlying signal Y in Hausdorff distance. Therefore, the tight bound in Proposition III.5 may be unsatisfactory. Moreover, a usual choice of f would be d_X, the distance function to X. But the bound in Proposition III.4 then becomes $\|\mathrm{d}_X - \mathrm{d}_Y\|_\infty = \mathrm{d_H}(X,Y)$. We address this issue in Section III.2 by considering an outliers-robust function f, the so-called distance-to-measure function (DTM).

III.1.3 Weighted Vietoris-Rips filtrations

Rather than computing the persistence of the Čech simplicial filtration of a point cloud $X \subseteq E$, one sometimes consider the corresponding Vietoris-Rips filtration, which is usually easier to compute.

If G is a graph with vertex set X, its corresponding clique complex is the simplicial complex over X consisting of the sets of vertices of cliques of G. If S is a simplicial complex, its corresponding flag complex is the clique complex of its 1-skeleton. We remind the reader that $\mathcal{N}(\mathcal{V}^t[X,f])$ denotes the nerve of $\mathcal{V}^t[X,f]$, where $\mathcal{V}^t[X,f]$ is the cover $\{\overline{\mathcal{B}}_f(x,t)\}_{x\in X}$ of $V^t[X,f]$.

Definition III.6. We denote by $\mathrm{Rips}(\mathcal{V}^t[X,f])$ the flag complex of $\mathcal{N}(\mathcal{V}^t[X,f])$, and by $\mathrm{Rips}(\mathcal{V}[X,f])$ the corresponding filtered simplicial complex. It is called the *weighted Rips complex with parameters* (X,f,p).

The following proposition states that the filtered simplicial complexes $\mathcal{N}(\mathcal{V}[X,f])$ and $\mathrm{Rips}(\mathcal{V}[X,f])$ are 2-interleaved multiplicatively, generalizing the classical case of the Čech and Vietoris-Rips filtrations (case $f=0$).

Proposition III.7. *For every $t \geq 0$,*

$$\mathcal{N}(\mathcal{V}^t[X,f]) \subseteq \mathrm{Rips}(\mathcal{V}^t[X,f]) \subseteq \mathcal{N}(\mathcal{V}^{2t}[X,f])$$

Proof. Let $t \geq 0$. The first inclusion follows from that $\mathrm{Rips}(\mathcal{V}^t[X,f]))$ is the clique complex of $\mathcal{N}(\mathcal{V}^t[X,f])$. To prove the second one, choose a simplex $\omega \in \mathrm{Rips}(\mathcal{V}^t[X,f]))$. It means that for every $x,y \in \omega$, $\overline{\mathcal{B}}_f(x,t) \cap \overline{\mathcal{B}}_f(y,t) \neq \emptyset$, i.e. $\overline{\mathcal{B}}(x,r_x(t)) \cap \overline{\mathcal{B}}(y,r_y(t)) \neq \emptyset$. We have to prove that $\omega \in \mathcal{N}(\mathcal{V}^{2t}[X,f])$, i.e. $\bigcap_{x\in\omega} \overline{\mathcal{B}}(x,r_x(2t)) \neq \emptyset$.

For every $x \in \omega$, one has $r_x(2t) \geq 2r_x(t)$. Indeed,

$$\begin{aligned} r_x(2t) &= \big((2t)^p - f(x)^p\big)^{\frac{1}{p}} \\ &= 2\big(t^p - (\frac{f(x)}{2})^p\big)^{\frac{1}{p}} \\ &\geq 2\big(t^p - f(x)^p\big)^{\frac{1}{p}} = 2r_x(t) \end{aligned}$$

Using the fact that doubling the radius of pairwise intersecting balls is enough to make them intersect globally, we obtain that $\omega \in \mathcal{N}(\mathcal{V}^{2t}[X,f])$.

Using Theorem 3.1 of [BLM+19a], the multiplicative interleaving $\mathrm{Rips}(\mathcal{V}^t[X,f]) \subseteq \mathcal{N}(\mathcal{V}^{2t}[X,f])$ can be improved to $\mathrm{Rips}(\mathcal{V}^t[X,f]) \subseteq \mathcal{N}(\mathcal{V}^{ct}[X,f])$, where $c = \sqrt{\frac{2n}{n+1}}$ and n is the dimension of the ambient space $E = \mathbb{R}^n$.

Note that weighted Rips filtration shares the same stability properties as the weighted Čech filtration. Indeed, the proofs of Proposition III.4 and III.5 immediately extend to this case.

In order to compute the flag complex $\mathrm{Rips}(\mathcal{V}^t[X,f])$, it is enough to know the filtration values of its 0- and 1-simplices. The following proposition describes these values.

Proposition III.8. *Let $p < +\infty$. The filtration value of a vertex $x \in X$ is given by $t_X(\{x\}) = f(x)$.*
The filtration value of an edge $\{x,y\} \subseteq E$ is given by

$$t_X(\{x,y\}) = \begin{cases} \max\{f(x),f(y)\} & \text{if } \|x-y\| \leq |f(x)^p - f(y)^p|^{\frac{1}{p}}, \\ t & \text{otherwise,} \end{cases}$$

where t is the only positive root of

$$\|x-y\| = (t^p - f(x)^p)^{\frac{1}{p}} + (t^p - f(y)^p)^{\frac{1}{p}} \tag{III.1}$$

When $\|x-y\| \geq |f(x)^p - f(y)^p|^{\frac{1}{p}}$, the positive root of Equation (III.1) does not always admit a closed form. We give some particular cases for which it can be computed.

- For $p=1$, the root is $t_X(\{x,y\}) = \frac{f(x)+f(y)+\|x-y\|}{2}$,
- for $p=2$, it is $t_X(\{x,y\}) = \frac{\sqrt{\left((f(x)+f(y))^2+\|x-y\|^2\right)\left((f(x)-f(y))^2+\|x-y\|^2\right)}}{2\|x-y\|}$,
- for $p=\infty$, the condition reads $\|x-y\| \geq \max\{f(x),f(y)\}$, and the root is $t_X(\{x,y\}) = \frac{\|x-y\|}{2}$. In either case, $t_X(\{x,y\}) = \max\{f(x),f(y),\frac{\|x-y\|}{2}\}$.

Proof. The filtration value of a vertex $x \in X$ is, by definition of the nerve, $t_X(\{x\}) = \inf\{s \in T, \overline{\mathcal{B}}_f(x,s) \neq \emptyset\}$. It is equal to $f(x)$.

Also by definition, the filtration value of an edge $\{x,y\}$, with $x,y \in X$ and $x \neq y$, is given by

$$t_X(\{x,y\}) = \inf\{s \in \mathbb{R}, \overline{\mathcal{B}}_f(x,s) \cap \overline{\mathcal{B}}_f(y,s) \neq \emptyset\}$$

Two cases may occur: the balls $\overline{\mathcal{B}}_f(x, t_X(\{x,y\}))$ and $\overline{\mathcal{B}}_f(x, t_X(\{x,y\}))$ have both positive radius, or one of these is a singleton. In the last case, $t_X(\{x,y\}) = \max\{f(x),f(y)\}$. In the first case, we have $\|x-y\| = r_x(t) + r_y(t)$, i.e. $\|x-y\| = (t^p - f(x)^p)^{\frac{1}{p}} + (t^p - f(y)^p)^{\frac{1}{p}}$. Notice that Equation (III.1) admits

only one solution since the function $t \mapsto (t^p - f(x)^p)^{\frac{1}{p}} + (t^p - f(y)^p)^{\frac{1}{p}}$ is strictly increasing on the interval $[\max\{f(x), f(y)\}, +\infty)$.

We close this subsection by discussing the influence of p on the weighted Čech and Rips filtrations. Let $\mathrm{Diagram}_0(\mathcal{N}(\mathcal{V}[X, f, p]))$ be the persistence diagram of the 0th-homology of $\mathcal{N}(\mathcal{V}[X, f, p])$. We say that a point (b, d) of $\mathrm{Diagram}_0(\mathcal{V}[X, f, p])$ is non-trivial if $b \neq d$. Let $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ be the persistence diagram of the 0th-homology of $\mathrm{Rips}(\mathcal{V}[X, f, p])$. Note that $\mathrm{Diagram}_0(\mathcal{N}(\mathcal{V}[X, f, p])) = \mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ since the corresponding filtrations share the same 1-skeleton.

Proposition III.9. *The number of non-trivial points in* $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ *is non-increasing with respect to* $p \in [1, +\infty)$. *The same property holds for* $\mathrm{Diagram}_0(\mathcal{N}(\mathcal{V}[X, f, p]))$.

Proof. The number of points in $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ is equal to the cardinal of X. Any $p \geq 1$ being fixed, we can pair every $x \in X$ with some edge $\{y, z\} \in \mathrm{Rips}(\mathcal{V}[X, f, p])$ such that the points of $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ are of the form $\big(t_X(\{x\}), t_X(\{y, z\})\big)$.

Notice that the filtration values of the points in X do not depend on p: for all $p \geq 1$ and $x \in X$, $t_X(\{x\}) = f(x)$. Moreover, the filtration values of the edges in $\mathrm{Rips}(\mathcal{V}[X, f, p])$ are non-increasing with respect to p. Indeed, for all $\{y, z\} \in \mathrm{Rips}(\mathcal{V}[X, f, p])$ with $y \neq z$, according to Proposition III.8, the filtration value $t_X(\{y, z\})$ is either $\max\{f(x), f(y)\}$ if $\|x - y\| \leq |f(x)^p - f(y)^p|^{\frac{1}{p}}$, or is the only positive root of Equation (III.1) otherwise. Note that the positive root of Equation (III.1) is greater than $\max\{f(x), f(y)\}$ and decreasing in p. Besides, the term $|f(x)^p - f(y)^p|^{\frac{1}{p}}$ is non-decreasing in p.

These two facts ensure that for every $x \in X$, the point of $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ created by x has an ordinate which is non-increasing with respect to p. In particular, the number of non-trivial points in $\mathrm{Diagram}_0(\mathrm{Rips}(\mathcal{V}[X, f, p]))$ is non-increasing with respect to p.

Figure III.7 in Subsection III.2.4 illustrates the previous proposition in the case of the DTM-filtrations. Greater values of p lead to sparser 0^{th}-homology diagrams.

Now, consider $k > 0$, and let $\mathrm{Diagram}_k(\mathcal{N}(\mathcal{V}[X, f, p]))$ be the persistence diagram of the kth-homology of $\mathcal{N}(\mathcal{V}[X, f, p])$. In this case, one can easily build examples showing that the number of non-trivial points of $\mathrm{Diagram}_k(\mathcal{N}(\mathcal{V}[X, f, p]))$ does not have to be non-increasing with respect to p. The same holds for the Vietoris-Rips version $\mathrm{Diagram}_k(\mathrm{Rips}(\mathcal{V}[X, f, p]))$.

III.2 DTM-filtrations

The results of the previous section suggest that, in order to construct a weighted Čech filtration $V[X, f]$ that is robust to outliers, it is necessary to choose a function f that depends on X and that is itself robust to outliers. The so-called distance-

to-measure function (DTM) satisfies such properties, motivating the introduction of the DTM-filtrations in this section.

III.2.1 The distance to measure (DTM)

Let μ be a measure on E and $m \in (0,1)$ a parameter. We remind the reader that the DTM with parameters μ and m is denoted $\mathrm{d}_{\mu,m}$ or d_μ, and has been defined in Subsection II.5.2. Besides, the 2-Wasserstein distance $\mathrm{W}_2(\cdot,\cdot)$ has been defined in Subsection II.5.1. We recall here two properties that shall be useful in what follows. The first one is a regularity property.

Proposition III.10 ([CCSM11], Corollary 3.7). *For every probability measure μ and $m \in [0,1)$, $\mathrm{d}_{\mu,m}$ is 1-Lipschitz.*

The second proposition states the stability of the DTM in the Wasserstein metric.

Proposition III.11 ([CCSM11], Theorem 3.5). *Let μ, ν be two probability measures, and $m \in (0,1)$. Then*

$$\|\mathrm{d}_{\mu,m} - \mathrm{d}_{\nu,m}\|_\infty \leq m^{-\frac{1}{2}} \mathrm{W}_2(\mu,\nu).$$

Notice that for every $x \in E$, $\mathrm{d}_\mu(x)$ is not lower than the distance from x to $\mathrm{supp}(\mu)$, the support of μ:

$$\mathrm{d}_\mu(x) \geq \mathrm{dist}(x, \mathrm{supp}(\mu)).$$

This remark, along with the Propositions III.10 and III.11, are the only properties of the DTM that will be used to prove the results in the rest of the chapter.

III.2.2 DTM-filtrations

In the following, the two parameters $p \in [1,+\infty]$ and $m \in (0,1)$ are fixed.

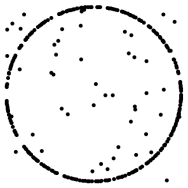
Definition III.12. Let $X \subseteq E$ be a finite point cloud, μ_X the empirical measure of X, and d_{μ_X} the corresponding DTM of parameter m. The weighted Čech filtration $V[X, \mathrm{d}_{\mu_X}]$, as defined in Definition III.1, is called the *DTM-filtration associated with the parameters* (X,m,p). It is denoted by $W[X,m,p]$, or simply $W[X]$. The corresponding persistence module is denoted by $\mathbb{W}[X,m,p]$ or $\mathbb{W}[X]$.

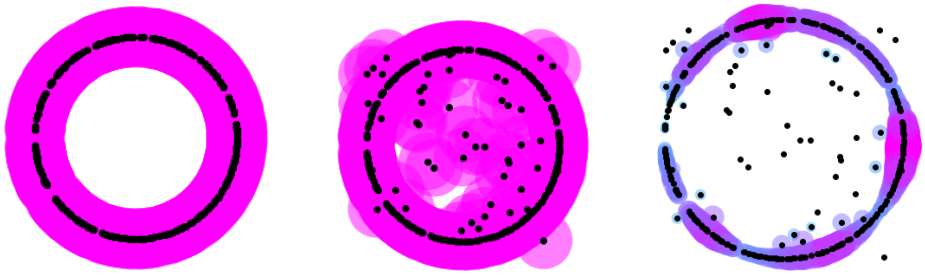
More generally, let μ be any probability measure. The weighted Čech filtration $V[\mathrm{supp}(\mu), \mathrm{d}_\mu]$ is called the *DTM-filtration associated with the parameters* (μ,m,p). It is denoted by $W[\mu,m,p]$, and the corresponding persistence module is denoted $\mathbb{W}[\mu,m,p]$.

Let $\mathcal{W}^t[X] = \mathcal{V}^t[X, \mathrm{d}_{\mu_X}]$ denote the cover of $W^t[X]$ as defined in section III.1, and let $\mathcal{N}(\mathcal{W}^t[X])$ be its nerve. The family $\mathcal{N}(\mathcal{W}[X])) = (\mathcal{N}(\mathcal{W}^t[X]))_{t\geq 0}$ is a filtered simplicial complex, and its persistent (simplicial) homology module is denoted by $\mathbb{W}_{\mathcal{N}}[X]$. By the persistent nerve lemma, the persistence modules $\mathbb{W}[X]$ and $\mathbb{W}_{\mathcal{N}}[X]$ are isomorphic.

As in Definition III.6, $\mathrm{Rips}(\mathcal{W}^t[X])$ denotes the flag complex of $\mathcal{N}(\mathcal{W}^t[X])$, and $\mathrm{Rips}(\mathcal{W}[X])$ the corresponding filtered simplicial complex.

Example III.13. Consider the point cloud X drawn on the right. It is the union of $\widetilde{X}$ and Γ, where $\widetilde{X}$ is a 50-sample of the uniform distribution on $[-1,1]^2 \subseteq \mathbb{R}^2$, and Γ is a 300-sample of the uniform distribution on the unit circle. We consider the weighted Čech filtrations $V[\Gamma,0]$ and $V[X,0]$, and the DTM-filtration $W[X]$, for $p=1$ and $m=0.1$. They are represented in Figure III.3 for the value $t=0.3$.





$V^t[\Gamma,0]$ $V^t[X,0]$ $W^t[X]$

Figure III.3: The sets $V^t[\Gamma,0]$, $V^t[X,0]$ and $W^t[X]$ for $p=1$, $m=0.1$ and $t=0.3$.

Because of the outliers $\widetilde{X}$, the value of t from which the sets $V^t[X,0]$ are contractible is small. On the other hand, we observe that the set $W^t[X]$ does not suffer too much from the presence of outliers.

We plot in Figure III.4 the persistence diagrams of the persistence modules associated to $\mathrm{Rips}(\mathcal{V}[\Gamma,0])$, $\mathrm{Rips}(\mathcal{V}[X,0])$ and $\mathrm{Rips}(\mathcal{W}[X])$ ($p=1$, $m=0.1$).

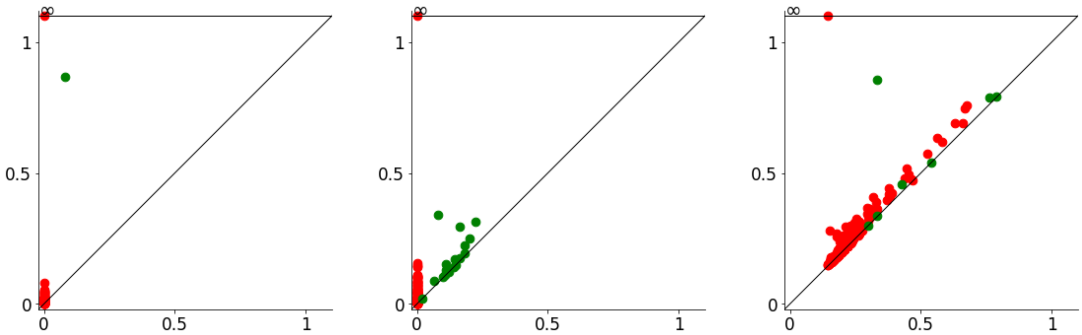


Diagram $(\mathrm{Rips}(\mathcal{V}[\Gamma,0]))$ Diagram $(\mathrm{Rips}(\mathcal{V}[X,0]))$ Diagram $(\mathrm{Rips}(\mathcal{W}[X]))$

Figure III.4: Persistence diagrams of some simplicial filtrations. Points in red (resp. green) represent the persistent homology in dimension 0 (resp. 1).

Observe that Diagram $(\mathrm{Rips}(\mathcal{V}[\Gamma,0]))$ and Diagram $(\mathrm{Rips}(\mathcal{W}[X]))$ appear to be close to each other, while Diagram $(\mathrm{Rips}(\mathcal{V}[X,0]))$ does not.

Applying the results of Section III.1, we immediately obtain the following proposition.

Proposition III.14. *Consider two measures μ, ν on E with compact supports X and Y. Then*

$$\mathrm{d_i}\left(W[\mu], W[\nu]\right) \leq m^{-\frac{1}{2}} \mathrm{W}_2\left(\mu, \nu\right) + 2^{\frac{1}{p}} \mathrm{d_H}\left(X, Y\right).$$

If X and Y are finite subsets of E, using the empirical measures $\mu = \mu_X$ and $\nu = \nu_Y$, we obtain

$$\mathrm{d_i}\left(W[X], W[Y]\right) \leq m^{-\frac{1}{2}} \mathrm{W}_2\left(X, Y\right) + 2^{\frac{1}{p}} \mathrm{d_H}\left(X, Y\right).$$

Proof. Remind that $W[\mu] = V[X, \mathrm{d}_\mu]$ and $W[\nu] = V[X, \mathrm{d}_\nu]$. We use the triangle inequality for the interleaving distance:

$$\mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[Y, \mathrm{d}_\nu]\right) \leq \underbrace{\mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[Y, \mathrm{d}_\mu]\right)}_{(1)} + \underbrace{\mathrm{d_i}\left(V[Y, \mathrm{d}_\mu], V[Y, \mathrm{d}_\nu]\right)}_{(2)}.$$

(1): From Proposition III.5, we have $\mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[Y, d_\mu]\right) \leq (1+c^p)^{\frac{1}{p}} \mathrm{d_H}\left(X, Y\right)$, where c is the Lipschitz constant of d_μ. According to Proposition III.10, $c = 1$. We obtain $\mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[Y, \mathrm{d}_\mu]\right) \leq 2^{\frac{1}{p}} \mathrm{d_H}\left(X, Y\right)$.
(2): From Proposition III.4, we have $\mathrm{d_i}\left(V[Y, \mathrm{d}_\mu], V[Y, \mathrm{d}_\nu]\right) \leq \|\mathrm{d}_\mu - \mathrm{d}_\nu\|_\infty$. According to Proposition III.11, $\|\mathrm{d}_\mu - \mathrm{d}_\nu\|_\infty \leq m^{-\frac{1}{2}} \mathrm{W}_2\left(\mu, \nu\right)$.

The second point of the proposition follow from the definition of the DTM-filtrations: $W[X] = V[X, \mathrm{d}_{\mu_X}]$ and $W[Y] = V[Y, \mathrm{d}_{\mu_Y}]$.

Note that this stability result is worse than the stability of the usual Čech filtrations, which only involves the Hausdorff distance. However, the term $\mathrm{W}_2\left(X, Y\right)$ is inevitable, as shown in the following example.

Let $E = \mathbb{R}$, and $\epsilon \in (0, 1)$. Define $\mu = \epsilon\delta_0 + (1-\epsilon)\delta_1$, and $\nu = (1-\epsilon)\delta_0 + \epsilon\delta_1$. We have $X = \mathrm{supp}(\mu) = \mathrm{supp}(\nu) = Y$. If $\epsilon \leq m \leq 1 - \epsilon$, then $\mathrm{d}_\nu(0) = 0$, while $\mathrm{d}_\mu(0) = \sqrt{1 - \frac{\epsilon}{m}}$. We deduce that $\mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[Y, \mathrm{d}_\nu]\right) \geq \mathrm{d}_\mu(0) - \mathrm{d}_\nu(0) = \sqrt{1 - \frac{\epsilon}{m}}$.

In comparison, the usual Čech filtrations $V[X, 0]$ and $V[Y, 0]$ are equal and does not depend on μ and ν. In this case, it would be more robust to consider these usual Čech filtrations. Now, in the case where the Hausdorff distance $\mathrm{d_H}\left(X, Y\right)$ is large, the usual Čech filtrations may be very distant. However, the DTM-filtrations may still be close, as we discuss in the next subsection.

III.2.3 Stability when $p = 1$

We first consider the case $p = 1$, for which the proofs are simpler and the results stronger.

We fix $m \in (0, 1)$. If μ is a probability measure on E with compact support $\mathrm{supp}(\mu)$, we define

$$c(\mu, m) = \sup_{\mathrm{supp}(\mu)} (\mathrm{d}_{\mu,m}). \tag{III.2}$$

If $\mu = \mu_\Gamma$ is the empirical measure of a finite set $\Gamma \subseteq E$, we denote it $c(\Gamma, m)$.

Proposition III.15. *Let μ be a probability measure on E with compact support Γ. Let d_μ be the corresponding DTM. Consider a set $X \subseteq E$ such that $\Gamma \subseteq X$. The weighted Čech filtrations $V[\Gamma, \mathrm{d}_\mu]$ and $V[X, \mathrm{d}_\mu]$ are $c(\mu, m)$-interleaved.*

Moreover, if $Y \subseteq E$ is another set such that $\Gamma \subseteq Y$, $V[X, \mathrm{d}_\mu]$ and $V[Y, \mathrm{d}_\mu]$ are $c(\mu, m)$-interleaved.

In particular, if Γ is a finite set and $\mu = \mu_\Gamma$ its empirical measure, $W[\Gamma]$ and $V[X, \mathrm{d}_{\mu_\Gamma}]$ are $c(\Gamma, m)$-interleaved.

Proof. Let $c = c(\mu, m)$. Since $\Gamma \subseteq X$, we have $V^t[\Gamma, \mathrm{d}_\mu] \subseteq V^t[X, \mathrm{d}_\mu]$ for every $t \geq 0$.

Let us show that, for every $t \geq 0$, $V^t[X, \mathrm{d}_\mu] \subseteq V^{t+c}[\Gamma, \mathrm{d}_\mu]$. Let $x \in X$, and choose $\gamma \in \Gamma$ a projection of x on the compact set Γ, i.e. one of the closest points to x in Γ. By definition of the DTM, we have that $\mathrm{d}_\mu(x) \geq \|x - \gamma\|$. Together with $\mathrm{d}_\mu(\gamma) \leq c$, we obtain

$$t + c - \mathrm{d}_\mu(\gamma) \geq t \geq t - \mathrm{d}_\mu(x) + \|x - \gamma\|,$$

which means that $\overline{\mathcal{B}}_{\mathrm{d}_\mu}(x, t) \subseteq \overline{\mathcal{B}}_{\mathrm{d}_\mu}(\gamma, t+c)$. The inclusion $V^t[X, \mathrm{d}_\mu] \subseteq V^{t+c}[\Gamma, \mathrm{d}_\mu]$ follows.

If Y is another set containing Γ, we obtain $V^t[X, \mathrm{d}_\mu] \subseteq V^{t+c}[\Gamma, \mathrm{d}_\mu] \subseteq V^{t+c}[Y, \mathrm{d}_\mu]$ for every $t \geq 0$.

Theorem III.16. *Consider two measures μ, ν on E with supports X and Y. Let μ', ν' be two measures with compact supports Γ and Ω such that $\Gamma \subseteq X$ and $\Omega \subseteq Y$. We have*

$$\begin{aligned}\mathrm{d_i}\left(W[\mu], W[\nu]\right) \leq & m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu, \mu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu', \nu'\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\nu', \nu\right) \\ & + c(\mu', m) + c(\nu', m).\end{aligned}$$

In particular, if X and Y are finite, we have

$$\mathrm{d_i}\left(W[X], W[Y]\right) \leq m^{-\frac{1}{2}}\mathrm{W}_2\left(X, \Gamma\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\Gamma, \Omega\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\Omega, Y\right) + c(\Gamma, m) + c(\Omega, m).$$

Moreover, with $\Omega = Y$, we obtain

$$\mathrm{d_i}\left(W[X], W[\Omega]\right) \leq m^{-\frac{1}{2}}\mathrm{W}_2\left(X, \Gamma\right) + m^{-\frac{1}{2}}\mathrm{W}_2\left(\Gamma, \Omega\right) + c(\Gamma, m) + c(\Omega, m).$$

Proof. Define $d_X = \mathrm{d}_\mu$, $d_Y = \mathrm{d}_\nu$, $d_\Gamma = \mathrm{d}_{\mu'}$ and $d_\Omega = \mathrm{d}_{\nu'}$. We prove the first assertion by introducing the following filtrations between $W[\mu] = V[X, d_X]$ and $W[\nu] = V[Y, d_Y]$:

$$V[X, d_X] \leftrightarrow V[X, d_\Gamma] \leftrightarrow V[\Gamma \cup \Omega, d_\Gamma] \leftrightarrow V[\Gamma \cup \Omega, d_\Omega] \leftrightarrow V[Y, d_\Omega] \leftrightarrow V[Y, d_Y].$$

We have:

- $d_i(V[X, d_X], V[X, d_\Gamma]) \leq m^{-\frac{1}{2}} \mathrm{W}_2(\mu, \mu')$ (Propositions III.11 and III.4),
- $d_i(V[X, d_\Gamma], V[\Gamma \cup \Omega, d_\Gamma]) \leq c(\mu', m)$ (Proposition III.15),
- $d_i(V[\Gamma \cup \Omega, d_\Gamma], V[\Gamma \cup \Omega, d_\Omega]) \leq m^{-\frac{1}{2}} \mathrm{W}_2(\mu', \nu')$ (Propositions III.11 and III.4),
- $d_i(V[\Gamma \cup \Omega, d_\Omega], V[Y, d_\Omega]) \leq c(\nu', m)$ (Proposition III.15),
- $d_i(V[Y, d_\Omega], V[Y, d_Y]) \leq m^{-\frac{1}{2}} \mathrm{W}_2(\nu', \nu)$ (Propositions III.11 and III.4).

The inequality with X and Y finite follows from defining μ, ν, μ' and ν' to be the empirical measures on X, Y, Γ and Ω, and by recalling that the DTM-filtrations $W[X]$ and $W[Y]$ are equal to the weighted Čech filtration $V[X, \mathrm{d}_\mu]$ and $V[Y, \mathrm{d}_\nu]$.

The last inequality of Theorem III.16 can be seen as an approximation result. Indeed, suppose that Ω is some underlying set of interest, and X is a sample of it with, possibly, noise or outliers. If one can find a subset Γ of X such that X and Γ are close to each other—in the Wasserstein metric—and such that Γ and Ω are also close, then the filtrations $W[X]$ and $W[\Omega]$ are close. Their closeness depends on the constants $c(\Gamma, m)$ and $c(\Omega, m)$. More generally, suppose that X is finite and μ' is a measure with compact support $\Omega \subset X$ not necessarily finite. Using the definition of the DTM-filtration $W[\mu'] = V[\Omega, \mathrm{d}_{\mu'}]$, the first inequality gives

$$\mathrm{d_i}\left(W[X], W[\mu']\right) \leq m^{-\frac{1}{2}} \mathrm{W}_2(X, \Gamma) + m^{-\frac{1}{2}} \mathrm{W}_2(\mu_\Gamma, \mu') + c(\Gamma, m) + c(\mu', m).$$

For any probability measure μ of support $\Gamma \subseteq E$, the constant $c(\mu, m)$ might be seen as a bias term, expressing the behaviour of the DTM over Γ. It relates to the concentration of μ on its support. A usual case is the following: a measure μ with support Γ is said to be (a, d)-standard, with $a, d \geq 0$, if for all $x \in \Gamma$ and $r \geq 0$, $\mu(\overline{\mathcal{B}}(x, r)) \geq \min\{ar^d, 1\}$. For example, the Hausdorff measure associated to a compact d-dimensional submanifold of E is (a, d)-standard for some $a > 0$. In this case, a simple computation shows that there exists a constant C, depending only on a and d, such that for all $x \in \Gamma$, $\mathrm{d}_{\mu,m}(x) \leq Cm^{\frac{1}{d}}$. Therefore, $c(\mu, m) \leq Cm^{\frac{1}{d}}$. We study this property more deeply in Subsection III.2.6.

Regarding the second inequality of Theorem III.16, suppose for the sake of simplicity that one can choose $\Gamma = \Omega$. The bound of Theorem III.16 then reads

$$d_i(W[X], W[Y]) \leq m^{-\frac{1}{2}} \mathrm{W}_2(X, \Gamma) + m^{-\frac{1}{2}} \mathrm{W}_2(\Gamma, Y) + 2c(\Gamma, m).$$

Therefore, the DTM-filtrations $W[X]$ and $W[Y]$ are close to each other if μ_X and μ_Y are both close to a common measure μ_Γ. This would be the case if X and Y are noisy samples of Γ. This bound, expressed in terms of Wasserstein distance rather than Hausdorff distance, shows the robustness of the DTM-filtration to outliers.

Notice that, in practice, for finite data sets X and Y, the constants $c(X, m)$ and $c(Y, m)$ can be explicitly computed, as it amounts to evaluating the DTM on X and Y. This remark holds for the bounds of Theorem III.16.

Example III.17. Consider the set $X = \widetilde{X} \cup \Gamma$ as defined in Example III.13. Figure III.5 displays the sets $W^t[X]$, $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ and $W^t[\Gamma]$ for the values $p = 1$,

$m = 0.1$ and $t = 0.4$, and the persistence diagrams of the corresponding weighted Rips filtrations. This illustrates the stability properties of Proposition III.15 and Theorem III.16.

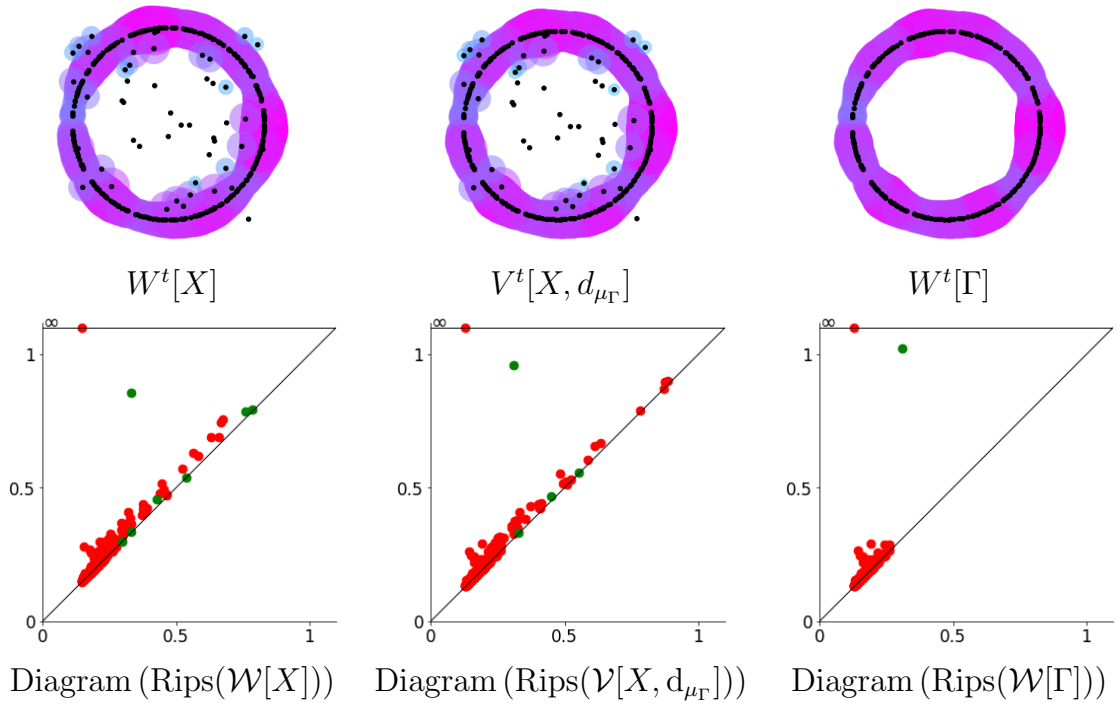


Figure III.5: Filtrations for $t = 0.4$, and their corresponding persistence diagrams.

The following proposition relates the DTM-filtration to the filtration of E by the sublevel sets of the DTM.

Proposition III.18. *Let μ be a probability measure on E with compact support K. Let $m \in [0, 1)$ and denote by V the sublevel sets filtration of d_μ. Consider a finite set $X \subseteq E$. Then*

$$\mathrm{d_i}\left(V, W[X]\right) \leq m^{-\frac{1}{2}} \mathrm{W}_2\left(\mu, \mu_X\right) + 2\epsilon + c(\mu, m),$$

with $\epsilon = \mathrm{d_H}\left(K \cup X, X\right)$.

Proof. First, notice that $V = V[E, \mathrm{d}_\mu]$. Indeed, for every $t \geq 0$, we have $V^t \subseteq V^t[E, \mathrm{d}_\mu]$ by definition of the weighted Čech filtration. To prove that $V^t[E, \mathrm{d}_\mu] \subseteq V^t$, let $x \in V^t[E, \mathrm{d}_\mu]$, and $y \in E$ such that $x \in \overline{\mathcal{B}}_{\mathrm{d}_\mu}(y, t)$. We have $\|x - y\| \leq t - f(y)$. For d_μ is 1-Lipschitz, we deduce

$$f(x) \leq f(y) + \|x - y\| \leq f(y) + t - f(y) \leq t,$$

hence $x \in V^t$. Then we compute:

$$\begin{aligned}\mathrm{d_i}\left(V, W[X]\right) &= \mathrm{d_i}\left(V[E, \mathrm{d}_\mu], V[X, \mathrm{d}_{\mu_X}]\right) \\ &\leq \mathrm{d_i}\left(V[E, \mathrm{d}_\mu], V[X \cup K, \mathrm{d}_\mu]\right) + \mathrm{d_i}\left(V[X \cup K, \mathrm{d}_\mu], V[X, \mathrm{d}_\mu]\right) \\ &\qquad + \mathrm{d_i}\left(V[X, \mathrm{d}_\mu], V[X, \mathrm{d}_{\mu_X}]\right) \\ &\leq c(\mu, m) + 2\epsilon + m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu, \mu_X\right),\end{aligned}$$

using Proposition III.11 for the first term, Proposition III.5 for the second one, and Proposition III.4 and Proposition III.15 for the third one.

As a consequence, one can use the DTM-filtration to approximate the persistent homology of the sublevel sets filtration of the DTM, which is expensive to compute in practice.

We close this subsection by noting that a natural strengthening of Theorem III.16 does not hold: let $m \in (0, 1)$ and $E = \mathbb{R}^n$ with $n \geq 1$. There is no constant C such that, for every probability measure μ, ν on E with supports X and Y, we have:

$$\mathrm{d_i}\left(V[X, \mathrm{d}_{\mu,m}], V[Y, \mathrm{d}_{\nu,m}]\right) \leq C\mathrm{W}_2\left(\mu, \nu\right).$$

The same goes for the weaker statement

$$\mathrm{d_i}\left(\mathbb{V}[X, \mathrm{d}_{\mu,m}], \mathbb{V}[Y, \mathrm{d}_{\nu,m}]\right) \leq C\mathrm{W}_2\left(\mu, \nu\right).$$

We shall prove the statement for $E = \mathbb{R}$. Let $q \in (0, 1)$ such that $q < m < 1-q$, and $\epsilon \in [0, q)$. Let $x \in (-1, 0)$ to be determined later. Define $\mu = q\delta_{-1} + (1-q)\delta_1$, and $\nu^\epsilon = (q-\epsilon)\delta_{-1} + (1-q)\delta_1 + \epsilon\delta_x$, with δ denoting the Dirac mass. Let $X = \{-1, 1\} \subset E$ and $Y = \{-1, x, 1\}$.

It is clear that $\mathrm{W}_2\left(\mu, \nu^\epsilon\right) = (x+1)\epsilon < \epsilon$. Using the triangle inequality, we have:

$$\begin{aligned}\mathrm{d_i}\left(\mathbb{V}[X, \mathrm{d}_{\mu,m}], \mathbb{V}[Y, \mathrm{d}_{\nu^\epsilon,m}]\right) &\geq \mathrm{d_i}\left(\mathbb{V}[X, \mathrm{d}_{\mu,m}], \mathbb{V}[Y, \mathrm{d}_{\mu,m}]\right) - \mathrm{d_i}\left(\mathbb{V}[Y, \mathrm{d}_{\nu^\epsilon,m}], \mathbb{V}[Y, \mathrm{d}_{\mu,m}]\right) \\ &\geq \mathrm{d_i}\left(\mathbb{V}[X, \mathrm{d}_{\mu,m}], \mathbb{V}[Y, \mathrm{d}_{\mu,m}]\right) - m^{-\frac{1}{2}}\epsilon\end{aligned}$$

Thus it is enough to show that $\mathrm{d_i}\left(\mathbb{V}[X, \mathrm{d}_{\mu,m}], \mathbb{V}[Y, \mathrm{d}_{\mu,m}]\right)$ is positive.

Since $1 - q > m$, we have $\mathrm{d}_{\mu,m}(1) = 0$. Using Proposition III.8, we deduce that the persistence barcode of the 0^{th} homology of $V[X, \mathrm{d}_\mu]$ consists of the bars $[0, +\infty)$ and $[\mathrm{d}_{\mu,m}(-1), \frac{1}{2}(\mathrm{d}_{\mu,m}(-1) + \mathrm{d}_{\mu,m}(1) + 2)]$.

Similarly, the persistence barcode of the 0th homology of $V[Y, \mathrm{d}_\mu]$ consists of the bars $[0, +\infty)$, $[\mathrm{d}_{\mu,m}(-1), \frac{1}{2}(\mathrm{d}_{\mu,m}(-1) + \mathrm{d}_{\mu,m}(x) + (1+x))]$ and $[\mathrm{d}_{\mu,m}(x), \frac{1}{2}(\mathrm{d}_{\mu,m}(x) + (1-x))]$.

Notice that, since $q > 0$ and $x < 0$, by definition of the DTM, we have $\mathrm{d}_{\mu,m}(x) < 1 - x$. Hence the last bar is not a singleton. Moreover, if x is close enough to 0, we have $\mathrm{d}_{\mu,m}(-1) < \mathrm{d}_{\mu,m}(x) + 1 - x$. Indeed, with $x = 0$, $\mathrm{d}_{\mu,m}(x) + 1 - x = 2$, and we have $\mathrm{d}_{\mu,m}(-1) = 2\sqrt{\frac{m-q}{m}} < 2$. Hence the second bar is not a singleton as well.

As a consequence, if x is close enough to 0, the interleaving distance between these two barcodes is positive.

III.2.4 Stability when $p > 1$

Now assume that $p > 1$, and $m \in (0, 1)$ is still fixed. In this case, stability properties turn out to be more difficult to establish. For small values of t, Lemma III.20 gives a tight non-additive interleaving between the filtrations. However, for large values of t, the filtrations are poorly interleaved. To overcome this issue we consider stability at the homological level, i.e., between the persistence modules associated to the DTM-filtrations.

Let us show first why one cannot expect a similar result as Proposition III.15. Consider the ambient space $E = \mathbb{R}^2$ and the sets $\Gamma = \{0\}$ and $X = \Gamma \cup \{1\}$. We have $\mathrm{d}_{\mu_\Gamma}(1) = 1$ and, for all $t \geq 1$, $W^t[\Gamma] = \overline{\mathcal{B}}(0, t)$ and $V^t[X, \mathrm{d}_{\mu_\Gamma}] = \overline{\mathcal{B}}(0, t) \cup \overline{\mathcal{B}}\left(1, (t^p - 1)^{\frac{1}{p}}\right)$. The sets $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ are represented in Figure III.6 for $t = 1.5$, $t = 5$ and several values of p.

For $p = 1$, the ball $\overline{\mathcal{B}}\left(1, (t^p - 1)^{\frac{1}{p}}\right)$ is contained in $\overline{\mathcal{B}}(0, t)$. Whereas for $p > 1$, the radius $(t^p - 1)^{\frac{1}{p}}$ is asymptotically equal to $t + o(\frac{1}{t^{p-1}})$ as $t \to +\infty$. Therefore, an $\epsilon \geq 0$ for which the ball $\overline{\mathcal{B}}\left(1, (t^p - 1)^{\frac{1}{p}}\right)$ would be included in $\overline{\mathcal{B}}(0, t + \epsilon)$ for all $t \geq 0$ should not be lower than $1 = \mathrm{d_H}(\Gamma, X)$. Therefore, $\mathrm{d_i}(W[\Gamma], V[X, \mathrm{d}_{\mu_\Gamma}]) = 1 = \mathrm{d_H}(\Gamma, X)$.

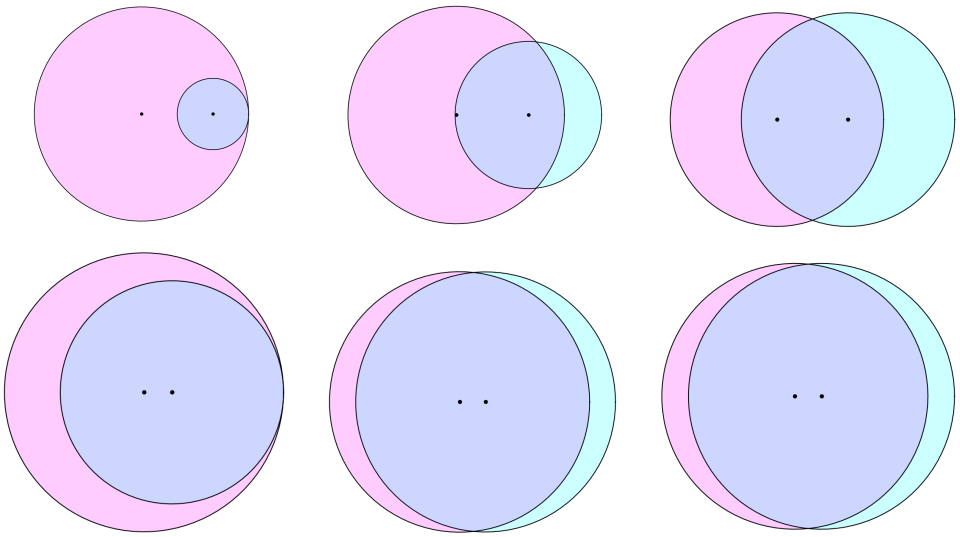


Figure III.6: Some sets $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ for $t = 1.5$ **(first row)** and $t = 5$ **(second row)**.

Even though the filtrations $W[\Gamma]$ and $V[X, \mathrm{d}_{\mu_\Gamma}]$ are distant, the set $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ is contractible for all $t \geq 0$, and therefore the interleaving distance between the persistence modules $\mathbb{W}[\Gamma]$ and $\mathbb{V}[X, \mathrm{d}_{\mu_\Gamma}]$ is 0.

In general, and in the same spirit as Proposition III.15, we can obtain a bound on the interleaving distance between the persistence modules $\mathbb{W}[\Gamma]$ and $\mathbb{V}[X, \mathrm{d}_{\mu_\Gamma}]$ which does not depend on X—see Proposition III.19.

If μ is a probability measure on E with compact support Γ, we define

$$c(\mu, m, p) = \sup_{\Gamma}(\mathrm{d}_{\mu,m}) + \kappa(p) t_\mu(\Gamma), \tag{III.3}$$

where $\kappa(p) = 1 - \frac{1}{p}$, and $t_\mu(\Gamma)$ is the filtration value of the simplex Γ in $\mathcal{N}(\mathcal{V}[\Gamma, \mathrm{d}_\mu])$, the (simplicial) weighted Čech filtration. Equivalently, $t_\mu(\Gamma)$ is the value t from which all the balls $\overline{\mathcal{B}}_{\mathrm{d}_\mu}(\gamma, t)$, $\gamma \in \Gamma$, share a common point.
If $\mu = \mu_\Gamma$ is the empirical measure of a finite set $\Gamma \subseteq E$, we denote it $c(\Gamma, m, p)$.

Note that we have the inequality $\frac{1}{2}\mathrm{diam}(\Gamma) \leq t_\mu(\Gamma) \leq 2\mathrm{diam}(\Gamma)$, where $\mathrm{diam}(\Gamma)$ denotes the diameter of Γ. This follows from writing $t_\mu(\Gamma) = \inf\left\{t \geq 0, \bigcap_{\gamma\in\Gamma} \overline{\mathcal{B}}_{\mathrm{d}_\mu}(\gamma, t) \neq \emptyset\right\}$ and using that $\forall \gamma \in \Gamma, \mathrm{d}_\mu(\gamma) \leq \mathrm{diam}(\Gamma)$.

Proposition III.19. *Let μ be a measure on E with compact support Γ, and d_μ be the corresponding DTM of parameter m. Consider a set $X \subseteq E$ such that $\Gamma \subseteq X$. The persistence modules $\mathbb{V}[\Gamma, \mathrm{d}_\mu]$ and $\mathbb{V}[X, \mathrm{d}_\mu]$ are $c(\mu, m, p)$-interleaved.*

Moreover, if $Y \subseteq E$ is another set such that $\Gamma \subseteq Y$, $\mathbb{V}[X, \mathrm{d}_\mu]$ and $\mathbb{V}[Y, \mathrm{d}_\mu]$ are $c(\mu, m, p)$-interleaved.

In particular, if Γ is a finite set and $\mu = \mu_\Gamma$ its empirical measure, $\mathbb{W}[\Gamma]$ and $\mathbb{V}[X, \mathrm{d}_{\mu_\Gamma}]$ are $c(\Gamma, m, p)$-interleaved.

The proof involves the two following ingredients, whose proofs are postponed to Subsection III.2.5. The first lemma gives a (non-additive) interleaving between the filtrations $W[\Gamma]$ and $V[X, \mathrm{d}_{\mu_\Gamma}]$, relevant for low values of t, while the second proposition gives a result for large values of t.

Lemma III.20. *Let μ, Γ and X be defined as in Proposition III.19. Let $\phi\colon t \mapsto 2^{1-\frac{1}{p}}t + \sup_\Gamma \mathrm{d}_\mu$. Then for every $t \geq 0$,*

$$V^t[\Gamma, \mathrm{d}_\mu] \subseteq V^t[X, \mathrm{d}_\mu] \subseteq V^{\phi(t)}[\Gamma, \mathrm{d}_\mu].$$

In the remainder of the section, we say that a homology group $H_n(\cdot)$ is trivial if it is of rank 0 when $n > 0$, or if it is of rank 1 when $n = 0$. We say that a homomorphism between homology groups $H_n(\cdot) \to H_n(\cdot)$ is trivial if the homomorphism is of rank 0 when $n > 0$, or if it is of rank 1 when $n = 0$.

Proposition III.21. *Let μ, Γ and X be as defined in Proposition III.19, and $c = c(\mu, m, p)$. Consider the map $v^t_*\colon \mathbb{V}^t[X, \mathrm{d}_\mu] \to \mathbb{V}^{t+c}[X, \mathrm{d}_\mu]$ induced in homology by the inclusion $v^t\colon V^t[X, \mathrm{d}_\mu] \to V^{t+c}[X, \mathrm{d}_\mu]$. If $t \geq t_\mu(\Gamma)$, then v^t is trivial.*

Proof of Proposition III.19. Denote $c = c(\mu, m, p)$. For every $t \geq 0$, denote by $v^t\colon V^t[X, \mathrm{d}_\mu] \to V^{t+c}[X, \mathrm{d}_\mu]$, $w^t\colon V^t[\Gamma, \mathrm{d}_\mu] \to V^{t+c}[\Gamma, \mathrm{d}_\mu]$ and $j^t\colon V^t[\Gamma, \mathrm{d}_\mu] \to V^t[X, \mathrm{d}_\mu]$ the inclusion maps, and v^t_*, w^t_*, and j^t_* the induced maps in homology.

Notice that, for $t \leq t_\mu(\Gamma)$, the term $2^{1-\frac{1}{p}}t + \sup_\Gamma \mathrm{d}_\mu$ which appears in Lemma

III.20 can be bounded as follows:

$$\begin{aligned} 2^{1-\frac{1}{p}}t + \sup_{\Gamma} \mathrm{d}_\mu &= t + (2^{1-\frac{1}{p}} - 1)t + \sup_{\Gamma} \mathrm{d}_\mu \\ &\leq t + (2^{1-\frac{1}{p}} - 1)t_\mu(\Gamma) + \sup_{\Gamma} \mathrm{d}_{\mu_\Gamma} \\ &\leq t + (1 - \frac{1}{p})t_\mu(\Gamma) + \sup_{\Gamma} \mathrm{d}_{\mu_\Gamma} \\ &= t + c \end{aligned}$$

where, for the second line, we used $2^{1-\frac{1}{p}} - 1 \leq 1 - \frac{1}{p}$ (Lemma III.27). Consequently, for every $t \leq t_\mu(\Gamma)$, we have $V^t[X, \mathrm{d}_\mu] \subseteq V^{t+c}[\Gamma, \mathrm{d}_\mu]$. Thus, for $t \geq 0$, we can define a map $\pi^t \colon \mathbb{V}^t[X, \mathrm{d}_\mu] \to \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu]$ as follows: π^t is the map induced by the inclusion if $t \leq t_\mu(\Gamma)$, and the zero map if $t \geq t_\mu(\Gamma)$.

The families $(\pi^t)_{t\geq 0}$ and $(j_*^t)_{t\geq 0}$ clearly are c-morphisms of persistence modules. Let us show that the pair $((\pi^t)_{t\geq 0}, (j_*^t)_{t\geq 0})$ defines a c-interleaving between $\mathbb{V}[\Gamma, \mathrm{d}_\mu]$ and $\mathbb{V}[X, \mathrm{d}_\mu]$.

Let $t \geq 0$. We shall show that the following diagrams commute:

$$\begin{array}{ccc} \mathbb{V}^t[X, \mathrm{d}_\mu] & \xrightarrow{v_*^t} & \mathbb{V}^{t+c}[X, \mathrm{d}_\mu] \\ & \searrow^{\pi^t} & \uparrow{\scriptstyle j_*^{t+c}} \\ & & \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu] \end{array} \qquad \begin{array}{ccc} \mathbb{V}^t[X, \mathrm{d}_\mu] & & \\ \uparrow{\scriptstyle j_*^t} & \searrow^{\pi^t} & \\ \mathbb{V}^t[\Gamma, \mathrm{d}_\mu] & \xrightarrow[w_*^t]{} & \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu] \end{array}$$

If $t \leq t_\mu(\Gamma)$, these diagrams can be obtained by applying the homology functor to the inclusions

$$V^t[\Gamma, \mathrm{d}_\mu] \subseteq V^t[X, \mathrm{d}_\mu] \subseteq V^{t+c}[\Gamma, \mathrm{d}_\mu] \subseteq V^{t+c}[X, \mathrm{d}_\mu].$$

If $t \geq t_\mu(\Gamma)$, the homology group $\mathbb{V}^t[\Gamma, \mathrm{d}_\mu]$ is trivial. Therefore the commutativity of the second diagram is obvious, and the commutativity of the first one follows from Proposition III.21. This shows that $\mathbb{V}[\Gamma, \mathrm{d}_\mu]$ and $\mathbb{V}[X, \mathrm{d}_\mu]$ are c-interleaved.

If Y is another set containing Γ, define, for all $t \geq 0$, the inclusions $u^t \colon V^t[Y, \mathrm{d}_\mu] \to V^{t+c}[Y, \mathrm{d}_\mu]$ and $k^t \colon V^t[\Gamma, \mathrm{d}_\mu] \to V^{t+c}[Y, \mathrm{d}_\mu]$. We can also define a map $\theta^t \colon \mathbb{V}^t[Y, \mathrm{d}_\mu] \to \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu]$ as we did for $\pi^t \colon \mathbb{V}^t[X, \mathrm{d}_\mu] \to \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu]$.

We can compose the previous diagrams to obtain the following:

$$\begin{array}{ccccc} \mathbb{V}^t[X, \mathrm{d}_\mu] & \xrightarrow{v_*^t} & \mathbb{V}^{t+c}[X, \mathrm{d}_\mu] & \xrightarrow{v_*^{t+c}} & \mathbb{V}^{t+2c}[X, \mathrm{d}_\mu] \\ & \searrow^{\pi^t} & \uparrow{\scriptstyle j_*^{t+c}} & \searrow^{\pi^{t+c}} & \uparrow{\scriptstyle j_*^{t+2c}} \\ & & \mathbb{V}^{t+c}[\Gamma, \mathrm{d}_\mu] & \xrightarrow[w_*^{t+c}]{} & \mathbb{V}^{t+2c}[\Gamma, \mathrm{d}_\mu] \\ & & \downarrow{\scriptstyle k_*^{t+c}} & \nearrow_{\theta^{t+c}} & \\ & & \mathbb{V}^{t+c}[Y, \mathrm{d}_\mu] & & \end{array}$$

Since all the triangles commute, so does the following:

$$\begin{array}{ccc} \mathbb{V}^t[X,\mathrm{d}_\mu] & \xrightarrow{\quad v_*^{t+2c} \quad} & \mathbb{V}^{t+2c}[X,\mathrm{d}_\mu] \\ & {\scriptstyle k_*^{t+c}\pi^t} \searrow \quad \nearrow {\scriptstyle j_*^{t+2c}\theta^{t+c}} & \\ & \mathbb{V}^{t+c}[Y,\mathrm{d}_\mu] & \end{array}$$

We can obtain the same interchanging X and Y. Therefore, by definition, the persistence modules $\mathbb{V}[X,\mathrm{d}_{\mu_\Gamma}]$ and $\mathbb{V}[Y,\mathrm{d}_{\mu_\Gamma}]$ are c-interleaved, with the interleaving $((k_*^{t+c}\pi^t)_{t\geq 0},(j_*^{t+c}\theta^t)_{t\geq 0})$.

Theorem III.22. *Consider two measures μ,ν on E with supports X and Y. Let μ',ν' be two measures with compact supports Γ and Ω such that $\Gamma\subseteq X$ and $\Omega\subseteq Y$. We have*

$$\begin{aligned}\mathrm{d_i}\left(\mathbb{W}[\mu],\mathbb{W}[\nu]\right) \leq\ & m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu,\mu'\right)+m^{-\frac{1}{2}}\mathrm{W}_2\left(\mu',\nu'\right)+m^{-\frac{1}{2}}\mathrm{W}_2\left(\nu',\nu\right)\\ &+c(\mu',m,p)+c(\nu',m,p).\end{aligned}$$

In particular, if X and Y are finite, we have

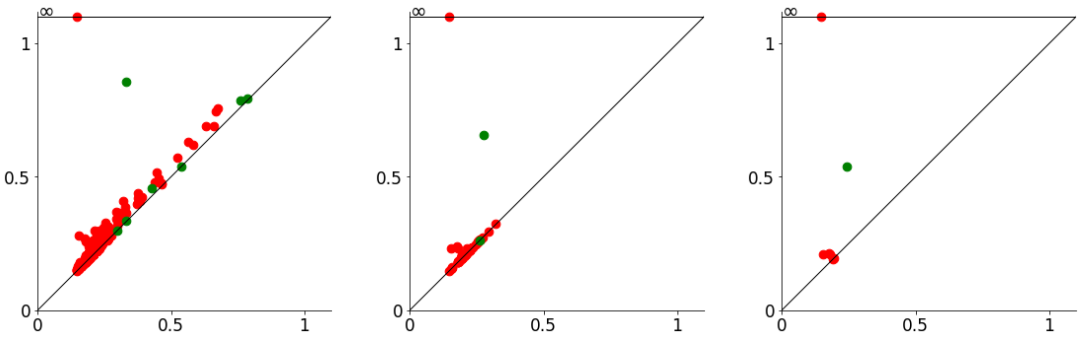
$$\begin{aligned}\mathrm{d_i}\left(\mathbb{W}[X],\mathbb{W}[Y]\right) \leq\ & m^{-\frac{1}{2}}\mathrm{W}_2\left(X,\Gamma\right)+m^{-\frac{1}{2}}\mathrm{W}_2\left(\Gamma,\Omega\right)+m^{-\frac{1}{2}}\mathrm{W}_2\left(\Omega,Y\right)\\ &+c(\Gamma,m,p)+c(\Omega,m,p).\end{aligned}$$

Moreover, with $\Omega=Y$, we obtain

$$\mathrm{d_i}\left(\mathbb{W}[X],\mathbb{W}[\Gamma]\right) \leq m^{-\frac{1}{2}}\mathrm{W}_2\left(X,\Gamma\right)+m^{-\frac{1}{2}}\mathrm{W}_2\left(\Gamma,\Omega\right)+c(\Gamma,m,p)+c(\Omega,m,p).$$

Proof. The proof is the same as Theorem III.16, using Proposition III.19 instead of Proposition III.15.

Notice that when $p=1$, the constant $c(\Gamma,m,p)$ is equal to the constant $c(\Gamma,m)$ defined in Subsection III.2.3, and we recover Theorem III.16 in homology.



$p=1$ $p=2$ $p=\infty$

Figure III.7: Persistence diagrams of the simplicial filtrations $\mathrm{Rips}(\mathcal{W}[X])$ for several values of p.

As an illustration of these results, we represent in Figure III.7 the persistence diagrams associated to the filtration $\mathrm{Rips}(\mathcal{W}[X])$ for several values of p. The point

cloud X is the one defined in Example III.13. Observe that, as stated in Proposition III.9, the number of red points (homology in dimension 0) is non-increasing with respect to p.

III.2.5 Proof of Lemma III.20 and Proposition III.21

We first prove the lemma stated in the previous subsection.

Proof of Lemma III.20. Denote $f = \mathrm{d}_\mu$. Let $x \in X$, and γ a projection of x on Γ. Let us show that for every $t \geq 0$,

$$\overline{\mathcal{B}}_f(x,t) \subseteq \overline{\mathcal{B}}_f(\gamma, 2^{1-\frac{1}{p}}t + f(\gamma)),$$

and the lemma will follow.

Define $d = f(\gamma)$. Let $u \in E$. By definition of the balls, we have

$$\begin{cases} u \in \overline{\mathcal{B}}_f(\gamma,t) \iff t \geq \left(\|u-\gamma\|^p + f(\gamma)^p\right)^{\frac{1}{p}}, \\ u \in \overline{\mathcal{B}}_f(x,t) \iff t \geq \left(\|u-x\|^p + f(x)^p\right)^{\frac{1}{p}}. \end{cases}$$

We shall only use

$$\begin{cases} u \in \overline{\mathcal{B}}_f(\gamma,t) \impliedby t \geq \|u-\gamma\| + d, \\ u \in \overline{\mathcal{B}}_f(x,t) \implies t \geq \left(\|u-x\|^p + \|x-\gamma\|^p\right)^{\frac{1}{p}}. \end{cases}$$

Let $u \in \overline{\mathcal{B}}_f(x,t)$. Let us prove that $u \in \overline{\mathcal{B}}_f(\gamma, 2^{1-\frac{1}{p}}t + d)$. If $\|u-\gamma\| \leq \|\gamma - x\|$, then $t \geq \|u-\gamma\|$, and we deduce $u \in \overline{\mathcal{B}}_f(\gamma, t+d) \subseteq \overline{\mathcal{B}}_f(\gamma, 2^{1-\frac{1}{p}}t + d)$.

Else, we have $\|u-\gamma\| \geq \|\gamma - x\|$. Consider the line segment $[\gamma, u]$ and the sphere $\mathcal{S}(\gamma, \|\gamma - x\|)$ of center γ and radius $\|\gamma - x\|$. The intersection $\mathcal{S}(\gamma, \|\gamma - x\|) \cap [\gamma, u]$ is a singleton. Call its element x'. The situation is represented in Figure III.8.

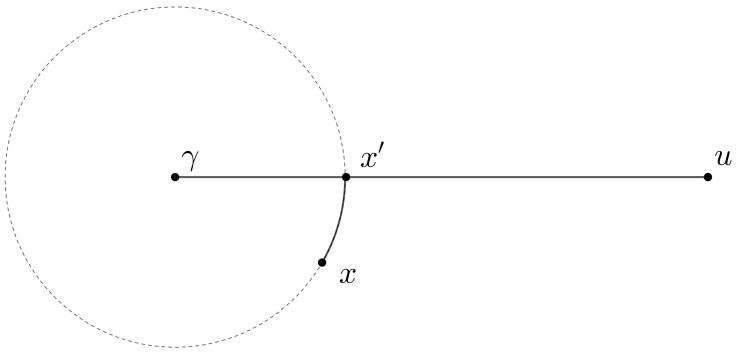


Figure III.8: Definition of the point x'.

We have $\|u-x'\| \leq \|u-x\|$ and $\|\gamma - x'\| = \|\gamma - x\|$. Therefore

$$\left(\|u-x'\|^p + \|x'-\gamma\|^p\right)^{\frac{1}{p}} \leq \left(\|u-x\|^p + \|x-\gamma\|^p\right)^{\frac{1}{p}}.$$

We also have $\|\gamma - u\| = \|\gamma - x'\| + \|x' - u\|$ and $\left(\|u-x\|^p + \|x-\gamma\|^p\right)^{\frac{1}{p}} \leq t$. Thus it follows from the last inequality that

$$\left(\|u-x'\|^p + (\|u-\gamma\| - \|u-x'\|)^p\right)^{\frac{1}{p}} \leq t.$$

The left-hand term of this inequality is not lower than $2^{\frac{1}{p}-1}\|u-\gamma\|$. Indeed, consider the function $s \mapsto \left(s^p + (\|u-\gamma\|-s)^p\right)^{\frac{1}{p}}$ defined for $s \in [0, \|u-\gamma\|]$. One shows directly, by computing its derivative, that its minimum is $2^{\frac{1}{p}-1}\|u-\gamma\|$, attained at $s = \frac{\|u-\gamma\|}{2}$.

We deduce that $2^{\frac{1}{p}-1}\|u-\gamma\| \leq t$, and $\|u-\gamma\| \leq 2^{1-\frac{1}{p}}t$. Thus $u \in \overline{\mathcal{B}}_f(\gamma, 2^{1-\frac{1}{p}}t + d)$.

Notice that the previous lemma gives a tight bound, as we can see with the following example. Consider set $\Gamma = \{0\} \subset \mathbb{R}$, $L > 0$, and $X = \Gamma \cup \{x\}$ with $x = \frac{L}{2}$. Let $m < \frac{1}{2}$, and $f = \mathrm{d}_{\mu_\Gamma}$, which is the function distance to Γ. For all $t \geq 2^{\frac{1}{p}-1}L$, we have $L \in \overline{\mathcal{B}}_f(x,t)$. Indeed, $r_x(2^{\frac{1}{p}-1}L) = \left((2^{\frac{1}{p}-1}L)^p - (\frac{L}{2})^p\right)^{\frac{1}{p}} = \frac{L}{2}$. In comparison, for every $t < \phi(2^{\frac{1}{p}-1}L) = L$, we have $L \notin \overline{\mathcal{B}}_f(0,t)$.

Following this example, we can find a lower bound on the interleaving distance between the persistence modules $\mathbb{W}[\Gamma]$ and $\mathbb{V}[X, d_{\mu_\Gamma}]$. Consider $L > 0$, the set $\Gamma = \{0, 2L\} \subset \mathbb{R}$, $x = \frac{L}{2}$, and $X = \Gamma \cup \{x, 2L - x\}$. Let $m < \frac{1}{2}$, and $f = \mathrm{d}_{\mu_\Gamma}$. The persistence diagram of the 0th-homology of $W[\Gamma]$ consists of two points, $(0, +\infty)$ and $(0, L)$. Regarding $V[X, f]$, the point of finite ordinate in the persistence diagram of its 0^{th}-homology is $(0, 2^{\frac{1}{p}-1}L)$. Indeed, for $t = 2^{\frac{1}{p}-1}L$, we have $L \in \overline{\mathcal{B}}_f(x,t)$ and $L \in \overline{\mathcal{B}}_f(L-x,t)$, hence the set $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ is connected. We deduce that these persistence modules are at least $\left(1 - 2^{\frac{1}{p}-1}\right)$ L-interleaved.

In comparison, the upper bound we prove in Proposition III.19 is $\left(1 - \frac{1}{p}\right) L$.

We now prove the proposition stated in the previous subsection.

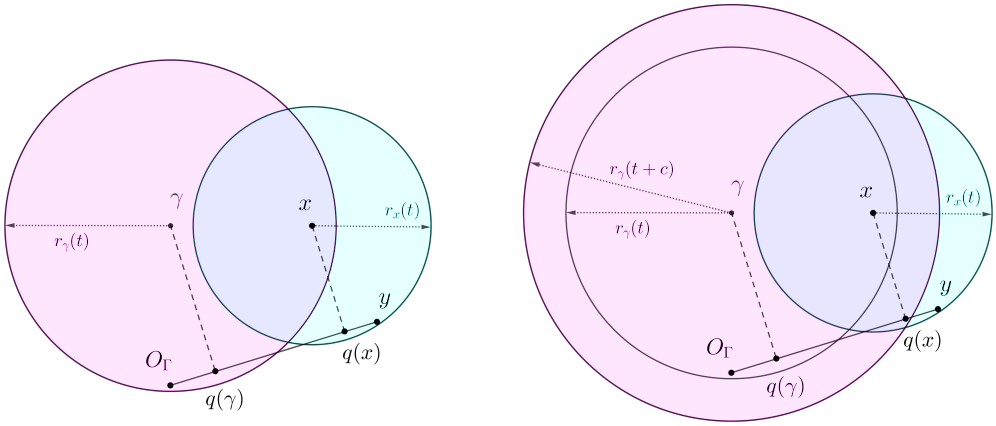
Proof of Proposition III.21. Denote $f = \mathrm{d}_\mu$. Let $t \geq t_\mu(\Gamma)$. By definition of $t_\mu(\Gamma)$, there exists a point $O_\Gamma \in \bigcap_{\gamma \in \Gamma} \overline{\mathcal{B}}_f(\gamma, t_\mu(\Gamma))$.

In order to show that $v^t_* \colon \mathbb{V}^t[X, \mathrm{d}_\mu] \to \mathbb{V}^{t+c}[X, \mathrm{d}_\mu]$ is trivial, we introduce an intermediate set between $V^t[X, \mathrm{d}_{\mu_\Gamma}]$ and $V^{t+c}[X, \mathrm{d}_{\mu_\Gamma}]$:

$$\begin{cases} V^t[X, \mathrm{d}_{\mu_\Gamma}] & = \bigcup_{x \in X \setminus \Gamma} \overline{\mathcal{B}}_f(x,t) \cup \bigcup_{\gamma \in \Gamma} \overline{\mathcal{B}}_f(\gamma, t), \\ \widetilde{V}^t & := \bigcup_{x \in X \setminus \Gamma} \overline{\mathcal{B}}_f(x,t) \cup \bigcup_{\gamma \in \Gamma} \overline{\mathcal{B}}_f(\gamma, t+c), \\ V^{t+c}[X, \mathrm{d}_{\mu_\Gamma}] & = \bigcup_{x \in X \setminus \Gamma} \overline{\mathcal{B}}_f(x, t+c) \cup \bigcup_{\gamma \in \Gamma} \overline{\mathcal{B}}_f(\gamma, t+c). \end{cases}$$

Since $t \geq t_\mu(\Gamma)$, we have $O_\Gamma \in \widetilde{V}^t$. Let us show that $\widetilde{V}^t$ is star-shaped around O_Γ.

Let $x \in X$ and consider γ a projection of x on Γ. We first prove that $\overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t+c)$ is star-shaped around O_Γ. Let $y \in \overline{\mathcal{B}}_f(x,t)$. We have to show that the line segment $[y, O_\Gamma]$ is a subset of $\overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t+c)$. Let D be the affine line going through y and O_Γ, and denote by q the orthogonal projection on D. We have $[y, O_\Gamma] \subseteq [y, q(x)] \cup [q(x), O_\Gamma]$. The first line segment $[y, q(x)]$ is a subset of $\overline{\mathcal{B}}_f(x,t)$. Regarding the second line segment $[q(x), O_\Gamma]$, let us show that $q(x) \in \overline{\mathcal{B}}_f(\gamma, t+c)$, and $[q(x), O_\Gamma] \subseteq \overline{\mathcal{B}}_f(\gamma, t+c)$ will follow. The situation is pictured in Figure III.9.



$\overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t)$ $\overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t+c)$

Figure III.9: Construction of an intermediate set $\widetilde{V}^t$.

According to Lemma III.28,

$$\|\gamma - q(x)\|^2 \leq \|x-\gamma\|^2 + \|x-q(x)\|(2\|\gamma - q(\gamma)\| - \|x-q(x)\|).$$

Let $d = \|x-q(x)\|$. Since $d = \|x-q(x)\| \leq \left(t^p - d_\mu(x)^p\right)^{\frac{1}{p}} \leq \left(t^p - \|x-\gamma\|^p\right)^{\frac{1}{p}}$, we have $\|x-\gamma\| \leq (t^p - d^p)^{\frac{1}{p}}$. Moreover, $\|\gamma - q(\gamma)\| \leq \|\gamma - O_\Gamma\| \leq t_\mu(\Gamma)$. The last inequality then gives

$$\|\gamma - q(x)\|^2 \leq (t^p - d^p)^{\frac{2}{p}} + d(2t_\mu(\Gamma) - d).$$

According to Lemma III.29, we obtain that $\|\gamma - q(x)\|$ is not greater than $t + \kappa(p) t_\mu(\Gamma)$. Therefore, we have the inequality

$$\left((t+\kappa(p)t_\mu(\Gamma)+f(\gamma))^p - f(\gamma)^p\right)^{\frac{1}{p}} \geq \left(t+\kappa(p)t_\mu(\Gamma)+f(\gamma)\right) - f(\gamma) \geq \|\gamma - q(x)\|,$$

and we deduce

$$q(x) \in \overline{\mathcal{B}}_f(\gamma, t + \kappa(p)t_\mu(\Gamma) + f(\gamma)) \subset \overline{\mathcal{B}}_f(\gamma, t+c).$$

In conclusion, $[y, O_\Gamma] \subset \overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t+c)$. This being true for every $y \in \overline{\mathcal{B}}_f(x,t)$, and obviously true for $y \in \overline{\mathcal{B}}_f(\gamma, t+c)$, we deduce that $\overline{\mathcal{B}}_f(x,t) \cup \overline{\mathcal{B}}_f(\gamma, t+c)$ is star-shaped around O_Γ. Finally, since $O_\Gamma \in \bigcap_{\gamma\in\Gamma} \overline{\mathcal{B}}_f(\gamma, t_X(\Gamma))$, we have that $\widetilde{V}^t$ is star-shaped around O_Γ.

To conclude the proof, notice that the map v^t_* factorizes through $H_n(\widetilde{V}^t)$. Indeed, consider the diagram of inclusions:

$$V^t[X, \mathrm{d}_{\mu_\Gamma}] \xrightarrow{\quad v^t \quad} \widetilde{V}^t \longrightarrow V^{t+c}[X, \mathrm{d}_{\mu_\Gamma}].$$

Applying the singular homology functor, we obtain

$$\mathbb{V}^t[X, \mathrm{d}_{\mu_\Gamma}] \xrightarrow{\quad v^t_* \quad} H_n(\widetilde{V}^t) \longrightarrow \mathbb{V}^{t+c}[X, \mathrm{d}_{\mu_\Gamma}].$$

Since $\widetilde{V}^t$ is star-shaped, $H_n(\widetilde{V}^t)$ is trivial, and so is v^t_*.

III.2.6 Consequences under the standard assumption

In this subsection, we state two useful consequences of Theorem III.16 and Proposition III.18, under a classic regularity assumption on the measure μ. We will use Corollary III.24 in the Chapter IV. We only study the case $p = 1$.

The assumption in question is known as the (a,d)-standard assumption. A measure μ on E is said *(a,d)-standard*, with $a, d > 0$, if the following is satisfied:

$$\forall x \in \mathrm{supp}(\mu),\ \forall r \geq 0,\ \text{we have}\ \ \mu(\overline{\mathcal{B}}(x,r)) \geq \min\{ar^d, 1\}.$$

Let $m \in (0,1)$. In order to obtain precise results, we will consider the following formulation of (a,d)-standardness:

$$\forall x \in \mathrm{supp}(\mu),\ \forall r \in \left[0, \left(\frac{m}{a}\right)^{\frac{1}{d}}\right],\ \text{we have}\ \ \mu(\overline{\mathcal{B}}(x,r)) \geq \min\{ar^d, 1\}. \quad \text{(III.4)}$$

These two formulations are equivalent, with potentially a different constant a.

Let $p = 1$, $m \in (0,1)$, and μ, ν be probability measures on E. We first prove a result announced earlier. We recall the reader that the quantity $c(\mu, m)$ is defined as

$$c(\mu, m) = \sup_{x \in \mathrm{supp}(\mu)} \mathrm{d}_{\mu,m}(x).$$

Lemma III.23. *Suppose that μ satisfies the (a,d)-standard assumption. Then $c(\mu,m) \leq C m^{\frac{1}{d}}$ with $C = a^{-\frac{1}{d}}$.*

Proof. By definition,

$$\delta_{\mu,t}(x) = \inf\left\{r \geq 0,\ \mu\left(\overline{\mathcal{B}}(x,r)\right) > t\right\} \qquad \text{and} \qquad \mathrm{d}^2_{\mu,m}(x) = \frac{1}{m}\int_0^m \delta^2_{\mu,t}(x)\mathrm{d}t.$$

Using the assumption $\mu(\mathcal{B}(x,r)) \geq ar^d$ for all $x \in \mathrm{supp}(\mu)$, we get $\delta_{\mu,t}(x) \leq (\frac{t}{a})^{\frac{1}{d}}$, and a simple computation yields

$$\mathrm{d}^2_{\mu,m}(x) \leq \frac{d}{d+2}\left(\frac{t}{a}\right)^{\frac{2}{d}} \leq \left(\frac{t}{a}\right)^{\frac{2}{d}}.$$

Now we can restate Theorem III.16 without mentioning the intermediate measures μ' and ν'.

Corollary III.24. *Let μ, ν be such that $\mathrm{W}_2(\mu,\nu) = w \leq \frac{1}{4}$. Suppose that μ satisfies the (a,d)-standard assumption. Then*

$$\mathrm{d_i}(W[\mu], W[\nu]) \leq C_1\left(\frac{w}{m}\right)^{\frac{1}{2}} + 2C_2 m^{\frac{1}{d}}$$

with $C_1 = 8\mathrm{diam}(\mathrm{supp}(\mu)) + 5$ and $C_2 = a^{-\frac{1}{d}}$.

Proof. Let π be an optimal transport plan for $w = \mathrm{W}_2(\mu,\nu)$. Denote $\alpha = w^{\frac{1}{2}}$ and $D = \mathrm{diam}(\mathrm{supp}(\mu))$. Define π' to be π restricted to the set $\{x, y \in E, \|x-y\| < \alpha\}$. We denote its marginals μ' and ν'. By Markov inequality, $1 - |\pi'| \leq \frac{w^2}{\alpha^2} = w$, where we recall that $|\pi'|$ denotes the total mass of π'. Consider the restricted probability measures $\overline{\mu'}$ and $\overline{\nu'}$ (defined in Subsection IV.3.1). Let us show that we have

$$\mathrm{W}_2\left(\mu, \overline{\mu'}\right) = 2D\alpha, \quad \mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right) \leq \alpha \quad \text{and} \quad \mathrm{W}_2\left(\nu, \overline{\nu'}\right) \leq 2(1+D)\alpha. \tag{III.5}$$

The first inequality is an application of Lemma IV.45 (stated in the following chapter):

$$\mathrm{W}_2\left(\mu, \overline{\mu'}\right) \leq 2(1-|\mu'|)^{\frac{1}{2}} D = 2(1-|\pi'|)^{\frac{1}{2}} D \leq 2w^{\frac{1}{2}} D.$$

To obtain the second inequality, we write

$$\begin{aligned}\mathrm{W}_2^2(\overline{\mu'}, \overline{\nu'}) = \int \|x-y\|^2 \,\mathrm{d}\overline{\pi'}(x,y) &= \int \|x-y\| \,\frac{\mathrm{d}\pi'(x,y)}{|\pi'|} \\ &\leq \frac{1}{|\pi'|}\int \|x-y\| \,\mathrm{d}\pi(x,y).\end{aligned}$$

Hence Jensen inequality leads to $\mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right) \leq \frac{w}{|\pi'|^{\frac{1}{2}}}$. Since $1 - |\pi'| \leq w$, we have $\frac{w}{|\pi'|^{\frac{1}{2}}} \leq \frac{w}{1-w}$, and the assumption $w \leq \frac{1}{4}$ yields $\frac{w}{1-w} \leq \alpha$. This proves the second point. Finally, we obtain the third inequality by applying the triangular inequality:

$$\mathrm{W}_2\left(\nu, \overline{\nu'}\right) \leq \mathrm{W}_2(\nu,\mu) + \mathrm{W}_2\left(\mu, \overline{\mu'}\right) + \mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right).$$

Next, let us deduce that

$$\begin{aligned} & c(\overline{\mu'}, m) \leq c(\mu) + m^{-\frac{1}{2}} 2D\alpha \\ \text{and} \qquad & c(\overline{\nu'}, m) \leq c(\mu, m) + \left(m^{-\frac{1}{2}} + m^{-\frac{1}{2}} 2D + 1\right) \alpha. \end{aligned} \tag{III.6}$$

The first inequality follows from Theorem III.11:

$$c(\overline{\mu'}, m) = \sup_{x \in \mathrm{supp}(\overline{\mu'})} \mathrm{d}_{\overline{\mu'}}(x) \leq \sup_{x \in \mathrm{supp}(\overline{\mu'})} \mathrm{d}_{\mu}(x) + m^{-\frac{1}{2}} \mathrm{W}_2\left(\overline{\mu'}, \mu\right),$$

and we conclude with $\mathrm{W}_2\left(\mu, \overline{\mu'}\right) = 2D\alpha$. In order to prove the second inequality, we also use Theorem III.11:

$$c(\overline{\nu'}, m) = \sup_{x \in \mathrm{supp}(\overline{\nu'})} \mathrm{d}_{\overline{\nu'}}(x) \leq \sup_{x \in \mathrm{supp}(\overline{\nu'})} \mathrm{d}_{\overline{\mu'}}(x) + m^{-\frac{1}{2}} \mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right).$$

Since π' has support included in $\{x, y \in E, \|x - y\| < \alpha\}$, we can use Proposition III.10 to obtain

$$\sup_{x \in \mathrm{supp}(\overline{\nu'})} \mathrm{d}_{\overline{\mu'}}(x) \leq \sup_{x \in \mathrm{supp}(\overline{\mu'})} \mathrm{d}_{\overline{\mu'}}(x) + \alpha = c(\mu', m) + \alpha$$

and we deduce

$$\begin{aligned} c(\overline{\nu'}, m) &\leq c(\mu', m) + \alpha + m^{-\frac{1}{2}} \mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right) \\ &\leq c(\mu, m) + (m^{-\frac{1}{2}} + m^{-\frac{1}{2}} 2D + 1)\alpha. \end{aligned}$$

To conclude, Theorem III.16 gives

$$\begin{aligned} d_i(W[\mu], W[\nu]) &\leq m^{-\frac{1}{2}} \mathrm{W}_2\left(\mu, \overline{\mu'}\right) + m^{-\frac{1}{2}} \mathrm{W}_2\left(\overline{\mu'}, \overline{\nu'}\right) + m^{-1} \mathrm{W}_2\left(\nu, \overline{\nu'}\right) \\ &\quad + c(\overline{\mu'}, m) + c(\overline{\nu'}, m) \\ &\leq \left(m^{-\frac{1}{2}}(4D + 1) + 4(D + 1)\right)\alpha + 2c(\mu, m), \end{aligned}$$

where we used Equations (III.5) and (III.6) on the last line. Since $m \leq 1$, we can simplify this expression into

$$\mathrm{d_i}\left(W[\mu], W[\nu]\right) \leq m^{-\frac{1}{2}}(8D + 5)\alpha + 2c(\mu, m).$$

We conclude the proof using $c(\mu, m) \leq a^{-\frac{1}{d}} m^{\frac{1}{d}}$ (Lemma III.23).

The following corollary is to be seen as a consistency result for DTM-filtrations.

Corollary III.25. *Let μ be a probability measure on E that satisfies the (a, d)-standard assumption. Let $m \in (0, 1)$ and denote by V the sublevel sets filtration of $\mathrm{d}_{\mu,m}$. Then*

$$d_i(V, W[\mu, m]) \leq C m^{\frac{1}{d}}$$

with $C = a^{-\frac{1}{d}}$.

Proof. Let V denote the sublevel-set filtration of the DTM $\mathrm{d}_{\mu,m}$. Following the proof of Proposition III.18, we see that $W[\mu, m]$ and V are $c(\mu, m)$-interleaved. We conclude by using Lemma III.23.

III.3 Conclusion

In this chapter we have introduced the DTM-filtrations that depend on a parameter $p \geq 1$. This new family of filtrations extends the filtration introduced in [BCOS16] that corresponds to the case $p = 2$.

The established stability properties are, as far as we know, of a new type: the closeness of two DTM-filtrations associated to two data sets relies on the existence of a well-sampled underlying object that approximates both data sets in the Wasserstein metric. This makes the DTM-filtrations robust to outliers. Even though large values of p lead to persistence diagrams with less points in the 0^{th} homology, the choice of $p = 1$ gives the strongest stability results. When $p > 1$, the interleaving bound is less significant since it involves the diameter of the underlying object, but the obtained bound is consistent with the case $p = 1$ as it converges to the bound for $p = 1$ as p goes to 1.

It is interesting to notice that the proofs rely on only a few properties of the DTM. As a consequence, the results should extend to other weight functions, such that the DTM with an exponent parameter different from 2, or kernel density estimators. Some variants concerning the radius functions in the weighted Čech filtration, are also worth considering. The analysis shows that one should choose radius functions whose asymptotic behaviour look like the one of the case $p = 1$. In the same spirit as in [She13, BCOS16] where sparse-weighted Rips filtrations were considered, it would also be interesting to consider sparse versions of the DTM-filtrations and to study their stability properties.

Last, the obtained stability results, depending on the choice of underlying sets, open the way to the statistical analysis of the persistence diagrams of the DTM-filtrations.

III.A Supplementary results for Section III.1

Lemma III.26. *Let c, ϵ and x be non-negative real numbers, and $t \geq a$. Define $\alpha = (1 + c^p)^{\frac{1}{p}}$ and $k = \epsilon\alpha$. Then $t + k \geq a + c\epsilon$, and*

$$\big((t+k)^p - (x+c\epsilon)^p\big)^{\frac{1}{p}} - (t^p - x^p)^{\frac{1}{p}} \geq \epsilon$$

Proof. Let $\mathcal{D} = \{(t, x), t \geq x \geq 0\} \subseteq \mathbb{R}^2$. Let us find the minimum of

$$\begin{array}{rccl} \Phi & : & \mathcal{D} & \longrightarrow & \mathbb{R} \\ & & (t,x) & \longmapsto & \big((t+\alpha\epsilon)^p - (x+c\epsilon)^p\big)^{\frac{1}{p}} - (t^p - x^p)^{\frac{1}{p}} \end{array}$$

An $x > 0$ being fixed, we study $\phi\colon t \mapsto \Phi(t, x)$ on the interval $(x, +\infty)$. Its

derivative is

$$\phi'(t) = \frac{(t+\alpha\epsilon)^{p-1}}{\big((t+\alpha\epsilon)^p - (x+c\epsilon)^p\big)^{1-\frac{1}{p}}} - \frac{t^{p-1}}{(t^p - x^p)^{1-\frac{1}{p}}}$$

We solve:

$$\begin{aligned}
\phi'(t) = 0 &\iff (t+\alpha\epsilon)^{p-1}(t^p - x^p)^{1-\frac{1}{p}} = t^{p-1}\big((t+\alpha\epsilon)^p - (x+c\epsilon)^p\big)^{1-\frac{1}{p}} \\
&\iff \underbrace{(t+\alpha\epsilon)^p}(t^p - x^p) = \underbrace{t^p}\big((t+\alpha\epsilon)^p - (x+c\epsilon)^p\big) \\
&\iff (t+\alpha\epsilon)^p x^p = t^p (x+c\epsilon)^p \\
&\iff \frac{t+\alpha\epsilon}{t} = \frac{x+c\epsilon}{x} \\
&\iff t = \frac{\alpha}{c}x
\end{aligned}$$

We obtain the second line by raising the equality to the power of $\frac{p}{p-1}$. Hence the derivative of ϕ vanishes only at $t = \frac{\alpha}{c}x$. Together with $\lim_{+\infty}\phi = +\infty$, we deduce that ϕ attains its minimum at $t = x$ or $t = \frac{\alpha}{c}x$.

Let us show that $\phi(\frac{\alpha}{c}x) = \epsilon$.

$$\begin{aligned}
\phi(\tfrac{\alpha}{c}x) = \Phi(\tfrac{\alpha}{c}x, x) &= \big((\tfrac{\alpha}{c}x + \alpha\epsilon)^p - (x+c\epsilon)^p\big)^{\frac{1}{p}} &- &\big((\tfrac{\alpha}{c}x)^p - x^p\big)^{\frac{1}{p}} \\
&= \big((\tfrac{\alpha}{c})^p (x+c\epsilon)^p - (x+c\epsilon)^p\big)^{\frac{1}{p}} &- &x\big((\tfrac{\alpha}{c})^p - 1\big)^{\frac{1}{p}} \\
&= (x+c\epsilon)\big((\tfrac{\alpha}{c})^p - 1\big)^{\frac{1}{p}} &- &x\big((\tfrac{\alpha}{c})^p - 1\big)^{\frac{1}{p}} \\
&= c\epsilon\big((\tfrac{\alpha}{c})^p - 1\big)^{\frac{1}{p}}
\end{aligned}$$

Using $\alpha = (1+c^p)^{\frac{1}{p}}$, one obtains that $c\big((\frac{\alpha}{c})^p - 1\big)^{\frac{1}{p}} = 1$. Therefore, $\phi(\frac{\alpha}{c}x) = \epsilon$.

Secondly, consider Φ on the interval $\{(x,x), x \geq 0\}$.
The function $t \mapsto \Phi(x,x) = ((x+\alpha\epsilon)^p - (x+c\epsilon)^p)^{\frac{1}{p}}$ is increasing. Its minimum is $\Phi(0,0) = ((\alpha\epsilon)^p - (c\epsilon)^p)^{\frac{1}{p}} = \epsilon(\alpha^p - c^p)^{\frac{1}{p}} = \epsilon$.

In conclusion, on every interval $(x, +\infty) \times \{x\} \subseteq \mathcal{D}$, Φ admits ϵ as a minimum. Therefore, ϵ is the minimum of Φ on $\mathcal{D}$.

III.B Supplementary results for Section III.2

Lemma III.27. *For all $p \geq 1$, we have $2^{1-\frac{1}{p}} - 1 \leq 1 - \frac{1}{p}$.*

Proof. The convexity property of the function $x \mapsto 2^x$ gives, for all $x \in [0,1]$, $2^x \leq x + 1$. Hence $2^{1-\frac{1}{p}} - 1 \leq 1 - \frac{1}{p}$.

Lemma III.28. *Let $\gamma, x \in E$, D an affine line, and $q(\gamma), q(x)$ the projections of γ and x on D. Then*

$$\|\gamma - q(x)\|^2 \leq \|x - \gamma\|^2 + \|x - q(x)\|(2\|\gamma - q(\gamma)\| - \|x - q(x)\|).$$

Proof. We first study the case where γ, x and D lie in the same affine plane. If γ and x are on opposite sides of D, the result is obvious. Otherwise, the points $\gamma, x, q(\gamma)$ and $q(x)$ form a right trapezoid (see Figure III.10).

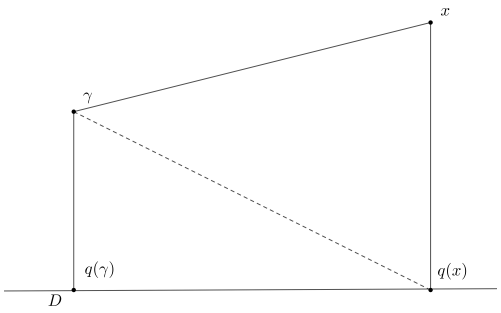


Figure III.10: The points $\gamma, x, q(\gamma)$ and $q(x)$ form a right trapezoid.

Using the Pythagorean theorem on the orthogonal vectors $\gamma - q(\gamma)$ and $q(\gamma) - q(x)$, and on $(\gamma - q(\gamma)) - (x - q(x))$ and $q(\gamma) - q(x)$, we obtain

$$\begin{cases} \|\gamma - q(\gamma)\|^2 + \|q(\gamma) - q(x)\|^2 = \|\gamma - q(x)\|^2, \\ \|(\gamma - q(\gamma)) - (x - q(x))\|^2 + \|q(\gamma) - q(x)\|^2 = \|\gamma - x\|^2. \end{cases}$$

Using that $\|(\gamma - q(\gamma)) - (x - q(x))\| = | \, \|\gamma - q(\gamma)\| - \|x - q(x)\| \, |$, the second equality rephrases as $\|q(\gamma) - q(x)\|^2 = \|\gamma - x\|^2 - (\|\gamma - q(\gamma)\| - \|x - q(x)\|)^2$. Combining these two equalities gives

$$\begin{aligned} \|\gamma - q(x)\|^2 &= \|\gamma - q(\gamma)\|^2 + \|q(\gamma) - q(x)\|^2 \\ &= \|\gamma - q(\gamma)\|^2 + \|\gamma - x\|^2 - (\|\gamma - q(\gamma)\| - \|x - q(x)\|)^2 \\ &= \|\gamma - x\|^2 + \|x - q(x)\|(2\|\gamma - q(\gamma)\| - \|x - q(x)\|). \end{aligned}$$

Now, if γ, x and D do not lie in the same affine plane, denote by P the affine plane containing D and x. Let $\tilde{\gamma}$ the point of P such that $\|\gamma - q(\gamma)\| = \|\tilde{\gamma} - q(\gamma)\|$ and $\|\gamma - q(x)\| = \|\tilde{\gamma} - q(x)\|$. Using the previous result on $\tilde{\gamma}$ and the inequality $\|\gamma - x\| \geq \|\tilde{\gamma} - x\|$, we obtain the result.

Lemma III.29. *Let $a, b, d \geq 0$ such that $b \leq a$ and $d \leq a$. Then*

$$(a^p - d^p)^{\frac{2}{p}} + d(2b - d) \leq (a + \kappa b)^2,$$

with $\kappa = 1 - \frac{1}{p}$.

Proof. The equation being homogeneous with respect to a, it is enough to show that

$$(1 - d^p)^{\frac{2}{p}} + d(2b - d) \leq (1 + \kappa b)^2$$

with $b \leq 1$ and $d \leq 1$. We shall actually show that $(1 - d^p)^{\frac{2}{p}} + d(2b - d) \leq 1 + 2\kappa b$.

Note that this is true when $d \leq \kappa$. Indeed, $(1-d^p)^{\frac{2}{p}} + d(2b-d) \leq 1+2db \leq 1+2\kappa b$. Now, notice that it is enough to show the inequality for $b = 1$. Indeed, it is equivalent to $(1-d^p)^{\frac{2}{p}} - 1 - d^2 \leq 2\kappa b - 2db = 2b(\kappa - d)$. For every $d \geq \kappa$, the right-hand side of this inequality is nonpositive, hence the worst case happens when $b = 1$. What is left to show is the following: $\forall d \in [\kappa, 1]$,

$$(1-d^p)^{\frac{2}{p}} + d(2-d) \leq 1 + 2\kappa.$$

The function $x \mapsto (1-x)^{\frac{1}{p}}$ being concave on $[0,1]$, we have $(1-x)^{\frac{1}{p}} \leq 1 - \frac{1}{p}x$ for all $x \in [0,1]$. Therefore, $(1-d^p)^{\frac{1}{p}} \leq 1 - \frac{1}{p}d^p$. Consider the function

$$\phi\colon d \mapsto (1 - \frac{1}{p}d^p)^2 + d(2-d).$$

Let us show that $\forall d \in [0,1], \phi(d) \leq 1 + 2\kappa$.

This inequality is obvious for $d = 0$. It is also the case for $d = 1$, since we obtain $(1 - \frac{1}{p}d^p)^2 + d(2-d) = (1-\frac{1}{p})^2 + 1 = \kappa^2 + 1$. On the interval $[0,1]$, the derivative of ϕ is $\phi'(d) = \frac{2}{p}d^{2p-1} - 2d^{p-1} - 2d + 2$. Let d_* be such that $\phi'(d_*) = 0$. Multiplying $\phi'(d_*)$ by $\frac{d_*}{2}$ gives the relation $\frac{1}{p}d_*^{2p} - d_*^p - d_*^2 + d_* = 0$. Subtracting this equality in $\phi(d_*)$ gives $\phi(d_*) = 1 - (\frac{1}{p} - \frac{1}{p^2})d_*^{2p} + (1 - \frac{2}{p})d_*^p + d_*$. We shall show that the following function ψ, defined for all $d \in [0,1]$, is not greater than $1 + 2\kappa$:

$$\psi\colon d \longmapsto 1 - \frac{1}{p}(1-\frac{1}{p})d^{2p} + (1-\frac{2}{p})d^p + d.$$

We consider the cases $p \geq 2$ and $p \leq 2$ separately. In each case, $1 - \frac{1}{p} \geq 0$. Assume that $p \geq 2$. Then $d^p \leq 1$ and $1 - \frac{2}{p} \geq 0$. Therefore $(1-\frac{2}{p})d^p \leq 1 - \frac{2}{p}$, and we obtain

$$\begin{aligned}\psi(d) &\leq 1 + (1-\frac{2}{p})d^p + d \\ &\leq 1 + (1-\frac{2}{p})d + d \\ &= 1 + 2(1-\frac{1}{p})\end{aligned}$$

Now assume that $p \leq 2$. We have the following inequality: $d - d^p \leq p - 1$. Indeed, by considering its derivative, one shows that the application $d \mapsto d - d^p$ is maximum for $d = p^{-\frac{1}{p-1}}$, for which

$$\begin{aligned}d - d^p = d(1-d^{p-1}) &= p^{-\frac{1}{p-1}}(1-p^{-1}) \\ &= p^{-\frac{1}{p-1}-1}(p-1) \\ &= p^{-\frac{p}{p-1}}(p-1) \leq p - 1.\end{aligned}$$

Using $(\frac{2}{p}-1) \geq 0$ and $d^p \geq d-(p-1)$, we obtain $(\frac{2}{p}-1)d^p \geq (\frac{2}{p}-1)d-(\frac{2}{p}-1)(p-1)$. Going back to $\psi(d)$, we have

$$\begin{aligned}\psi(d) &= 1-\frac{1}{p}(1-\frac{1}{p})d^{2p}-(\frac{2}{p}-1)d^p+d \\ &\leq 1-\frac{1}{p}(1-\frac{1}{p})d^{2p}-(\frac{2}{p}-1)d+(\frac{2}{p}-1)(p-1)+d \\ &= 1-\frac{1}{p}(1-\frac{1}{p})d^{2p}+(2-\frac{2}{p})d+(\frac{2}{p}-1)(p-1).\end{aligned}$$

Let us verify that $d \mapsto 1-\frac{1}{p}(1-\frac{1}{p})d^{2p}+2(1-\frac{1}{p})d+(\frac{2}{p}-1)(p-1)$ is increasing. Its derivative is

$$\begin{aligned}-2p\frac{1}{p}(1-\frac{1}{p})d^{2p-1}+2(1-\frac{1}{p}) &\geq -2p\frac{1}{p}(1-\frac{1}{p})+2(1-\frac{1}{p}) \\ &= 0\end{aligned}$$

We deduce that $\psi(d) \leq \psi(1)$ for all $d \in [0,1]$. The value $\psi(1)$ is $1-\frac{1}{p}(1-\frac{1}{p})+2(1-\frac{1}{p})+(\frac{2}{p}-1)(p-1)$. Moreover, we have $-\frac{1}{p}(1-\frac{1}{p})+(\frac{2}{p}-1)(p-1) \leq 0$. Indeed, $-\frac{1}{p}(1-\frac{1}{p})+(\frac{2}{p}-1)(p-1) = -\frac{(p-1)^3}{p^2}$. Therefore $\psi(1) \leq 1+2(1-\frac{1}{p})$.

IV Topological inference for immersed manifolds

Abstract. Given a sample of an abstract manifold immersed in some Euclidean space, we describe a way to recover the singular homology of the original manifold. It consists in estimating its tangent bundle—seen as subset of another Euclidean space—in a measure theoretic point of view, and in applying measure-based filtrations for persistent homology. The construction we propose is consistent and stable, and does not involve the knowledge of the dimension of the manifold.

IV.1 Preliminaries . . . 132
IV.1.1 Model and hypotheses . . . 132
IV.1.2 Index of constants . . . 134
IV.2 Reach of an immersed manifold . . . 135
IV.2.1 Geodesic bounds under curvature conditions . . . 135
IV.2.2 Normal reach . . . 139
IV.2.3 Probabilistic bounds under normal reach conditions . . . 143
IV.2.4 Quantification of the normal reach . . . 150
IV.3 Tangent space estimation . . . 156
IV.3.1 Local covariance matrices and lifted measure . . . 156
IV.3.2 Consistency of the estimation . . . 158
IV.3.3 Stability of the estimation . . . 163
IV.3.4 An approximation theorem . . . 169
IV.4 Topological inference with the lifted measure . . . 171
IV.4.1 Overview of the method . . . 171
IV.4.2 Homotopy type estimation with the DTM . . . 175
IV.4.3 Persistent homology with DTM-filtrations . . . 177
IV.5 Conclusion . . . 180
IV.A Supplementary material for Section IV.2 . . . 180
IV.B Supplementary material for Section IV.3 . . . 184

Numerical experiments. A `Python` notebook, containing numerical illustrations, can be found at `https://github.com/raphaeltinarrage/ImmersedManifolds/blob/master/Demo.ipynb`.

Publication. This work has been presented at the Young Researchers Forum of the Symposium on Computational Geometry, June 2020. A journal-formatted version is available at [Tin19].

Organisation of the chapter. The model we consider is described in Section IV.1. In Section IV.2, we introduce the normal reach, and derive certain probability bounds based on it. In Section IV.3, we study the tangent space estimation of an immersed manifold via local covariance matrices. We gather these results in Section IV.4 to obtain estimation guarantees for our method. For the clarity of the chapter, the proofs of several results have been postponed to the appendices. Please refer to Subsection I.2.3 for an introduction to this chapter.

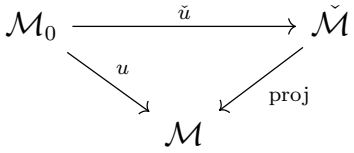
IV.1 Preliminaries

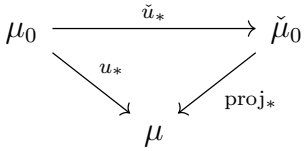
IV.1.1 Model and hypotheses

Model. We consider an abstract $\mathcal{C}^2$-manifold $\mathcal{M}_0$ of dimension d, $E = \mathbb{R}^n$ and a $\mathcal{C}^2$-immersion $u\colon \mathcal{M}_0 \to E$. We denote $\mathcal{M} = u(\mathcal{M}_0)$. Moreover, we write $T_{x_0}\mathcal{M}_0$ for the (abstract) tangent space of $\mathcal{M}_0$ at x_0, and $T_x\mathcal{M}$ for $d_{x_0}u(T_{x_0}\mathcal{M}_0)$, which is an affine subspace of E. Let $\check{u}$ be the application

$$\begin{aligned} \check{u}\colon\ \mathcal{M}_0 &\longrightarrow E \times \mathrm{M}(E) \\ x_0 &\longmapsto (x,\ p_{T_x\mathcal{M}})\,, \end{aligned}$$

where $p_{T_x\mathcal{M}}$ is the orthogonal projection matrix on $T_x\mathcal{M}$, and $\mathrm{M}(E)$ the space of $n \times n$ matrices. We denote $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$. We also consider a probability measure μ_0 on $\mathcal{M}_0$, and define $\mu = u_*\mu_0$ and $\check{\mu}_0 = \check{u}_*\mu_0$. These several sets and measures fit in the following commutative diagrams:





The aim of this work is to estimate the measure $\check{\mu}_0$, from the observation of μ, or a close measure ν. We explain our method in Subsection IV.4.1.

Besides, we endow $\mathcal{M}_0$ with the Riemannian structure given by the immersion u. For every $x_0 \in \mathcal{M}_0$, the second fundamental form of $\mathcal{M}_0$ at x_0 is denoted

$$\mathrm{II}_{x_0}\colon\ T_{x_0}\mathcal{M}_0 \times T_{x_0}\mathcal{M}_0 \longrightarrow (T_x\mathcal{M})^\perp,$$

and the exponential map is denoted

$$\exp_{x_0}^{\mathcal{M}_0}\colon\ T_{x_0}\mathcal{M}_0 \longrightarrow \mathcal{M}_0.$$

We shall also consider the application $\exp_x^{\mathcal{M}}\colon T_x\mathcal{M} \to \mathcal{M}$, the exponential map seen in $\mathcal{M}$, defined as $u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1}$.

Notation conventions. In the rest of this chapter, symbols with 0 as a subscript shall refer to quantities associated to $\mathcal{M}_0$. For instance, a point of $\mathcal{M}_0$ may be denoted x_0, and a curve on $\mathcal{M}_0$ may be denoted γ_0. Symbols with a caron accent shall refer to quantities associated to $\check{\mathcal{M}}$, such as a point $\check{x}$, or a curve $\check{\gamma}$. Symbols

with no such subscript or accent shall refer to quantities associated to $\mathcal{M}$, such as x or γ.

In order to simplify the notations, we consider the following convention:

> Dropping the 0 subscript to a symbol shall correspond to applying the map u. Dropping the 0 subscript to a symbol and adding a caron accent shall correspond to applying the map $\check{u}$.

For instance, if x_0 is a point of $\mathcal{M}_0$, then x represents $u(x_0)$, and $\check{x}$ represents $\check{u}(x_0)$. Note that it is possible to have $x = y$ but $T_x\mathcal{M} \neq T_y\mathcal{M}$. Similarly, if $\gamma_0 \colon I \to \mathcal{M}_0$ is a map, then γ represents $u \circ \gamma$, and $\check{\gamma}$ represents $\check{u} \circ \gamma$.

Hypotheses. We shall refer to the following hypotheses:

> **Hypothesis 1.** For every $x_0, y_0 \in \mathcal{M}_0$ such that $x_0 \neq y_0$ and $x = y$, we have $T_x\mathcal{M} \neq T_y\mathcal{M}$.

> **Hypothesis 2.** The operator norm of the second fundamental form of $\mathcal{M}_0$ at each point is bounded by $\rho > 0$.

> **Hypothesis 3.** The measure μ_0 admits a density f_0 on $\mathcal{M}_0$. Moreover, f_0 is L_0-Lipschitz (with respect to the geodesic distance) and bounded by $f_{\min}, f_{\max} > 0$.

Note that Hypothesis 1 ensures that $\check{u}$ is injective, hence that the set $\check{\mathcal{M}}$ is a submanifold of $E \times \mathrm{M}(E)$. The manifolds $\mathcal{M}_0$ and $\check{\mathcal{M}}$ are $\mathcal{C}^1$-diffeomorphic via $\check{u}$. Hypothesis 2 implies the following key property: if $\gamma_0 \colon I \to \mathcal{M}_0$ is an arc-length parametrized geodesic of class $\mathcal{C}^2$, then for all $t \in I$, we have $\|\ddot{\gamma}(t)\| \leq \rho$ (see Equation (IV.1) in Subsection IV.2.1). Last, in Hypothesis 3, we consider that $\mathcal{M}_0$ is endowed with the natural Hausdorff measure $\mathcal{H}^d_{\mathcal{M}_0}$, obtained by pulling back the d-dimensional Hausdorff measure $\mathcal{H}^d$ on E via the immersion u.

In Subsection IV.2.2, we define an application $\lambda_0 \colon \mathcal{M}_0 \to \mathbb{R}^+$, called the normal reach. The notation λ_0^r refers to the sublevel set $\lambda_0^{-1}([0, r])$. We consider the following hypothesis:

> **Hypothesis 4.** There exists $c_3 \geq 0$ and $r_3 > 0$ such that, for every $r \in [0, r_3)$, $\mu_0(\lambda_0^r) \leq c_3 r$.

We think that this hypothesis is a consequence of Hypotheses 1, 2 and 3, but we have not been able to prove it yet. As a partial result, we prove that it holds when the dimension of $\mathcal{M}_0$ is 1 (Proposition IV.21).

IV.1.2 Index of constants

In this chapter, we will refer to constants that are collected here. In the following list, each constant is preceded by the result where it appeared first. If a constant is defined from the others, it is indicated here. It is not necessary to read this list, since the constants will be introduced throughout the text.

1. (Hypothesis 2) ρ,
2. (Hypothesis 3) L_0, $f_{\min}$ and $f_{\max}$,
3. (Hypothesis 4) c_3 and r_3,
4. (Lemma IV.9) $c_4 \colon t \mapsto \frac{1}{t}\left(1-\sqrt{1-2t}\right)$,
5. (Lemma IV.12) $J_{\min} = (\frac{23}{24})^d$ and $J_{\max} = (\frac{5}{4})^d$,
6. (Remark IV.13) $L = 2L_0$.
7. (Lemma IV.14) $c_7 = 4L_0 J_{\max} + \frac{d}{2}\rho f_{\max}$,
8. (Proposition IV.16), $\begin{cases} c_8 = c_7 + f_{\max} J_{\max} d 2^d \rho \\ \quad = 4L_0 J_{\max} + d\rho f_{\max}(2^{-1} + J_{\max} 2^d), \end{cases}$
9. (Proposition IV.16 and Hypothesis 5) $c_9 = f_{\min} J_{\min} V_d$,
10. (Proposition IV.16 and Hypothesis 6) $c_{10} = d 2^d f_{\max} J_{\max} V_d$,
11. (Proposition IV.17 and Hypothesis 7) $c_{11} = \frac{f_{\max} J_{\max}}{f_{\min} J_{\min}} \left(\frac{\rho}{\sqrt{4-\sqrt{13}}}\right)^d d 2^{2d}\sqrt{3}$,
12. (Subsection IV.2.4) $\mathcal{C}_0$, $\mathcal{D}_0$, Θ and Δ,
13. (Lemma IV.20) $c_{13} = \frac{2}{\sin(\Theta)}$,
14. (Proposition IV.21) $c_{14} = |C_0| f_{\max} J_{\max} c_{13} = |C_0| f_{\max} J_{\max} \frac{2}{\sin(\Theta)}$,
15. (Lemma IV.26) $\begin{cases} c_{15} = 6\rho + 4\dfrac{c_7}{f_{\min} J_{\min}} + \dfrac{f_{\max}}{f_{\min} J_{\min}} 2^d d\rho + \dfrac{c_8}{f_{\min} J_{\min}} \\ \quad = 6\rho + \dfrac{1}{f_{\min} J_{\min}} \left(20 L_0 J_{\max} + f_{\max} d\rho \left(\dfrac{5}{2} + 2^d + 2^d J_{\max}\right)\right), \end{cases}$
16. (Proposition IV.29) $\begin{cases} c_{16} = 4(1 + c_{17}) = 4(4 + c_{25} + c_{26} + c_{27}), \\ c'_{16} = 4c_{28}, \end{cases}$
17. (Lemma IV.31) $c_{17} = 3 + c_{25} + c_{26} + c_{27}$,
18. (Remark IV.32) $c'_{17} = 4 + c_{25} + c_{27}$,
19. (Theorem IV.33) $c_{19} = 2 + \frac{1}{2} c'_{16} = 2(1 + c_{28})$,
20. (Corollary IV.35) $c_{20} = c_{19}(c_3)^{\frac{1}{p}} + c_{16} + c_{15}$,
21. (Subsection IV.4.2) $\check{\rho}_\gamma$, $\check{f}_{\min,\gamma}$ and $\check{c}_{9,\gamma} = \check{f}_{\min,\gamma} J_{\min} V_d$,
22. (Lemma IV.40) $c_{22} = a^{-\frac{1}{d}}$ with $a = \check{c}_{9,\gamma}$,

23. (Corollary IV.41) $c_{23} = 8\mathrm{diam}(\mathrm{supp}(\mu)) + 5$,

24. (Corollary IV.42) $\check{c}_{23,\gamma} = 8\mathrm{diam}(\mathcal{M}) + 8\gamma + 5$ and $\check{c}_{22,\gamma} = (c_{5,\gamma})^{-\frac{1}{d}}$,

25. (Lemma IV.46)
$$\begin{cases} c_{25} = 2(1 + 2^2 5^{d-1} 3^{-d}) \dfrac{c_{10}}{c_9} \\ \quad = d 2^{d+1} (1 + 2^2 5^{d-1} 3^{-d}) \dfrac{f_{\max} J_{\max}}{f_{\min} J_{\min}}, \end{cases}$$

26. (Lemma IV.47)
$$\begin{cases} c_{26} = (2 + 2^{\frac{5}{2}} 5^{d-\frac{1}{2}} 3^{-d}) \dfrac{c_{11}}{c_9} \\ \quad = (2 + 2^{\frac{5}{2}} 5^{d-\frac{1}{2}} 3^{-d}) \dfrac{f_{\max} J_{\max} \left(\frac{\rho}{\sqrt{4-\sqrt{13}}} \right)^d d 2^{2d} \sqrt{3}}{(f_{\min} J_{\min})^2 V_d}, \end{cases}$$

27. (Lemma IV.48)
$$\begin{cases} c_{27} = \dfrac{2^{d-1}}{c_9} + 2 \dfrac{12 \cdot 5^{d-1} c_{10} + 1}{3^d c_9} + 2^{d+3} \dfrac{(\frac{3}{2})^{d-1} c_{10} + 1}{c_9} \\ \quad = \dfrac{2}{f_{\min} J_{\min} V_d} \left[(3^{-d} + 2^{d+2}) + (5^{d-1} 3^{-d} + 3^{d-1}) d 2^{d+3} f_{\max} J_{\max} V_d \right], \end{cases}$$

28. (Lemma IV.49)
$$\begin{cases} c_{28} = \dfrac{2^{d-2}}{c_9} + \dfrac{4 \cdot 3^{\frac{1}{2}} 5^{d-\frac{1}{2}} c_{11} + 4^{d-\frac{1}{2}}}{3^d c_9} + 2 \cdot 4^d \dfrac{2 c_{11} (\frac{3}{2})^{d-\frac{1}{2}} + 1}{3^d c_9} \\ \quad = \dfrac{1}{f_{\min} J_{\min} V_d} \Bigg[(2^{d-2} + 2^{d-1} 3^{-d} 5) \\ \qquad + (2^2 3^{\frac{1}{2}-d} 5^{d-\frac{1}{2}} + 2^{d+\frac{3}{2}} 3^{-\frac{1}{2}}) \dfrac{f_{\max} J_{\max}}{f_{\min} J_{\min}} \left(\dfrac{\rho}{\sqrt{4-\sqrt{13}}} \right)^d d 2^{2d} \sqrt{3} \Bigg]. \end{cases}$$

IV.2 Reach of an immersed manifold

In this section, we introduce a new notion of reach, adapted to the immersed manifolds. Basic facts about the reach can be found in Subsection II.3.3, and about Riemannian geometry in Subsection II.2.5.

IV.2.1 Geodesic bounds under curvature conditions

Before introducing the normal reach, we inspect some technical consequences of Hypothesis 2 that shall be used in the rest of the chapter.

We consider the immersion $u\colon \mathcal{M}_0 \to \mathcal{M} \subset E$ as in Section IV.1. The manifold $\mathcal{M}_0$ is equipped with the Riemannian structure induced by u. For every $x_0 \in \mathcal{M}_0$, the second fundamental form at x_0 is denoted

$$\mathrm{II}_{x_0}\colon\ T_{x_0}\mathcal{M}_0 \times T_{x_0}\mathcal{M}_0 \longrightarrow (T_x\mathcal{M})^\perp.$$

Let $x_0 \in \mathcal{M}_0$ and consider an arc-length parametrized geodesic $\gamma_0\colon I \to \mathcal{M}_0$ such that $\gamma_0(0) = x_0$ and $\dot{\gamma}(0) = v_0$. The following relation can be found in [NSW08, Section 6] or [BLW19, Section 3]:

$$\mathrm{II}_{x_0}(v_0, v_0) = \ddot{\gamma}_0(0).$$

According to Hypothesis 2, the operator norm of II_{x_0} is bounded by ρ. We deduce that

$$\|\ddot{\gamma}_0(0)\| \leq \rho. \tag{IV.1}$$

Denoting $\gamma = u \circ \gamma_0$, we also have $\|\ddot{\gamma}(0)\| \leq \rho$.

The following lemma is based on this observation. Its second point can be seen as an equivalent of Theorem II.12, where the Euclidean distance is replaced with the geodesic distance on $\mathcal{M}_0$, and where the quantity $\frac{1}{\rho}$ plays the role of the reach of $\mathcal{M}$.

Lemma IV.1. *Let $x_0 \in \mathcal{M}_0$ and $\gamma_0 \colon I \to \mathcal{M}_0$ an arc-length parametrized geodesic starting from x_0. Let $\gamma = u \circ \gamma_0$ and $v = \dot{\gamma}(0)$. For all $t \in I$, we have*

- $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2} t^2$.

As a consequence, for every $y_0 \in \mathcal{M}_0$, denoting $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$, we have

- $\mathrm{dist}\,(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2$,
- $(1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$.

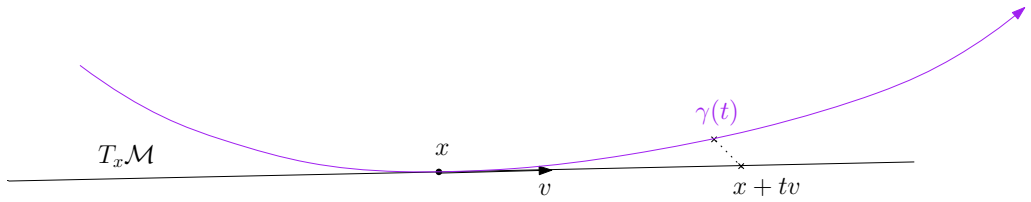


Figure IV.1: Deviation of a geodesic from its initial direction.

Proof. Consider the application $f \colon t \mapsto \|\gamma(t) - (x + tv)\|$. Since γ is a geodesic, it is of class $\mathcal{C}^2$, and Equation (IV.1) gives $\sup_I \|\ddot{\gamma}\| \leq \rho$. We can apply Taylor-Lagrange formula to get $f(t) \leq \sup_I \|\ddot{\gamma}\| \frac{1}{2}t^2 \leq \frac{\rho}{2}t^2$. Therefore, for all $t \in I$, we have $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2}t^2$, and the first claim is proven.

Next, let $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$. By Hopf-Rinow Theorem (Theorem II.7) there exists a length-minimizing geodesic γ_0 from x_0 to y_0. Using the last inequality for $t = \delta$ yields

$$\|y - (x + \delta v)\| = \|\gamma(\delta) - (x + \delta v)\| \leq \frac{\rho}{2}\delta^2,$$

and we deduce that $\mathrm{dist}\,(y - x, T_x\mathcal{M}) \leq \|(y - x) - \delta v\| \leq \frac{\rho}{2}\delta^2$.

We prove the last point by applying the triangular inequality:

$$\|x - y\| \geq \|x - (x + \delta v)\| - \|(x + \delta v) - y\| \geq \delta - \frac{\rho}{2}\delta^2.$$

Remark IV.2. The last point of Lemma IV.1 implies the following fact: for all $x_0 \in \mathcal{M}_0$, the map u is injective on the open (geodesic) ball $\mathcal{B}_{\mathcal{M}_0}(x_0, \frac{2}{\rho})$. Indeed, if $x_0, y_0 \in \mathcal{M}_0$ are such that $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0) < \frac{2}{\rho}$, we get $0 < (1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$, hence $x \neq y$.

Remark IV.3. We can also deduce the following: for every $y_0 \in \mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$ such that $y_0 \neq x_0$, the vector $y - x$ is not orthogonal to $T_x\mathcal{M}$ nor $T_y\mathcal{M}$. To see this, notice that the inequality $\delta < \frac{1}{\rho}$ and the second point of Lemma IV.1 yields

$$\operatorname{dist}(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2 < \frac{1}{2}\delta.$$

Besides, the third point gives $\delta < 2\|y - x\|$, and we deduce that $\operatorname{dist}(y - x, T_x\mathcal{M}) < \|y - x\|$. Equivalently, $y - x$ is not orthogonal to $T_x\mathcal{M}$. Similarly, one proves that $y - x$ is not orthogonal to $T_y\mathcal{M}$.

Consider two points $x_0, y_0 \in \mathcal{M}_0$. We wish to compare their geodesic distance $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ and their Euclidean distance $\|y - x\|$. A first inequality is true in general:

$$\|y - x\| \leq \mathrm{d}_{\mathcal{M}_0}(x_0, y_0).$$

Moreover, if they are close enough in geodesic distance—say $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq \frac{1}{\rho}$ for instance—then Lemma IV.1 third point yields

$$\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq 2\|x - y\|.$$

However, without any assumption on $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$, such an inequality does not hold in general. Figure IV.2 represents a pair of points which are close in Euclidean distance, but far away with respect to the geodesic distance. In the next subsection, we prove an inequality of the form $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0) \leq \text{constant} \cdot \|x - y\|$, but imposing a constraint on $\|x - y\|$ instead of $\mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$ (see Lemma IV.9).

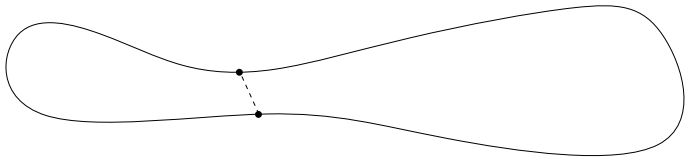


Figure IV.2: Pair of points for which the geodesic distance is large compared to the Euclidean distance.

We now state a technical lemma. It gives how much time it takes for a geodesic to exit a ball. Its proof is deferred to Appendix IV.A (page 180).

Lemma IV.4. *Let $x_0, y_0 \in \mathcal{M}_0$ and $\gamma_0 \colon I \mapsto \mathcal{M}_0$ an arc-length parametrized geodesic with $\gamma_0(0) = y_0$. Define $v = \dot{\gamma}(0)$. Define $l = \|y - x\|$, and let r be such that $l \leq r < \frac{1}{\rho}$. Consider the application $\phi \colon t \in I \mapsto \|\gamma(t) - x\|^2$.*

- *If $\langle v, y - x\rangle \geq 0$, then $\phi > \phi(0)$ on $(0, T_1)$, where $T_1 = \frac{2}{\rho}\sqrt{1 - \rho l}$.*
- *If $\langle v, y - x\rangle = 0$, then ϕ is increasing on $[0, T_2]$ where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2 - \sqrt{3 + \rho^2 l^2}}$.*

Let b be the first value of t such that $\|\gamma(t) - x\| = r$.

- *For all $t \in [0, b]$, we have $\ddot{\phi}(t) \geq 2(1 - \rho r)$.*
- *If $\langle v, y - x\rangle \leq 0$, then $b \geq (1 + \rho r)^{-\frac{1}{2}}\sqrt{r^2 - l^2}$.*

- *If* $\langle v, y-x\rangle \geq 0$, *then* $b \leq \left(\frac{1-\rho r}{2}\right)^{-\frac{1}{2}} \sqrt{r^2-l^2}$. *Note that if* $r < \frac{1}{2\rho}$, *then* $b < 2r < \frac{1}{\rho}$.

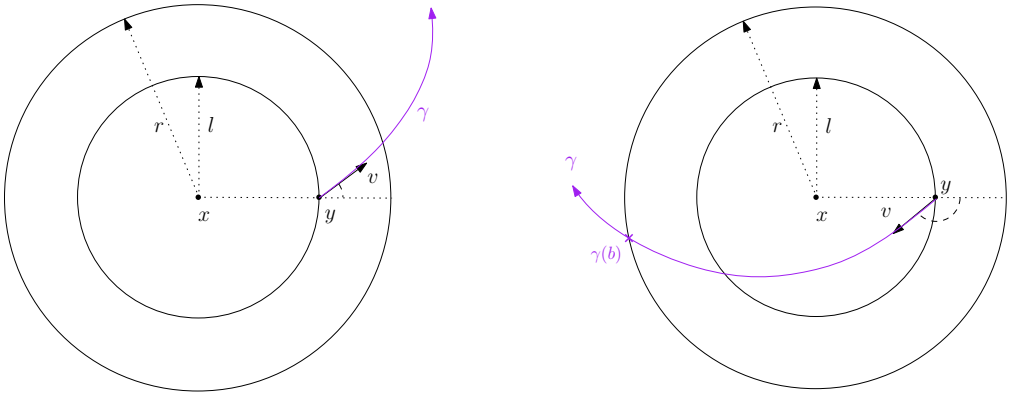


Figure IV.3: Illustration of Lemma IV.4 first point **(left)** and fourth point **(right)**.

We close this subsection by studying the exponential map of $\mathcal{M}_0$, denoted

$$\exp_{x_0}^{\mathcal{M}_0} : T_{x_0}\mathcal{M}_0 \to \mathcal{M}_0.$$

According to [AB06, Corollary 4, Point 1], the map $\exp_{x_0}^{\mathcal{M}_0}$ is injective on the open ball $\mathcal{B}_{T_{x_0}\mathcal{M}_0}(0, \frac{\pi}{\rho})$ of $T_{x_0}\mathcal{M}_0$, and is a diffeomorphism onto its image $\mathcal{B}_{\mathcal{M}_0}(0, \frac{\pi}{\rho})$. We also have a quantitative control of its regularity. Let $x_0 \in \mathcal{M}_0$ and $v_0 \in T_{x_0}\mathcal{M}_0$. The d-dimensional Jacobian of $\exp_{x_0}^{\mathcal{M}_0}$ at v_0 is defined as

$$J_{v_0} = \sqrt{\det\left(A^t A\right)},$$

where $A = d_{v_0} \exp_{x_0}^{\mathcal{M}_0}$ is the differential of the exponential map, seen as a $d \times n$ matrix.

Lemma IV.5. *If* $\|v\| = r < \frac{\pi}{2\sqrt{2}\rho}$, *the Jacobian* J_v *of* $\exp_{x_0}^{\mathcal{M}_0}$ *at* v *satisfies*

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d.$$

Proof. The proof is almost identical to [Aam18, Proposition III.22]. From the Gauss equation [dC92, Theorem 2.5 p 130], we get that the sectional curvature $K(v,w)$ of $\mathcal{M}_0$, with v and w orthonormal vectors in $T_{x_0}\mathcal{M}_0$, satisfies

$$K(v,w) = \langle \mathrm{II}_{x_0}(v,v), \mathrm{II}_{x_0}(w,w)\rangle - \|\mathrm{II}_{x_0}(v,w)\|^2.$$

Using Hypothesis 2, we obtain

$$-2\rho^2 \leq K(v,w) \leq \rho^2.$$

Now, let $v \in T_{x_0}\mathcal{M}_0$ and $w \in T_v\left(T_{x_0}\mathcal{M}_0\right) \simeq T_{x_0}\mathcal{M}_0$. As a consequence of the Rauch theorem [DVW15, Lemma 8], the differential of $\exp_{x_0}^{\mathcal{M}_0}$ at v admits the

following bound:

$$\left(1 - \frac{(\rho\,\|v\|)^2}{6}\right)\|w\| \leq \left\|d_v \exp_{x_0}^{\mathcal{M}_0}(w)\right\| \leq \left(1 + (\rho\,\|v\|)^2\right)\|w\|\,.$$

Next, denote $A = d_v \exp_{x_0}^{\mathcal{M}_0}$, the differential of the exponential map seen as a $d \times n$ matrix. The last inequality shows that the eigenvalues λ of A^tA are bounded by

$$\left(1 - \frac{(\rho\,\|v\|)^2}{6}\right)^2 \leq \lambda \leq \left(1 + (\rho\,\|v\|)^2\right)^2.$$

Since $\det(A^tA)$ is the product of the d eigenvalues of A^tA, we obtain the result.

IV.2.2 Normal reach

We still consider an immersion $u\colon \mathcal{M}_0 \to \mathcal{M} \subset E$ which satisfies Hypothesis 2.

Definition IV.6. For every $x_0 \in \mathcal{M}_0$, let $\Lambda(x_0) = \{y_0 \in \mathcal{M}_0, y_0 \neq x_0, x - y \perp T_y\mathcal{M}\}$. The *normal reach* of $\mathcal{M}_0$ at x_0 is defined as:

$$\lambda_0(x_0) = \inf_{y_0 \in \Lambda(x_0)} \|x - y\|\,.$$

Observe that if x_0, y_0 are distinct points of $\mathcal{M}_0$ with $x = y$, then $x - y$ is orthogonal to any vector, hence $\lambda_0(x_0) = \|x - y\| = 0$.

Moreover, note that $\Lambda(x_0)$ is closed, hence the infimum of Definition IV.6 is attained. Indeed, we can write $\Lambda(x_0) = L \backslash \{x_0\}$, with $L = \{y_0 \in \mathcal{M}_0, x - y \perp T_y\mathcal{M}\}$. L is a closed set since it is the preimage of $\{0\}$ by the continuous map $y_0 \mapsto \left\|p_{T_y\mathcal{M}}(x - y)\right\|$. Furthermore, $\{x_0\}$ is an isolated point of $\Lambda(x_0)$, since Remark IV.3 says that, for every y_0 in the geodesic ball $\mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$ such that $y_0 \neq x_0$, the vector $x - y$ is not orthogonal to $T_y\mathcal{M}$, hence $y_0 \notin L$.

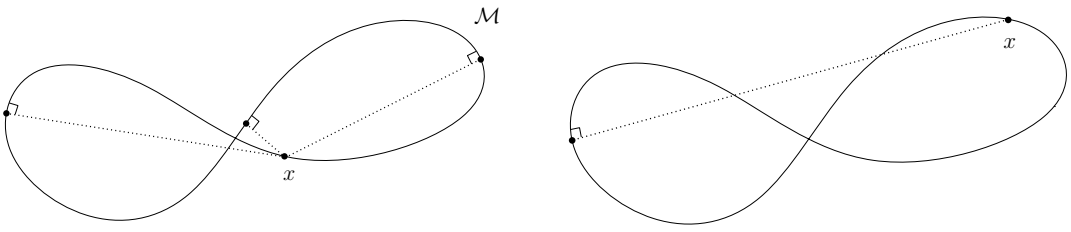


Figure IV.4: The set $\Lambda(x_0)$ from Definition IV.6, for two different points x_0.

Observe that if a point $x \in \mathcal{M}$ has several preimages by u, then for all $x_0 \in u^{-1}(\{x\})$, we have $\lambda_0(x_0) = 0$. Hence we can define the *normal reach seen in* $\mathcal{M}$, denoted $\lambda\colon \mathcal{M} \to \mathbb{R}$, as

$$\lambda(x) = \begin{cases} \lambda_0(u^{-1}(x)) & \text{if } x \text{ has only one preimage,} \\ 0 & \text{else.} \end{cases}$$

It satisfies the relation $\lambda_0 = \lambda \circ u$.

Example IV.7. Suppose that $\mathcal{M}$ is the lemniscate of Bernoulli, with diameter 2. Figure IV.5 represents the values of the normal reach $\lambda\colon \mathcal{M} \to \mathbb{R}$. Observe that λ is not continuous.

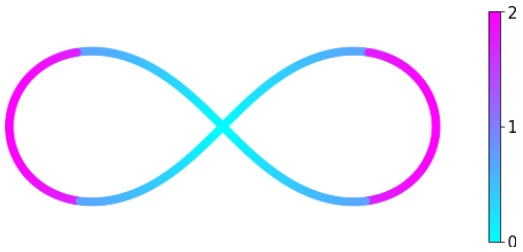


Figure IV.5: Values of the normal reach on the lemniscate of Bernoulli.

Here is a key property of the normal reach:

Lemma IV.8. *Let $x_0 \in \mathcal{M}_0$. Let $r > 0$ such that $r < \lambda(x)$. Then $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}\left(x, r\right)\right)$ is connected.*

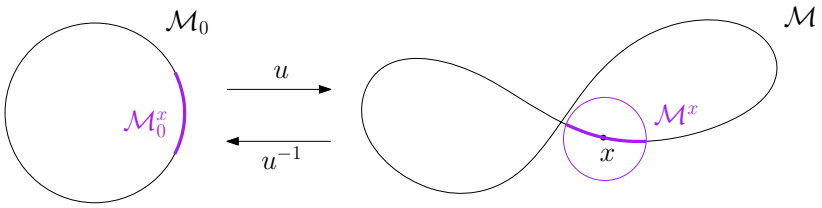


Figure IV.6: The set $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}\left(x, r\right)\right)$, with $r < \lambda(x)$, is connected.

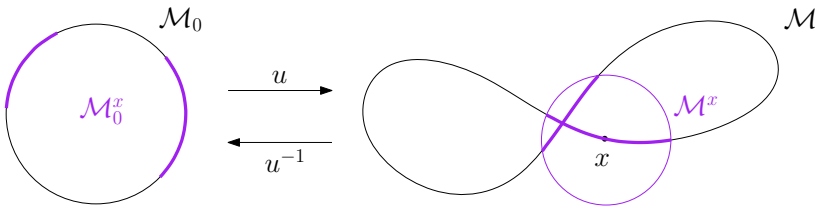


Figure IV.7: The set $u^{-1}\left(\mathcal{M} \cap \overline{\mathcal{B}}\left(x, r\right)\right)$, with $r \geq \lambda(x)$, may not be connected.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}\left(x, r\right) \cap \mathcal{M}$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$. Let us prove that $\mathcal{M}_0^x$ is connected. Suppose that it is not the case. Let $C \subset \mathcal{M}_0^x$ be a connected component which does not contain x_0. Since C is compact, we can consider a minimizer y_0 of $\{\|x - y\|, y_0 \in C\}$. Let us show that $x - y \perp T_y\mathcal{M}$, which will lead to a contradiction.

Two cases may occur: y is in the open ball $\mathcal{B}(x, r)$, or y is on its boundary $\partial\mathcal{B}(x, r)$. If $y \in \mathcal{B}(x, r)$, then there exists a neighborhood $V_0 \subseteq \mathcal{M}_0$ of y such that $V_0 \subseteq \mathcal{M}_0^x$. Hence y satisfies $x - y \perp T_y\mathcal{M}$, otherwise it would not be a local minimizer. Now, suppose that $y \in \partial\mathcal{B}(x, r)$. Since y_0 is a minimizer, there

exists a neighborhood $V_0 \subseteq C$ of y_0 such that $V \cap \mathcal{B}(x,r) = \emptyset$. We deduce the existence of a neighborhood $V_0' \subseteq \mathcal{M}_0$ of y_0 such that $V' \cap \mathcal{B}(x,r) = \emptyset$. For instance, take a ball $\mathcal{B} = \mathcal{B}_{\mathcal{M}_0}(y_0, s)$ such that $\mathcal{B} \cap C \subseteq V_0$, and define $V_0' = \mathcal{B}$. We deduce that $y - x \perp T_y\mathcal{M}$.

To conclude, the properties $x - y \perp T_y\mathcal{M}$ and $x_0 \neq y_0$ imply that $\|x - y\| \geq \lambda(x)$, which contradicts $r < \lambda(x)$.

The following lemma is an equivalent of [NSW08, Proposition 6.3] for the normal reach. It allows to compare the geodesic and Euclidean distance by only imposing a condition on the last one.

Lemma IV.9. *Let $x_0, y_0 \in \mathcal{M}_0$. Denote $r = \|x - y\|$ and $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0, y_0)$. Suppose that $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$. Then*

$$\delta \leq c_4(\rho r) r \qquad \text{where} \qquad c_4(t) = \frac{1}{t}\left(1 - \sqrt{1-2t}\right).$$

In other words, the following inclusion holds: $u^{-1}(\overline{\mathcal{B}}(x,r)) \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_4(\rho r) r)$.

Note that, for $t < \frac{1}{2}$, we have the inequalities $1 \leq c_4(t) \leq 1 + 2t < 2$.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$, $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$ and $\delta = \mathrm{d}_{\mathcal{M}_0}(x,y)$.

Step 1: Let us prove that $\mathcal{M}_0^x \cap \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$, with $\delta_{\min} = c_4(\rho r) r$, where $c_4(t) = \frac{1}{t}\left(1 - \sqrt{1-2t}\right)$ and ϵ is small enough. Choose $y_0 \in \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$. According to Lemma IV.1, we have

$$\|x - y\| \geq \left(1 - \frac{\rho}{2}(\delta_{\min} + \epsilon)\right)\left(\delta_{\min} + \epsilon\right). \tag{IV.2}$$

Consider the polynomial $\phi\colon t \mapsto \left(1 - \frac{\rho}{2}t\right)t - r$. Its discriminant is $1 - 2\rho r > 0$, and we deduce that $\phi(t)$ is positive if and only if

$$t \in \left(\frac{1}{\rho}\left(1 - \sqrt{1-2\rho r}\right), \frac{1}{\rho}\left(1 + \sqrt{1-2\rho r}\right)\right).$$

Observe that the first value $\frac{1}{\rho}\left(1 - \sqrt{1-2\rho r}\right)$ is equal to $c_4(\rho r) r = \delta_{\min}$. Hence $\phi(\delta_{\min} + \epsilon) > 0$ for $0 < \epsilon < \frac{2}{\rho}\sqrt{1-2\rho r}$, and Equation (IV.2) gives $\|x - y\| > r$.

In other words, $y \notin \overline{\mathcal{B}}(x,r)$. This being true for every $y_0 \in \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$, we have $\mathcal{M}_0^x \cap \partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$.

Step 2: Let us deduce that $\mathcal{M}_0^x \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min})$. By contradiction, if a point $z_0 \in \mathcal{M}_0$ with $\|z - x\| > \delta_{\min}$ were to be in $\mathcal{M}_0^x$, it would be in the connected component of x_0 in $\mathcal{M}_0^x$, since it is connected by Lemma IV.8. But since $\mathcal{M}_0$ is a manifold, this would imply the existence of a continuous path from x_0 to z_0 in $\mathcal{M}_0^x$. But such a path would go through a sphere $\partial\mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$, which contradicts Step 1.

The following proposition connects the normal reach to the usual notion of reach.

Proposition IV.10. *Suppose that $u\colon \mathcal{M}_0 \to \mathcal{M} \subset E$ is an embedding. Let $\tau > 0$ be the reach of $\mathcal{M}$. We have*

$$\tau = \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*,$$

where ρ_* is the supremum of the operator norms of the second fundamental forms of $\mathcal{M}_0$, and $\lambda_* = \inf_{x\in\mathcal{M}} \lambda(x)$ is the infimum of the normal reach.

Proof. We first prove that $\tau \geq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$. According to Theorem II.13, two cases may occur: the reach is either caused by a bottleneck or by curvature. In the first case, there exists $x, y \in \mathcal{M}$ and $z \in \operatorname{med}(\mathcal{M})$ with $\|x-y\| = 2\tau$ and $\|x-z\| = \|y-z\| = \tau$. We deduce that $x - y \perp T_y\mathcal{M}$. Hence by definition of $\lambda(x)$,

$$\lambda(x) \leq \|x-y\| = 2\|x-z\| \leq 2\tau.$$

In the second case, there exists $x \in \mathcal{M}$ and an arc-length parametrized geodesic $\gamma\colon I \to \mathcal{M}$ such that $\gamma(0) = x$ and $\|\ddot{\gamma}(0)\| = \frac{1}{\tau}$. But $\|\ddot{\gamma}(0)\| \leq \rho_*$, hence $\frac{1}{\tau} \leq \rho_*$.

This disjunction shows that $\tau \geq \frac{1}{\rho} \wedge \frac{1}{2}\lambda_{\min}$.

We now prove that $\tau \leq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$. The inequality $\tau \leq \frac{1}{\rho_*}$ appears in [NSW08, Proposition 6.1]. To prove $\tau \leq \frac{1}{2}\lambda_*$, consider any $x_0 \in \mathcal{M}_0$. Let $y_0 \in \Lambda(x_0)$ such that $\|x-y\|$ is minimal. Using Theorem II.12 and the property $x - y \perp T_y\mathcal{M}$, we immediately have

$$\tau \leq \frac{\|x-y\|^2}{2\operatorname{dist}(y-x, T_y\mathcal{M})} = \frac{\|x-y\|}{2} = \frac{\lambda(x)}{2}.$$

In the case where u is not an embedding, $\mathcal{M}$ may have zero reach. However, as shown by the following theorem, the normal reach gives a scale at which $\mathcal{M}$ still behaves well. Note that we shall not make use of this result in the rest of the chapter.

Theorem IV.11. *Assume that $\mathcal{M}_0$ satisfies Hypothesis 2. Let $x \in \mathcal{M}_0$ and $r < \frac{1}{4\rho} \wedge \lambda(x)$. Then $\overline{\mathcal{B}}(x,r) \cap \mathcal{M}$ is a set of reach at least $\frac{1-2\rho r}{\rho}$. In particular, it is greater than $\frac{1}{2\rho}$.*

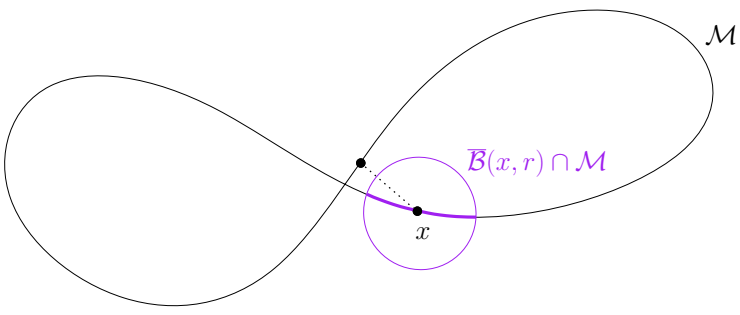


Figure IV.8: The set $\overline{\mathcal{B}}(x,r) \cap \mathcal{M}$ has positive reach.

Proof. Denote $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$.

Step 1: Let us prove that for every $y_0, z_0 \in \mathcal{M}_0^x$,

$$\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2 .$$

Let $y_0, z_0 \in \mathcal{M}_0^x$, and $\delta = \mathrm{d}_{\mathcal{M}_0}(y_0, z_0)$. Lemma IV.1 Point 3 gives $\delta \leq \frac{1}{1-\frac{\rho}{2}\delta} \|y-z\|$. Moreover, $\delta \leq \mathrm{d}_{\mathcal{M}_0}(y_0, x_0) + \mathrm{d}_{\mathcal{M}_0}(x_0, z_0) \leq 2c_4(\rho r)r$. Hence,

$$\frac{1}{1-\frac{\rho}{2}\delta} \leq \frac{1}{1-c_4(\rho r)\rho r} = \frac{1}{\sqrt{1-2\rho r}},$$

and we deduce that

$$\delta \leq \frac{1}{\sqrt{1-2\rho r}} \|y-z\| . \tag{IV.3}$$

Besides, Lemma IV.1 Point 2 gives $\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2}\delta^2$, and combining these two inequalities yields $\operatorname{dist}(z-y, T_y\mathcal{M}) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2$.

Step 2: Let us prove that

$$\operatorname{dist}(z-y, \operatorname{Tan}(\mathcal{M}^x, y)) \leq \frac{\rho}{2(1-2\rho r)} \|z-y\|^2 , \tag{IV.4}$$

where $\operatorname{Tan}(\mathcal{M}^x, y)$ is the tangent cone at y of the closed set $\mathcal{M}^x$.

If $y \in \mathcal{B}(x,r)$, then $\operatorname{Tan}(\mathcal{M}^x, y) = T_y\mathcal{M}$, and the inequality follows from Step 1. Otherwise, suppose that $y \in \partial\mathcal{B}(x,r)$ and that $z \neq y$. Let $\delta = \mathrm{d}_{\mathcal{M}_0}(y_0, z_0)$. According to Equation (IV.3), the inequality $\|y-z\| \leq 2r$ and the assumption $r < \frac{1}{4\rho}$, we have $\delta < \frac{1}{\rho}$. Consider a length-minimizing geodesic $\gamma_0 \colon [0,\delta] \to \mathcal{M}_0$ from y_0 to z_0, and denote $v = \dot{\gamma}(0)$. Let us show that $v \in \operatorname{Tan}(\mathcal{M}^x, y)$, and we will conclude with Step 1.

Since $\mathcal{M}^x = \overline{\mathcal{B}}(x,r) \cap \mathcal{M}$, $v \in \operatorname{Tan}(\mathcal{M}^x, y)$ is implied by $\langle v, y-x\rangle < 0$. Suppose by contradiction that $\langle v, y-x\rangle \geq 0$. Hence, according to Lemma IV.4 Point 1, with $l = r < \frac{1}{2\rho}$, we have $T_1 = \frac{2}{\rho}\sqrt{1-\rho l} > \frac{\sqrt{2}}{\rho} > \delta$, and

$$\|z-x\| = \|\gamma(\delta)-x\| > \|\gamma(0)-x\| = \|y-x\| = r.$$

We deduce the contradiction $z \notin \overline{\mathcal{B}}(x,r)$.

To conclude the proof, it follows from Theorem II.12 and Equation (IV.4) that $\mathcal{M}^x$ has reach at least $\frac{1-2\rho r}{\rho}$.

IV.2.3 Probabilistic bounds under normal reach conditions

We now consider $\mathcal{M}_0$ and μ_0 which satisfy the Hypotheses 2 and 3. The aim of this subsection is to provide a quantitative control of the measure $\mu = u_*\mu_0$ (Propositions IV.16 and IV.17). We do so by pulling-back μ on the tangent spaces $T_x\mathcal{M}$, where it is simpler to compute integrals (Lemma IV.14).

Recall that the exponential map of $\mathcal{M}_0$ at a point x_0 is denoted

$$\exp_{x_0}^{\mathcal{M}_0} : T_{x_0}\mathcal{M}_0 \to \mathcal{M}_0.$$

To ease the reading of this subsection, we introduce the exponential map seen in $\mathcal{M}$, denoted $\exp_x^{\mathcal{M}} : T_x\mathcal{M} \to \mathcal{M}$. It is defined as

$$\exp_x^{\mathcal{M}} = u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1}.$$

It fits in the following commutative diagram:

$$\begin{array}{ccc} \mathcal{M}_0 & \xrightarrow{u} & \mathcal{M} \\ {\scriptstyle \exp_{x_0}^{\mathcal{M}_0}}\uparrow & & \uparrow{\scriptstyle \exp_x^{\mathcal{M}}} \\ T_{x_0}\mathcal{M}_0 & \xrightarrow{d_{x_0}u} & T_x\mathcal{M} \end{array}$$

We also define the map $\overline{\exp}_x^{\mathcal{M}}$ as the restriction of $\exp_x^{\mathcal{M}}$ to the closed ball $\overline{\mathcal{B}}_{T_x\mathcal{M}}(0, \frac{\pi}{\rho})$. It is injective by Lemma IV.5. The next lemma gather results of the last subsections. The d-dimensional Jacobian of $\overline{\exp}_x^{\mathcal{M}}$ at v is defined as

$$J_v = \sqrt{\det\left(A_v^t A_v\right)},$$

where $A_v = d_v\overline{\exp}_x^{\mathcal{M}}$ is the differential of the exponential map seen as a $d \times n$ matrix.

Lemma IV.12. *Let $x_0 \in \mathcal{M}_0$ and $r < \frac{1}{2\rho} \wedge \lambda(x)$. Denote $\overline{\mathcal{B}} = \overline{\mathcal{B}}(x, r)$ and $\overline{\mathcal{B}}^T = \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}\right)$. We have the inclusions*

$$\overline{\mathcal{B}}_{T_x\mathcal{M}}(0, r) \;\subseteq\; \overline{\mathcal{B}}^T \;\subseteq\; \overline{\mathcal{B}}_{T_x\mathcal{M}}(0, c_4(\rho r)r).$$

Moreover, for all $v \in \overline{\mathcal{B}}^T$, the Jacobian J_v of $\overline{\exp}_x^{\mathcal{M}}$, is bounded by

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d,$$

and these terms are bounded by $J_{\min} = (\frac{23}{24})^d$ and $J_{\max} = (\frac{5}{4})^d$.

Proof. The inclusions come from Lemma IV.9. The bounds on the Jacobian come from Lemma IV.5 and the fact that $c_4(\rho r)r \leq 2r \leq \frac{1}{\rho} \leq \frac{\pi}{2\sqrt{2}\rho}$ when $r < \frac{1}{2\rho}$.

We now study the measure μ. An application of the coarea formula shows that μ admits the following density against $\mathcal{H}_{\mathcal{M}}^d$, the d-dimensional Hausdorff measure restricted to $\mathcal{M}$:

$$f(x) = \sum_{x_0 \in u^{-1}(\{x\})} f_0(x_0).$$

In particular, if x has only one preimage by u—i.e., if $\lambda(x) > 0$—then $f(x) = f_0 \circ u^{-1}(x)$. In the rest of the chapter, we will only use f on points x such that $\lambda(x) > 0$.

Remark IV.13. Recall that, by Hypothesis 3, the density f_0 is L_0-Lipschitz with respect to the geodesic distance: for all $x_0, y_0 \in \mathcal{M}_0$,

$$|f_0(x_0) - f_0(y_0)| \leq L_0 \cdot \mathrm{d}_{\mathcal{M}_0}(x_0, y_0).$$

We can deduce the following: for all $x_0, y_0 \in \mathcal{M}_0$ such that $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$, we have

$$|f(x) - f(y)| \leq L \|x - y\|$$

with $L = 2L_0$. To prove this, we start with the case where y has only one preimage by u. Since $\|x - y\| < \lambda(x)$ by assumption, we have $0 < \lambda(x)$, hence x also has only one preimage. Now we can write

$$\begin{aligned} |f(x) - f(y)| &= \left|f_0 \circ u^{-1}(x) - f_0 \circ u^{-1}(y)\right| \\ &\leq L_0 \cdot \mathrm{d}_{\mathcal{M}_0}(u^{-1}(x), u^{-1}(y)) \\ &\leq 2L_0 \|x - y\|, \end{aligned}$$

where we used Lemma *IV*.9 on the last inequality. Now we prove that $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$ implies that y has only one preimage. Let $r = \|x - y\|$, and suppose by contradiction that y_0, y_1 are two distinct preimages. According to Remark IV.2, $\mathrm{d}_{\mathcal{M}_0}(y_0, y_1) \geq \frac{2}{\rho}$. But Lemma IV.9 says that $u^{-1}(\mathcal{B}(x, r)) \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, 2r) \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$, which contradicts $\mathrm{d}_{\mathcal{M}_0}(y_0, y_1) \geq \frac{2}{\rho}$.

Lemma IV.14. *Let $x_0 \in \mathcal{M}_0$ and $r < \frac{1}{2\rho} \wedge \lambda(x)$. Consider μ_x, the measure μ restricted to $\overline{\mathcal{B}}(x, r)$, and define*

$$\nu_x = \left(\overline{\exp}_x^{\mathcal{M}}\right)_*^{-1} \mu_x.$$

The measure ν_x admits the following density against the d-dimensional Hausdorff measure on $T_x\mathcal{M}$:

$$g(v) = f\left(\overline{\exp}_x^{\mathcal{M}}(v)\right) \cdot J_v \cdot 1_{\overline{\mathcal{B}}^T}(v).$$

Moreover, for all $v \in \overline{\mathcal{B}}^T$, the map g satisfies

$$|g(v) - g(0)| \leq c_7 r,$$

where $c_7 = 4L_0 J_{\max} + \frac{d}{2}\rho f_{\max}$.

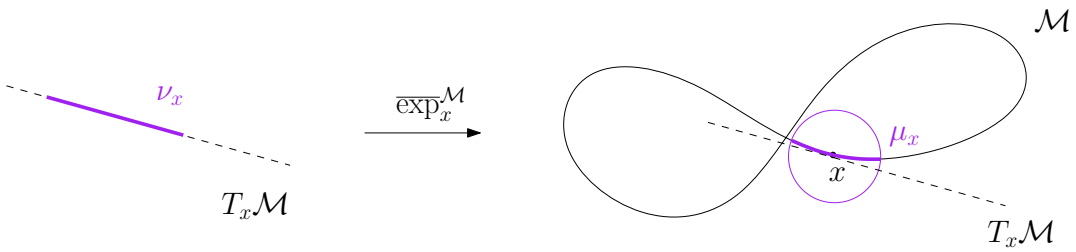


Figure IV.9: Measures involved in Lemma IV.14.

Proof. The expression of g comes from the area formula [Fed14, Theorem 3.2.5]. To prove the inequality, observe that we can decompose

$$\begin{aligned} g(v) - g(0) &= f\big(\exp_x^{\mathcal{M}}(v)\big) J_v - f\big(\exp_x^{\mathcal{M}}(0)\big) J_0 \\ &= \Big[f\big(\exp_x^{\mathcal{M}}(v)\big) - f\big(\exp_x^{\mathcal{M}}(0)\big) \Big] J_v + (J_v - J_0)\, f\big(\exp_x^{\mathcal{M}}(0)\big) \end{aligned}$$

On the one hand, using Remark IV.13, we get

$$\begin{aligned} \left| f\left(\overline{\exp}_x^{\mathcal{M}}(v)\right) - f\left(\overline{\exp}_x^{\mathcal{M}}(0)\right) \right| &\leq L \left\| \overline{\exp}_x^{\mathcal{M}}(v) - \overline{\exp}_x^{\mathcal{M}}(0) \right\| \\ &= L \left\| u \circ \exp_{x_0}^{\mathcal{M}_0}(v) - u \circ \exp_{x_0}^{\mathcal{M}_0}(0) \right\| \\ &\leq L \cdot \mathrm{d}_{\mathcal{M}_0}(\overline{\exp}_{x_0}^{\mathcal{M}_0}(v), x_0) = L \left\| v \right\|. \end{aligned}$$

On the other hand, $J_0 = 1$ and $\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq \left(1 + (r\rho)^2\right)^d$ yield $|J_v - J_0| \leq d(\rho r)^2 \leq \frac{d}{2}\rho r$. We eventually obtain

$$g(v) - g(0) \leq L \left\| v \right\| J_{\max} + f_{\max} \frac{d}{2} \rho r \leq \left(2L J_{\max} + f_{\max} \frac{d}{2} \rho \right) r.$$

Remark IV.15. In the same vein as Lemma IV.14, define $\overline{\exp}_{x_0}^{\mathcal{M}_0}$ to be the map $\exp_{x_0}^{\mathcal{M}_0}$ restricted to $\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}(0, \frac{\pi}{\rho})$. For any $x_0 \in \mathcal{M}_0$, let $\mu_0^{x_0}$ be the measure μ_0 restricted to $\overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$, and define the measure

$$\nu_0 = (\overline{\exp}_{x_0}^{\mathcal{M}_0})^{-1} \mu_0^{x_0}.$$

Using the area formula, one shows that ν_0 admits the following density over the d-dimensional Hausdorff measure on $T_{x_0}\mathcal{M}_0$:

$$g_0(v) = f_0\left(\overline{\exp}_{x_0}^{\mathcal{M}_0}(v)\right) \cdot J_v \cdot 1_{\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}(0, \frac{1}{\rho})}(v).$$

Now we can use the density g of Lemma IV.14 to derive explicit bounds on μ.

Proposition IV.16. *Let $x_0 \in \mathcal{M}_0$, $r \leq \frac{1}{2\rho} \wedge \lambda(x)$ and $s \in [0, r]$. We have*

- $\mu\left(\overline{\mathcal{B}}(x, r)\right) \geq c_9 r^d$
- $\left| \frac{\mu(\overline{\mathcal{B}}(x,r))}{V_d r^d} - f(x) \right| \leq c_8 r$
- $\mu\left(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)\right) \leq c_{10} r^{d-1}(r - s)$

with $c_9 = f_{\min} J_{\min} V_d$, $c_8 = c_7 + f_{\max} J_{\max} d 2^d \rho$ and $c_{10} = d 2^d f_{\max} J_{\max} V_d$.

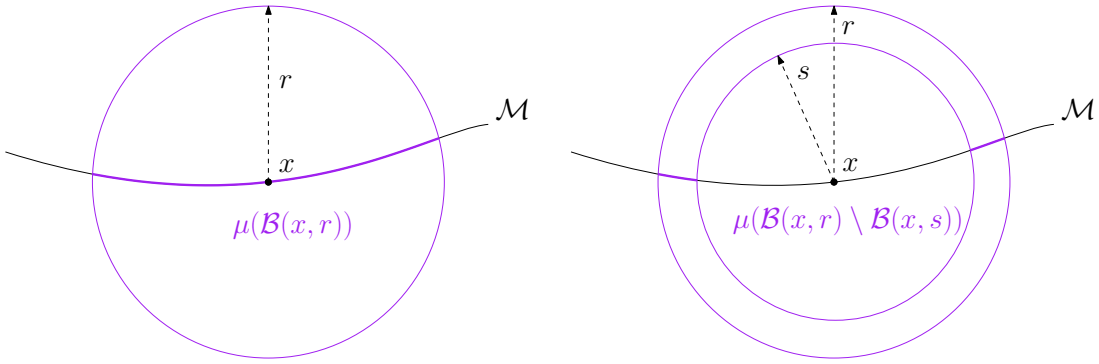


Figure IV.10: Representation of Proposition IV.16 first point **(left)** and third point **(right)**.

Proof. Consider the map $\overline{\exp}_x^{\mathcal{M}}$ and the measure ν_x as defined in Lemma IV.14. In the following, we write $T = T_x\mathcal{M}$, and $\overline{\mathcal{B}}^T = \left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}\,(x,r)\right)$.

Point (1): We have $\mu\left(\overline{\mathcal{B}}\,(x,r)\right) = \nu_x\left(\overline{\mathcal{B}}^T\right)$. Writing down the density g of ν_x yields

$$\nu_x\left(\overline{\mathcal{B}}^T\right) = \int_{\overline{\mathcal{B}}^T} g(v)\mathrm{d}\mathcal{H}^d(v).$$

According to the expression of g in Lemma IV.14, we have $g \geq f_{\min}J_{\min}$. Therefore,

$$\int_{\overline{\mathcal{B}}^T} g(v)\mathrm{d}\mathcal{H}^d(v) \geq \int_{\overline{\mathcal{B}}^T} f_{\min}J_{\min}\mathrm{d}\mathcal{H}^d(v) = f_{\min}J_{\min}\mathcal{H}^d\left(\overline{\mathcal{B}}^T\right).$$

Besides, since $\overline{\mathcal{B}}^T \supset \overline{\mathcal{B}}_T(0,r)$, we have

$$\mathcal{H}^d\left(\overline{\mathcal{B}}^T\right) \geq \mathcal{H}^d\left(\overline{\mathcal{B}}_T(0,r)\right) = V_d r^d.$$

We finally obtain $\nu_x\left(\overline{\mathcal{B}}^T\right) \geq f_{\min}J_{\min}V_d r^d$.

Point (2): Observe that $\int_{\overline{\mathcal{B}}_T(0,r)} f(x)\mathrm{d}\mathcal{H}^d(v) = f(x)V_d r^d$. Hence

$$\begin{aligned}\left|\mu(\overline{\mathcal{B}}\,(x,r)) - f(x)V_d r^d\right| &= \left|\int_{\overline{\mathcal{B}}^T} g(v)\mathrm{d}\mathcal{H}^d(v) - \int_{\overline{\mathcal{B}}_T(0,r)} f(x)\mathrm{d}\mathcal{H}^d(v)\right| \\ &\leq \underbrace{\left|\int_{\overline{\mathcal{B}}_T(0,r)} (f(x) - g(v))\mathrm{d}\mathcal{H}^d(v)\right|}_{(1)} + \underbrace{\left|\int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v)\mathrm{d}\mathcal{H}^d(v)\right|}_{(2)}.\end{aligned}$$

To bound Term (1), notice that $g(0) = f(\exp_x^{\mathcal{M}}(0))J_0 = f(x)$. Hence we can write:

$$\left|\int_{\overline{\mathcal{B}}_T(0,r)} (f(x) - g(v))\mathrm{d}\mathcal{H}^d(v)\right| \leq \int_{\overline{\mathcal{B}}_T(0,r)} \left|g(0) - g(v)\right|\mathrm{d}\mathcal{H}^d(v).$$

Now, Lemma IV.14 gives $|g(v)-g(0)| \leq c_7 r$, and we obtain $\left|\int_{\overline{\mathcal{B}}_T(0,r)}(f(x)-g(v))\mathrm{d}\mathcal{H}^d(v)\right| \leq c_7 r V_d r^d$.

On the other hand, we bound Term (2) thanks to the inclusion $\overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_T(0, c_4(\rho r)r)$. Denote $\mathcal{A} = \overline{\mathcal{B}}_T(0, c_4(\rho r)r) \setminus \overline{\mathcal{B}}_T(0,r)$. We have $\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r) \subset \mathcal{A}$, hence

$$\begin{aligned}\int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v)\mathrm{d}\mathcal{H}^d(v) &\leq \int_{\mathcal{A}} g(v)\mathrm{d}\mathcal{H}^d(v) \\ &\leq f_{\max} J_{\max} \mathcal{H}^d(\mathcal{A}).\end{aligned}$$

Moreover, we have

$$\begin{aligned}\mathcal{H}^d(\mathcal{A}) &= \mathcal{H}^d\left(\overline{\mathcal{B}}_T(0, c_4(\rho r)r)\right) - \mathcal{H}^d\left(\overline{\mathcal{B}}_T(0,r)\right) \\ &= V_d\left(c_4(\rho r)^d - 1\right) r^d.\end{aligned}$$

We can use $c_4(\rho r) \leq 1 + 2\rho r \leq 2$ and the inequality $A^d - 1 \leq d(A-1)A^{d-1}$, where $A \geq 1$, to get

$$\begin{aligned}\left(c_4(\rho r)^d - 1\right) &\leq d \cdot (c_4(\rho r) - 1) \cdot c_4(\rho r)^{d-1} \\ &\leq d \cdot 2\rho r \cdot 2^{d-1}.\end{aligned}$$

We finally deduce the following bound on Term (2):

$$\int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)} g(v)\mathrm{d}\mathcal{H}^d(v) \leq f_{\max} J_{\max} V_d r^d d \cdot \rho r 2^d.$$

Gathering Term (1) and (2), we obtain

$$\left|\mu(\overline{\mathcal{B}}(x,r)) - f(x)V_d r^d\right| \leq r\left(c_7 + f_{\max} J_{\max} d\rho 2^d\right) V_d r^d.$$

Point (3): Let us write

$$\begin{aligned}\mu\left(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)\right) &= \nu_x\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)\right)\right) \\ &= \int_{\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)} g(v)\mathrm{d}\mathcal{H}^d(v).\end{aligned}$$

In spherical coordinates, this integral reads

$$\int_{\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)} g(v)\mathrm{d}\mathcal{H}^d(v) = \int_{v\in\partial\mathcal{B}_T(0,1)} \int_{t=a(v)}^{b(v)} g(tv)t^{d-1}\mathrm{d}t\mathrm{d}v, \quad \text{(IV.5)}$$

where a and b are defined as follows: for every $v \in \partial\mathcal{B}_T(0,1) \subset T_x\mathcal{M}$, let γ_0 be a arc-length parametrized geodesic with $\gamma(0) = x$ and $\dot{\gamma}(0) = v$, and set $a(v)$ and $b(v)$ to be the first positive values such that $\|\gamma(a(v)) - x\| = s$ and $\|\gamma(b(v)) - x\| = r$.

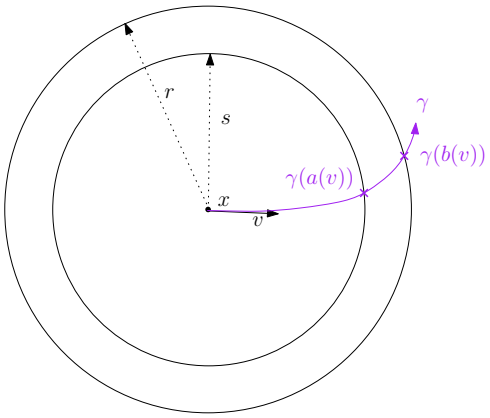


Figure IV.11: Illustration of $a(v)$ and $b(v)$ in the proof of Proposition IV.16.

Let us show that

$$b(v) - a(v) \leq \frac{1}{1-\rho r}(r-s) \tag{IV.6}$$

Consider the application $\phi\colon t \mapsto \|\gamma(t) - x\|^2$. According to Lemma IV.4 Point 3 with $l = 0$, we have $\ddot{\phi}(t) \geq 2(1-\rho r)$ for $t \in [0, b(v)]$. It follows that $\dot{\phi}(t) \geq 2(1-\rho r)t$, and that

$$\begin{aligned}\phi(b(v)) - \phi(a(v)) = \int_{a(v)}^{b(v)} \dot{\phi}(t)\mathrm{d}t &\geq \int_{a(v)}^{b(v)} 2(1-\rho r)t\mathrm{d}t \\ &= (1-\rho r)(b(v)^2 - a(v)^2).\end{aligned}$$

Since $r^2 - s^2 = \phi(b(v)) - \phi(a(v))$, we deduce that

$$r^2 - s^2 \geq (1-\rho r)(b(v)^2 - a(v)^2). \tag{IV.7}$$

Writing $r^2 - s^2 = (r+s)(r-s)$ and $b(v)^2 - a(v)^2 = \big(b(v)+a(v)\big)\big(b(v)-a(v)\big)$ leads to

$$(r-s)\frac{1}{1-\rho r}\frac{r+s}{b(v)+a(v)} \geq b(v) - a(v).$$

But $b(v) + a(v) \geq r + s$, hence $(r-s)\frac{1}{1-\rho r} \geq b(v) - a(v)$, as wanted.

Now, notice that we have $b(v) \leq 2r$. Indeed, $b < \frac{1}{\rho}$ by Lemma IV.4 Point 5 with $l = 0$, and we conclude with Lemma IV.1 Point 2. Hence we have

$$\int_{t=a(v)}^{b(v)} g(tv)t^{d-1}\mathrm{d}t \leq \int_{t=a(v)}^{b(v)} f_{\max}J_{\max}(2r)^{d-1}\mathrm{d}t.$$

Using Equation (IV.6), we get

$$\begin{aligned}\int_{t=a(v)}^{b(v)} f_{\max}J_{\max}(2r)^{d-1}\mathrm{d}t &= (b(v)-a(v))f_{\max}J_{\max}(2r)^{d-1}\mathrm{d}t \\ &\leq \frac{1}{1-\rho r}(r-s)f_{\max}J_{\max}(2r)^{d-1}.\end{aligned}$$

From these last two equations we deduce

$$\int_{v\in\partial\mathcal{B}(0,1)}\int_{t=a(v)}^{b(v)} g(tv)t^{d-1}\mathrm{d}t\mathrm{d}v \leq \frac{1}{1-\rho r}(r-s)f_{\max}J_{\max}(2r)^{d-1}\int_{v\in\partial\mathcal{B}(0,1)}\mathrm{d}v$$
$$= \frac{1}{1-\rho r}(r-s)f_{\max}J_{\max}(2r)^{d-1}\cdot dV_d.$$

Going back to Equation (IV.5), we obtain

$$\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right) = \frac{2^{d-1}dV_d f_{\max}J_{\max}}{1-\rho r}(r-s)r^{d-1},$$

and we conclude with $r\leq\frac{1}{2\rho}$:

$$\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right) = 2^d dV_d f_{\max}J_{\max}(r-s)r^{d-1}.$$

The following proposition is a weaker form of Proposition IV.16, without normal reach condition. Its proof, based on the same ideas, is given in Appendix IV.A (page 181).

Proposition IV.17. *Let $x_0\in\mathcal{M}_0$, $r\leq\frac{1}{2\rho}$ and $s\in[0,r]$. We have*

- $\mu\left(\overline{\mathcal{B}}(x,r)\right)\geq c_9 r^d$
- $\mu\left(\overline{\mathcal{B}}(x,r)\setminus\overline{\mathcal{B}}(x,s)\right)\leq c_{11}r^{d-\frac{1}{2}}(r-s)^{\frac{1}{2}}$

with $c_9 = f_{\min}J_{\min}V_d$ and $c_{11} = \frac{f_{\max}J_{\max}}{f_{\min}J_{\min}}(\frac{\rho}{\sqrt{4-\sqrt{13}}})^d d2^{2d}\sqrt{3}$.

IV.2.4 Quantification of the normal reach

In this subsection, we suppose that the dimension of the manifold $\mathcal{M}_0$ is $d=1$, and we assume the Hypotheses 1, 2 and 3. We give an upper bound on the measure $\mu_0(\lambda_0^t)$, i.e., the measure of points $x_0\in\mathcal{M}_0$ with normal reach not greater than t. This proves a result announced in Section IV.1: Hypothesis 4 is a consequence of Hypotheses 1, 2 and 3.

We shall use two quantities related to the immersion $\mathcal{M}_0$. Let $\mathcal{D}_0$ be the set of critical points of the Euclidean distance on $\mathcal{M}_0$, that is,

$$\mathcal{D}_0 = \{(x_0,y_0)\in\mathcal{M}_0,\ x_0\neq y_0,\ x-y\perp T_y\mathcal{M}\ \text{ and }\ x-y\perp T_x\mathcal{M}\}. \tag{IV.8}$$

Also, let $\mathcal{C}_0$ be the set of self-intersections of $\mathcal{M}_0$:

$$\mathcal{C}_0 = \{(x_0,y_0)\in\mathcal{M}_0,\ x_0\neq y_0\ \text{ and }\ x=y\}. \tag{IV.9}$$

As a consequence of Remark IV.2 and the compacity of $\mathcal{M}_0$, the set $\mathcal{C}_0$ is finite. For every $(x_0,y_0)\in\mathcal{C}_0$, let $\theta(x_0,y_0)\in\left[0,\frac{\pi}{2}\right]$ be the angle formed by the lines $T_x\mathcal{M}$ and $T_y\mathcal{M}$. Define

$$\Theta = \inf\{\theta(x_0,y_0),\ (x_0,y_0)\in\mathcal{C}_0\}. \tag{IV.10}$$

Note that, according to Hypothesis 1, we have $\Theta > 0$. Besides, on the set $\mathcal{D}_0 \setminus \mathcal{C}_0$, consider the quantity

$$\Delta = \inf \{ \|x - y\|, \ (x_0, y_0) \in \mathcal{D}_0 \setminus \mathcal{C}_0 \}. \tag{IV.11}$$

Since $\mathcal{C}_0$ consists of isolated points of $\mathcal{D}_0$, the set $\mathcal{D}_0 \setminus \mathcal{C}_0$ is closed, hence the previous infimum is attained. Therefore, $\Delta > 0$.

In order to bound the measure $\mu_0(\lambda_0^t)$, we first prove that the sublevel set λ_0^t is included in a thickening of $\mathcal{C}_0$ (Lemma IV.20). By bounding the measure of this thickening, we obtain the main result of this subsection (Proposition IV.21). We start with a lemma which describes the situation around self-intersection points of $\mathcal{M}_0$.

Lemma IV.18. *Let $(x_0^*, y_0^*) \in \mathcal{C}_0$. Denote by θ the angle formed by the lines $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$. Let $x_0, y_0 \in \mathcal{M}_0$. Denote $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0^*, x_0)$ and $\delta' = \mathrm{d}_{\mathcal{M}_0}(y_0^*, y_0)$. If $\delta' \leq \delta \leq \frac{\sin(\theta)}{2\rho}$, then $\|x - y\| \geq \frac{\sin(\theta)}{2}\delta$.*

Proof. Let γ_0 be an arc-length parametrized geodesic connecting x_0^* to x_0, and η_0 connecting y_0^* to y_0. Let $v_0 = \dot{\gamma}_0(0)$ and $\overline{x} = x^* + \delta v$. Accordingly, let $w_0 = \dot{\eta}_0(0)$ and $\overline{y} = y^* + \delta' w = x^* + \delta' w$.

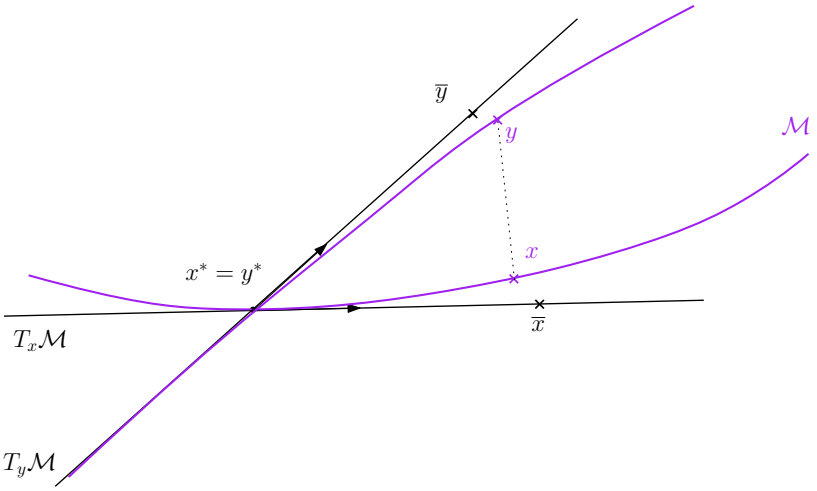


Figure IV.12: Situation in Lemma IV.18.

The triangular inequality yields

$$\|x - y\| \geq \|\overline{x} - \overline{y}\| - \|x - \overline{x}\| - \|y - \overline{y}\|.$$

According to Lemma IV.1, we have $\|x - \overline{x}\| \leq \frac{\rho}{2}\delta^2$ and $\|y - \overline{y}\| \leq \frac{\rho}{2}\delta'^2 \leq \frac{\rho}{2}\delta^2$. Moreover, $\|\overline{x} - \overline{y}\|$ is not lower than $\|\overline{x} - z\|$, where z is the projection of $\overline{x}$ on the line $T_{y^*}\mathcal{M}$. Elementary trigonometry shows that $\|\overline{x} - z\| = \sin(\theta)\delta$. Hence

the previous equation yields

$$\begin{aligned}\|x-y\| &\geq \sin(\theta)\delta - \frac{\rho}{2}\delta^2 - \frac{\rho}{2}\delta^2 \\ &= \sin(\theta)\delta\left(1 - \frac{\rho}{\sin(\theta)}\delta\right),\end{aligned}$$

and we conclude using $\delta \leq \frac{\sin(\theta)}{2\rho}$.

Remark IV.19. A similar proof leads to the following result: let $x_0, y_0, z_0 \in \mathcal{M}_0$. Denote $\delta = \mathrm{d}_{\mathcal{M}_0}(x_0^*, x_0)$ and $\delta' = \mathrm{d}_{\mathcal{M}_0}(y_0^*, y_0)$. Suppose that x_0 and y_0 are in opposite orientation around z_0, that is, there exists a unit vector $v \in T_{z_0}\mathcal{M}_0$ such that $x_0 = \exp_{z_0}^{\mathcal{M}_0}(\delta v)$ and $y_0 = \exp_{z_0}^{\mathcal{M}_0}(-\delta' v)$. If $\delta', \delta \leq \frac{1}{\rho}$, then $\|x-y\| \geq \frac{1}{2}(\delta + \delta')$.

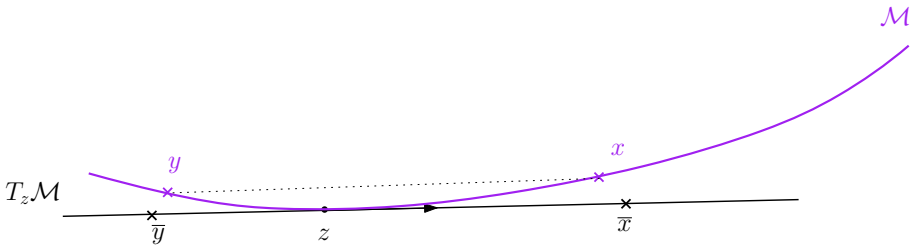


Figure IV.13: Situation in Remark IV.19.

The following lemma associates every point of $\mathcal{M}_0$ with small normal reach to a point with zero normal reach.

Lemma IV.20. *Let $x_0 \in \mathcal{M}_0$ with $\lambda_0(x_0) < \Delta \wedge \frac{\sin(\Theta)^2}{4\rho}$. Then there exists $x_0^* \in \mathcal{M}_0$ with*

$$\lambda_0(x_0^*) = 0 \qquad \text{and} \qquad \mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq c_{13}\lambda_0(x_0),$$

where $c_{13} = \frac{2}{\sin(\Theta)}$.

Proof. Let $y_0 \in \mathcal{M}_0$ such that $\|x-y\| = \lambda_0(x_0)$ and $x - y \perp T_y\mathcal{M}$. In order to find a point x_0^*, consider the following vector field on $\mathcal{M}_0 \times \mathcal{M}_0$:

$$\begin{array}{ccc}\mathcal{M}_0 \times \mathcal{M}_0 & \longrightarrow & T\mathcal{M}_0 \times T\mathcal{M}_0 \\ \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} & \longmapsto & \begin{pmatrix} p_{T_x\mathcal{M}}(y-x) \\ p_{T_y\mathcal{M}}(x-y) \end{pmatrix},\end{array}$$

where $p_{T_x\mathcal{M}}$ and $p_{T_y\mathcal{M}}$ denote the orthogonal projection on $T_x\mathcal{M}$ and $T_y\mathcal{M}$. We implicitly use the identifications $T_x\mathcal{M} \simeq T_{x_0}\mathcal{M}_0$. Since $\mathcal{M}_0$ is $\mathcal{C}^2$, this vector field is of regularity $\mathcal{C}^1$, and we can apply Cauchy-Lipschitz theorem. Let u_0 be a maximal integral curve for this field, with initial value $u_0(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Since $\mathcal{M}_0 \times \mathcal{M}_0$ is compact, the solution u_0 is global.

In order to study the convergence of u_0, we shall consider a Lyapunov map. Let $H\colon E \to \mathbb{R}$ be defined as $H(u) = \|u\|^2$. A computation shows that

$$\begin{aligned} H\left(\gamma(t)-\eta(t)\right)' &= -2\left\langle\, p_{T_{\gamma(t)}\mathcal{M}}\left(\gamma(t)-\eta(t)\right)\ ,\ \gamma(t)-\eta(t)\,\right\rangle \\ &\quad -2\left\langle\, p_{T_{\eta(t)}\mathcal{M}}\left(\gamma(t)-\eta(t)\right)\ ,\ \gamma(t)-\eta(t)\,\right\rangle \\ &= -2\left\|p_{T_{\gamma(t)}\mathcal{M}}\left(\gamma(t)-\eta(t)\right)\right\|^2 - 2\left\|p_{T_{\eta(t)}\mathcal{M}}\left(\gamma(t)-\eta(t)\right)\right\|^2. \end{aligned} \tag{IV.12}$$

This quantity is nonpositive, hence the map $t \mapsto H\left(\gamma(t)-\eta(t)\right)$ is nonincreasing. Note that for $t=0$, we have $H\left(\gamma(0)-\eta(0)\right) = \lambda_0(x_0)$. Note also that for every $t \in \mathbb{R}^+$, we have $H\left(\gamma(t)-\eta(t)\right) \neq 0$, since the relation $\gamma(t)=\eta(t)$ corresponds to a stationary point of the system.

We divide the rest of the proof in five steps.

Step 1. Let us prove that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t),\eta_0(t)) > \frac{1}{\rho}$ for every $t \in \mathbb{R}^+$. By contradiction, suppose that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t),\eta_0(t)) \leq \frac{1}{\rho}$ for some t. As a consequence of Remark IV.3, we have $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(0),\eta_0(0)) \geq \frac{1}{\rho}$. Therefore there exists a value $s \in [0,t]$ such that $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s),\eta_0(s)) = \frac{1}{\rho}$.

Let z_0 be a (geodesic) midpoint between $\gamma_0(s)$ and $\eta_0(s)$. We have

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s),z_0) = \mathrm{d}_{\mathcal{M}_0}(\eta_0(s),z_0) = \frac{1}{2\rho},$$

hence we can apply Remark IV.19 to get

$$\|\gamma(s)-\eta(s)\| \geq \frac{1}{2}\left(\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s),z_0) + \mathrm{d}_{\mathcal{M}_0}(\gamma_0(s),z_0)\right) = \frac{1}{2\rho}.$$

Besides, we have seen that the map $t \mapsto \|\gamma(t)-\eta(t)\|$ is bounded above by $\|\gamma(0)-\eta(0)\| = \lambda_0(x_0)$. The inequality $\frac{1}{2\rho} \leq \|\gamma(s)-\eta(s)\| \leq \lambda_0(x_0)$ now contradicts the assumption $\lambda_0(x_0) < \frac{\sin(\Theta)^2}{4\rho}$.

Step 2. Let us show that $\gamma(t)-\eta(t)$ goes to zero. Let v_0 denote the map $v_0(t) = \gamma_0(t) - \eta_0(t)$, and $v(t) = \gamma(t) - \eta(t)$. It is enough to show that H is a strict Lyapunov map, i.e., there exists a constant $c > 0$ such that

$$H\left(v(t)\right)' \leq -cH\left(v(t)\right). \tag{IV.13}$$

According to Equation (IV.12), we can write $H\left(v(t)\right)' = -2c(t)\left\|v(t)\right\|^2$ with

$$c(t) = \frac{1}{\left\|v(t)\right\|^2}\left(\left\|p_{T_{\gamma(t)}\mathcal{M}}\left(v(t)\right)\right\|^2 + \left\|p_{T_{\eta(t)}\mathcal{M}}\left(v(t)\right)\right\|^2\right) \tag{IV.14}$$

$$= \left\|p_{T_{\gamma(t)}\mathcal{M}}\left(\frac{v(t)}{\|v(t)\|}\right)\right\|^2 + \left\|p_{T_{\eta(t)}\mathcal{M}}\left(\frac{v(t)}{\|v(t)\|}\right)\right\|^2. \tag{IV.15}$$

To prove Equation (IV.13), it remains to show that $c(t)$ is bounded below.

By contradiction, suppose that it is not the case. This implies that there exists an increasing sequence $(t_n)_{n\geq 0}$ such that the sequence $(c(t_n))_{n\geq 0}$ goes to 0. By compacity of $\mathcal{M}_0$, we can assume that $(x_0(t_n))_{n\geq 0}$ and $(y_0(t_n))_{n\geq 0}$ admit a limit, that we denote x_0^* and y_0^*. By compacity of the unit sphere of E, we can also assume that $\left(\frac{v(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$ admits a limit v^*, as well as $\left(\frac{\gamma(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$ and $\left(\frac{\eta(t_n)}{\|v(t_n)\|}\right)_{n\geq 0}$. Note already the following facts: $\|v^*\| = 1$, and v^* is included in the 2-dimensional affine space spanned by $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$.

According to Step 1, we have $x_0^* \neq y_0^*$. Let us prove that $x^* = y^*$. By contradiction suppose that it is not the case. Then $(v(t_n))_{n\geq 0}$ goes to the nonzero vector $x^* - y^*$. Using that $c(t_n)$ goes to zero, Equation (IV.14) yields

$$\left\|p_{T_{x^*}\mathcal{M}}\left(x^* - y^*\right)\right\| = \left\|p_{T_{y^*}\mathcal{M}}\left(x^* - y^*\right)\right\| = 0.$$

Hence the pair (x^*, y^*) is an element of $\mathcal{D}_0$ (defined in Equation (IV.8)). By definition of Δ (Equation (IV.11)), we obtain $\|x^* - y^*\| \geq \Delta$. Besides, since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ is non-increasing, we get $\|x^* - y^*\| \leq \|x - y\|$, which is lower than Δ by assumption. This is a contradiction.

Now, we have $x^* = y^*$. Still using that $c(t_n)$ goes to zero, Equation (IV.15) yields

$$\left\|p_{T_{x^*}\mathcal{M}}\left(v^*\right)\right\| = \left\|p_{T_{y^*}\mathcal{M}}\left(v^*\right)\right\| = 0.$$

But $x^* = y^*$ implies that $T_{x^*}\mathcal{M} \neq T_{y^*}\mathcal{M}$, according to Hypothesis 1. In conclusion, v^* is a vector of the affine space spanned by $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$, and v^* is orthogonal to both these lines. Hence v^* has to be zero, which is absurd since it has norm 1. We deduce that $c(t)$ is bounded below, and that H is a strict Lyapunov map.

Step 3. Let us prove that u_0 admits a limit $\begin{pmatrix} x_0^* \\ y_0^* \end{pmatrix}$ when $t \to +\infty$, with $x_0^* \neq y_0^*$ and $x^* = y^*$. By compacity of $\mathcal{M}_0 \times \mathcal{M}_0$, we can pick two accumulation points x_0^* and y_0^* of γ_0 and η_0. Let us prove that, for every $\epsilon > 0$, there exists a $t \geq 0$ such that for every $s \geq t$, the geodesic distances $\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*)$ and $\mathrm{d}_{\mathcal{M}_0}(\eta_0(s), y_0^*)$ are upper bounded by ϵ. This would imply that γ_0 and η_0 admit x_0^* and y_0^* as limits. Let $\epsilon > 0$. We can assume that $\epsilon < \frac{\sin(\Theta)}{2\rho}$, where Θ is defined in Equation (IV.10).

According to Step 2, we have $x^* = y^*$. Hence the tangent spaces $T_{x^*}\mathcal{M}$ and $T_{y^*}\mathcal{M}$ are different. Let $\theta \in \left(0, \frac{\pi}{2}\right]$ be the angle they form. Since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ goes to zero, there exists a $t \geq 0$ such that for every $s \geq t$, we have

$$\|\gamma(t) - \eta(t)\| < \frac{\sin(\theta)}{2}\epsilon. \tag{IV.16}$$

Now, by definition of the accumulation points x^* and y^*, there exists a $t' \geq t$ such that

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t'), x_0^*) \leq \epsilon \qquad \text{and} \qquad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t'), y_0^*) \leq \epsilon. \tag{IV.17}$$

We shall deduce that for every $s \geq t'$, we have

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*) \leq \epsilon \qquad \text{and} \qquad \mathrm{d}_{\mathcal{M}_0}(\eta_0(s), y_0^*) \leq \epsilon. \tag{IV.18}$$

Let us prove it by contradiction. From Equation (IV.17) and the assumption that Equation (IV.18) is false, we deduce that there exists a first value $s \geq t'$ such that $\delta = \mathrm{d}_{\mathcal{M}_0}(\gamma_0(s), x_0^*) = \epsilon$ or $\delta' = \mathrm{d}_{\mathcal{M}_0}(\eta_0(s), x_0^*) = \epsilon$. Since $\epsilon < \frac{\sin(\Theta)}{2\rho}$, we can use Lemma IV.18 to deduce

$$\|\gamma_0(s) - \eta_0(s)\| \geq \frac{\sin(\theta)}{2}\epsilon.$$

But this contradicts Equation (IV.16).

Step 4. Let us show that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{\sin(\theta)}{2\rho}$ and $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) \leq \frac{\sin(\theta)}{2\rho}$. By contradiction, suppose that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) > \frac{\sin(\theta)}{2\rho}$ or $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) > \frac{\sin(\theta)}{2\rho}$. According to the limits $\gamma_0 \to x_0^*$ and $\eta_0 \to y_0^*$, there exists $t \in \mathbb{R}^+$ such that

$$\mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), x_0^*) = \frac{\sin(\theta)}{2\rho} \quad \text{and} \quad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t), y_0^*) \leq \frac{\sin(\theta)}{2\rho}$$
$$\text{or} \qquad \mathrm{d}_{\mathcal{M}_0}(\gamma_0(t), x_0^*) \leq \frac{\sin(\theta)}{2\rho} \quad \text{and} \quad \mathrm{d}_{\mathcal{M}_0}(\eta_0(t), y_0^*) = \frac{\sin(\theta)}{2\rho}.$$

In both cases, we can apply Lemma IV.18 to get

$$\|\gamma(t) - \eta(t)\| \geq \frac{\sin(\theta)}{2} \cdot \frac{\sin(\theta)}{2\rho} = \frac{\sin(\theta)^2}{4\rho}. \tag{IV.19}$$

Since the map $t \mapsto \|\gamma(t) - \eta(t)\|$ is non-increasing, we have

$$\|\gamma(t) - \eta(t)\| \leq \|\gamma(0) - \eta(0)\| = \|x - y\| = \lambda_0(x_0).$$

But $\lambda_0(x_0) < \frac{\sin(\Theta)^2}{4\rho}$ by assumption. Hence $\|\gamma(t) - \eta(t)\| < \frac{\sin(\theta)^2}{4\rho}$, which contradicts Equation (IV.19).

Step 5. Let us show that $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \geq \frac{2}{\sin(\theta)}\lambda_0(x_0)$. According to Step 4, we have $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{\sin(\theta)}{2\rho}$ and $\mathrm{d}_{\mathcal{M}_0}(y_0, y_0^*) \leq \frac{\sin(\theta)}{2\rho}$. Therefore, Lemma IV.18 gives

$$\|x - y\| \geq \frac{\sin(\theta)}{2}\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*).$$

Using $\|x - y\| = \lambda_0(x_0)$ and $\sin(\theta) \geq \sin(\Theta)$, we obtain

$$\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq \frac{2}{\sin(\theta)}\|x - y\| \leq \frac{2}{\sin(\Theta)}\lambda_0(x_0).$$

We can now deduce the main result of this subsection: Hypothesis 4 holds in dimension 1.

Proposition IV.21. *For every $r < \Delta \wedge \frac{\sin(\Theta)^2}{4\rho}$, we have*

$$\mu_0(\lambda_0^r) \leq c_{14} r$$

where $c_{14} = |C_0| f_{\max} J_{\max} c_{13}$ and $|C_0|$ is the number of self-intersection points of $\mathcal{M}_0$.

Proof. Let C_0 denote the set of self-intersection points of $\mathcal{M}_0$, i.e.,

$$C_0 = \{x_0 \in \mathcal{M}_0, \lambda_0(x_0) = 0\}.$$

Observe that C_0 is closely related to the set $\mathcal{C}_0$ defined in Equation (IV.9). Using Lemma IV.20, we can pair every $x_0 \in \lambda_0^r$ to a point $x_0^* \in C_0$ with $\mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq c_{13}\lambda_0(x_0)$. In other words, the sublevel set λ_0^r is included in the (geodesic) thickening

$$C_0^{c_{13}r} = \{x_0 \in \mathcal{M}_0, \exists x_0^* \in C_0, \mathrm{d}_{\mathcal{M}_0}(x_0, x_0^*) \leq c_{13} r\}.$$

Now, C_0 is a finite set, and we write its thickening as a union of geodesic balls:

$$C_0^{c_{13}r} = \bigcup_{x_0 \in \mathcal{C}_0} \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{13} r)$$

Thanks to Hypothesis 3, we can relate the measure μ_0 to the 1-dimensional Hausdorff measure $\mathcal{H}^1$. As in the proof of Proposition IV.16, we get

$$\mu_0\left(\overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{13} r)\right) \leq f_{\max} J_{\max} c_{13} r.$$

Therefore, if $|C_0|$ denotes the cardinal of C_0, we obtain

$$C_0^{c_{13}r} \leq |C_0| f_{\max} J_{\max} c_{13} r.$$

IV.3 Tangent space estimation

In this section, we show that one can estimate the tangent spaces of $\mathcal{M}$ based on a sample of it, via the computation of local covariance matrices. We study the consistency of this estimation in Subsection IV.3.2, which is based on the results of the last section. In Subsection IV.3.3 we prove that this estimation is stable, based on lighter hypotheses than 1, 2 and 3.

IV.3.1 Local covariance matrices and lifted measure

Definition IV.22. Let ν be any probability measure on E. Let $r > 0$ and $x \in \mathrm{supp}(\nu)$. The *local covariance matrix of ν around x at scale r* is the following matrix:

$$\Sigma_\nu(x) = \int_{\overline{\mathcal{B}}(x,r)} (x - y)^{\otimes 2} \frac{\mathrm{d}\nu(y)}{\nu(\overline{\mathcal{B}}(x, r))}.$$

We also define the *normalized local covariance matrix* as $\overline{\Sigma}_\nu(x) = \frac{1}{r^2}\Sigma_\nu(x)$.

Note that $\Sigma_\nu(x)$ and $\overline{\Sigma}_\nu(x)$ depend on r, which is not made explicit in the notation. The normalization factor $\frac{1}{r^2}$ of the normalized local covariance matrix is justified by Proposition IV.24. Moreover, we introduce the following notations: for every $r > 0$ and $x \in \mathrm{supp}(\nu)$,

- ν_x is the restriction of ν to the ball $\overline{\mathcal{B}}(x, r)$,
- $\overline{\nu_x} = \frac{1}{\nu(\overline{\mathcal{B}}(x,r))}\nu_x$ is the corresponding probability measure.

Thus the local covariance matrix can be written as $\Sigma_\nu(x) = \int (x-y)^{\otimes 2} \mathrm{d}\overline{\nu_x}(y)$.

The collection of probability measures $\{\overline{\nu_x}\}_{x\in\mathrm{supp}(\nu)}$ is called in [MSW19, Section 3.3] the local truncation of ν at scale r. The application $x \mapsto \Sigma_\nu(x)$ is called in [MMM18, Section 2.2] the multiscale covariance tensor field of ν associated to the truncation kernel.

We remind the reader that the aim of this work is to estimate the measure $\check{\mu}_0$, defined on $E \times \mathrm{M}(E)$ as $\check{\mu}_0 = \check{u}_*\mu_0$. We call it the *exact lifted measure*. In other words, it can be defined as

$$\check{\mu}_0 = (u_*\mu_0)(x_0) \otimes \left\{\delta_{\frac{1}{d+2}p_{T_x\mathcal{M}}}\right\}$$

by disintegration of measure. Here is another alternative definition of $\check{\mu}_0$: for any $\phi\colon E \times \mathrm{M}(E) \to \mathbb{R}$ with compact support,

$$\int \phi(x, A)\mathrm{d}\check{\mu}_0(x, A) = \int \phi\left(u(x_0), \frac{1}{d+2}p_{T_x\mathcal{M}}\right)\mathrm{d}\mu_0(x_0). \tag{IV.20}$$

In order to approximate $\check{\mu}_0$, we consider the following construction.

Definition IV.23. if ν is any measure on E, we denote by $\check{\nu}$ the measure on $E \times \mathrm{M}(E)$ defined by

$$\check{\nu} = \nu(x) \otimes \left\{\delta_{\overline{\Sigma}_\nu(x)}\right\}.$$

It is called the *lifted measure* associated to ν. In other words, for every $\phi\colon E \times \mathrm{M}(E) \to \mathbb{R}$ with compact support, we have

$$\int \phi(x, A)\mathrm{d}\check{\nu}(x, A) = \int \phi\Big(x, \overline{\Sigma}_\nu(x)\Big)\mathrm{d}\nu(x).$$

In accordance with the local covariance matrices, the lifted measure $\check{\nu}$ depends on the parameter r which is not made explicit in the notation. In order to compare these measures, we consider a Wasserstein-type distance on the space $E \times \mathrm{M}(E)$. Fix $\gamma > 0$, and let $\|\cdot\|_\gamma$ be the Euclidean norm on $E \times \mathrm{M}(E)$ defined as

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_{\mathrm{F}}^2, \tag{IV.21}$$

where $\|\cdot\|$ represents the usual Euclidean norm on E and $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius norm on $\mathrm{M}(E)$. Let $p \geq 1$. We denote by $\mathrm{W}_{p,\gamma}(\cdot, \cdot)$ the p-Wasserstein distance with

respect to this metric. By definition, if α, β are probability measures on $E \times \mathrm{M}(E)$, then $\mathrm{W}_{p,\gamma}(\alpha, \beta)$ can be written as

$$\mathrm{W}_{p,\gamma}(\alpha,\beta) = \inf_{\pi} \left(\int_{(E\times \mathrm{M}(E))^2} \|(x,A)-(y,B)\|_\gamma^p \, \mathrm{d}\pi\big((x,A),(y,B)\big) \right)^{\frac{1}{p}}, \quad \text{(IV.22)}$$

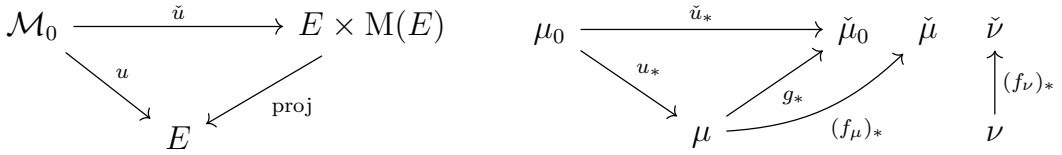
where the infimum is taken over all measures π on $(E \times \mathrm{M}(E))^2$ with marginals α and β.

We subdivise the rest of this section in three subsections. They respectively consists in showing that

- **Consistency:** if μ_0 is a measure satisfying the Hypotheses 2 and 3, then $\mathrm{W}_{p,\gamma}(\check{\mu}_0, \check{\mu})$ is small (Proposition IV.27),
- **Stability:** in addition, if ν is a measure on E such that $\mathrm{W}_p(\mu, \nu)$ is small, then so is $\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu})$ (Proposition IV.29)
- **Approximation:** under the previous hypotheses, $\mathrm{W}_{p,\gamma}(\check{\mu}_0, \check{\nu})$ is small (Theorem IV.33).

The first point means that the lifted measure $\check{\mu}$ is close to the exact lifted measure $\check{\mu}_0$. In other words, construction we propose is consistent. If we are not observing μ but a close measure ν, the second point states that the lifted measure $\check{\nu}$ is still close to $\check{\mu}$. Combining these two statements gives the third one: the lifted measure $\check{\nu}$ is close the exact lifted measure $\check{\mu}_0$.

These several measures fit in a commutative diagram:



where the maps g, f_μ and $f_\nu \colon E \to E \times \mathrm{M}(E)$ are defined as

$$g \colon x \longmapsto \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}}\right), \qquad f_\mu \colon x \longmapsto \Big(x, \overline{\Sigma}_\mu(x)\Big), \qquad f_\nu \colon x \longmapsto \Big(x, \overline{\Sigma}_\nu(x)\Big).$$

Note that the map g is well-defined only on points $x \in \mathcal{M}$ that are not self-intersection points, i.e., points x such that $\lambda(x) > 0$. Under Hypothesis 4, g is well-defined μ-almost surely. The maps f_μ and f_ν are defined respectively on $\mathrm{supp}(\mu)$ and $\mathrm{supp}(\nu)$.

IV.3.2 Consistency of the estimation

In this subsection, we assume that $\mathcal{M}_0$ and μ_0 satisfy the Hypotheses 2 and 3.

The following proposition shows that the normalized covariance matrix approximates the tangent spaces of $\mathcal{M}$, as long as the parameter r is chosen smaller than the normal reach. A similar result appears in [ACLZ17, Lemma 13] in the case where $\mathcal{M}$ is a submanifold and μ is the uniform distribution on $\mathcal{M}$. Based on this result,

we deduce that the lifted measure $\check{\mu}$ is close to the exact lifted measure $\check{\mu}_0$. The quality of this approximation depends on the measure of points with small normal reach, i.e., points where the tangent spaces are not well-estimated.

Proposition IV.24. *Let $x_0 \in \mathcal{M}_0$ and $r < \lambda(x) \wedge \frac{1}{2\rho}$. Denote by $p_{T_x\mathcal{M}}$ the orthogonal projection on the tangent space $T_x\mathcal{M}$, seen as a matrix. We have*

$$\left\| \overline{\Sigma}_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq c_{15} r.$$

Proposition IV.24 is a direct consequence of the two following lemmas.

Lemma IV.25 ([ACLZ17, Lemma 11]). *The following matrix is equal to $r^2 \frac{1}{d+2} p_{T_x\mathcal{M}}$:*

$$\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{V_d r^d}.$$

Lemma IV.26. *Still denoting $\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{V_d r^d}$, we have*

$$\left\| \Sigma_\mu(x) - \Sigma_* \right\|_{\mathrm{F}} \leq c_{15} r^3,$$

where $c_{15} = 6\rho + 4\frac{c_7}{f_{\min} J_{\min}} + \frac{f_{\max}}{f_{\min} J_{\min}} 2^d d\rho + \frac{c_8}{f_{\min} J_{\min}}$.

Proof. We use the notations of Lemmas IV.14 and IV.16. We write $T = T_x\mathcal{M}$, $\overline{\mathcal{B}} = \overline{\mathcal{B}}(x,r)$ and $\overline{\mathcal{B}}^T = (\overline{\exp}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}})$. We shall consider the following intermediate matrices:

$$\Sigma_1 = \int_{\overline{\mathcal{B}}} \left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x') \right)^{\otimes 2} \mathrm{d}\overline{\mu_x}(x')$$
$$\Sigma_2 = \int_{\overline{\mathcal{B}}^T} g(0) \cdot y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}$$
$$\Sigma_3 = \int_{\overline{\mathcal{B}}_T(0,r)} g(0) \cdot y^{\otimes 2} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}$$

Let us write the triangle inequality:

$$\left\| \Sigma_\mu(x) - \Sigma_* \right\|_{\mathrm{F}} \leq \underbrace{\left\| \Sigma_\mu(x) - \Sigma_1 \right\|_{\mathrm{F}}}_{(1)} + \underbrace{\left\| \Sigma_1 - \Sigma_2 \right\|_{\mathrm{F}}}_{(2)} + \underbrace{\left\| \Sigma_2 - \Sigma_3 \right\|_{\mathrm{F}}}_{(3)} + \underbrace{\left\| \Sigma_3 - \Sigma_* \right\|_{\mathrm{F}}}_{(4)}.$$

Term (1): By definition of the local covariance matrix, we have

$$\Sigma_\mu(x) = \int_{\overline{\mathcal{B}}(x,r)} (x - x')^{\otimes 2} \overline{\mu_x}(x').$$

We use the majoration

$$\begin{aligned}\left\|\Sigma_\mu(x)-\Sigma_1\right\|_{\mathrm{F}} &\leq \int_{\overline{\mathcal{B}}(x,r)} \left\|(x-x')^{\otimes 2}-\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\right\|_{\mathrm{F}} \mathrm{d}\overline{\mu_x}(x') \\ &\leq \sup_{x'\in\overline{\mathcal{B}}(x,r)\cap\mathcal{M}} \left\|(x-x')^{\otimes 2}-\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\right\|_{\mathrm{F}}.\end{aligned}$$

Let $x'\in\mathcal{B}(x,r)\cap\mathcal{M}$. According to Lemma IV.9, we have $\left\|\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right\| \leq 2r$. Moreover, $\|x-x'\|\leq r$, and Lemma IV.44 gives

$$\left\|(x-x')^{\otimes 2}-\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\right\|_{\mathrm{F}} \leq (r+2r)\left\|(x'-x)-\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right\|. \tag{IV.23}$$

Now, let us justify that

$$\left\|(x'-x)-\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right\| \leq \frac{\rho}{2}\mathrm{d}_{\mathcal{M}_0}(x_0,x_0')^2. \tag{IV.24}$$

If we write $x'=\gamma(\delta)$ with γ a geodesic such that $\gamma(0)=x$ and $\delta=\mathrm{d}_{\mathcal{M}_0}(x_0,x_0')$, then $\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')=\delta\dot\gamma(0)$, and we get

$$\begin{aligned}\left\|(x'-x)-\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right\| &= \|\gamma(\delta)-(x+\delta\dot\gamma(0))\| \\ &\leq \frac{\rho}{2}\delta^2,\end{aligned}$$

where we used Lemma IV.1 for the last inequality. Hence Equation (IV.24) is true. Combined with Lemma IV.9, which gives $\mathrm{d}_{\mathcal{M}_0}(x_0,x_0')\leq 2\|x-x'\|\leq 2r$, we obtain

$$\left\|(x-x')^{\otimes 2}-\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\right\|_{\mathrm{F}} \leq \frac{\rho}{2}(2r)^2 = 2\rho r^2.$$

To conclude, we use Equation (IV.23) to deduce $\|\Sigma_\mu(x)-\Sigma_1\|_{\mathrm{F}}\leq (r+2r)2\rho r^2 = 6\rho r^3$.

Term (2): By transfer, we can write Σ_1 as

$$\Sigma_1 = \int_{\overline{\mathcal{B}}}\left(\left(\overline{\exp}_x^{\mathcal{M}}\right)^{-1}(x')\right)^{\otimes 2}\frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|} = \int_{\overline{\mathcal{B}}^T} g(y)y^{\otimes 2}\frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

We deduce the majoration

$$\|\Sigma_1-\Sigma_2\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}^T}|g(0)-g(y)|\,\|y^{\otimes 2}\|\,\frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

According to Lemma IV.44, $\|y^{\otimes 2}\| = \|y\|^2 \leq (2r)^2$, and from Lemma IV.14 we get $|g(y) - g(0)| \leq c_7 r$. Therefore,

$$\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}} \leq 4r^2 \cdot c_7 r \cdot \frac{\mathcal{H}^d\big(\overline{\mathcal{B}}^T\big)}{|\mu_x|}.$$

To conclude, note that $|\mu_x| \geq f_{\min} J_{\min} \mathcal{H}^d\big(\overline{\mathcal{B}}^T\big)$ (as in Lemma IV.14), so we obtain $\|\Sigma_1 - \Sigma_2\|_{\mathrm{F}} \leq 4\frac{c_7}{f_{\min} J_{\min}} r^3$.

Term (3): As for the previous terms, we use the majoration

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}_T(0,r) \setminus \overline{\mathcal{B}}^T} \left\|g(0) \cdot y^{\otimes 2}\right\|_{\mathrm{F}} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

On the one hand, $\|g(0) \cdot y^{\otimes 2}\|_{\mathrm{F}} \leq g(0) r^2 \leq f_{\max} r^2$, and we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq f_{\max} r^2 \frac{\mathcal{H}^d\left(\overline{\mathcal{B}}_T(0,r) \setminus \overline{\mathcal{B}}^T\right)}{|\mu_x|}.$$

On the other hand, since $\overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_T(x, c_4(\rho r) r)$, we have

$$\mathcal{H}^d\left(\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0,r)\right) = (c_4(\rho r) r)^d V_d - r^d V_d.$$

The inequality $A^d - 1 \leq d(A-1)A^{d-1}$, where $A \geq 1$, gives

$$(c_4(\rho r) r)^d V_d - r^d V_d \leq V_d r^d \cdot d(c_4(\rho r) - 1) 2^{d-1}.$$

Combined with the inequalities $c_4(\rho r) \leq 1 + 2\rho r$ and $|\mu_x| \geq f_{\min} J_{\min} V_d r^d$, we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathrm{F}} \leq \frac{f_{\max}}{f_{\min} J_{\min}} 2^d d \rho r^3.$$

Term (4): Let us write Σ^* as

$$\Sigma_* = \int_{\overline{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{|\mu_x|}{V_d r^d} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

Hence we have

$$\|\Sigma_3 - \Sigma_*\|_{\mathrm{F}} \leq \int_{\overline{\mathcal{B}}_T(0,r)} \left| \frac{|\mu_x|}{V_d r^d} - f(x) \right| \left\|y^{\otimes 2}\right\|_{\mathrm{F}} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|}.$$

According to Lemma IV.16 point 2, $\left|\frac{|\mu_x|}{V_d r^d} - f(x)\right| \leq c_8 r$. Moreover, $\|y^{\otimes 2}\|_{\mathrm{F}} \leq r^2$ and $\int_{\overline{\mathcal{B}}_T(0,r)} \frac{\mathrm{d}\mathcal{H}^d(y)}{|\mu_x|} \leq \frac{1}{f_{\min} J_{\min}}$. Therefore,

$$\|\Sigma_3 - \Sigma_*\|_{\mathrm{F}} \leq \frac{c_8}{f_{\min} J_{\min}} r^3.$$

We now deduce a result concerning the lifted measures $\check{\mu}$ and $\check{\mu}_0$ (defined in Subsection IV.3.1). We remind the reader that the notation λ^r refers to the sublevel set $\lambda^{-1}([0,r])$. Hence the quantity $\mu(\lambda^r)$ is the measure of points $x \in \mathcal{M}$ such that $\lambda(x) \leq t$.

Proposition IV.27. *Let $r < \frac{1}{2\rho}$. Then*

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma \left(2\mu(\lambda^r)^{\frac{1}{p}} + c_{15} r \right).$$

Proof. Define the map $\phi\colon \mathcal{M}_0 \to (E \times \mathrm{M}(E)) \times (E \times \mathrm{M}(E))$ as

$$\phi\colon x_0 \mapsto \left(\left(x, \overline{\Sigma}_\mu(x) \right), \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}} \right) \right),$$

and consider the measure $\pi = \phi_* \mu_0$. It is a transport plan between $\check{\mu}$ and $\check{\mu}_0$. By definition of the Wasserstein distance,

$$\mathrm{W}^p_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \int \|(x,T) - (x',T')\|^p_\gamma \, \mathrm{d}\pi \left((x,T), (x',T') \right),$$

hence we can write

$$\begin{aligned} \mathrm{W}^p_{p,\gamma}(\check{\mu}, \check{\mu}_0) &\leq \int \left\| \left(x, \frac{1}{r^2}\Sigma_\mu(x) \right) - \left(x, \frac{1}{d+2} p_{T_x\mathcal{M}} \right) \right\|^p_\gamma \mathrm{d}\mu(x) \\ &= \gamma^p \int \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|^p_{\mathrm{F}} \mathrm{d}\mu(x). \end{aligned}$$

We split this last integral into the sets $A = \lambda^r$ and $B = E \setminus \lambda^r$.

On A, we use the majoration $\left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq \left\| \frac{1}{r^2}\Sigma_\mu(x) \right\|_{\mathrm{F}} + \left\| \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|_{\mathrm{F}} \leq 1 + 1$ to obtain

$$\int_A \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|^p_{\mathrm{F}} \mathrm{d}\mu(x) \leq 2^p \mu(A).$$

On B, we use Proposition IV.24 to get

$$\int_B \left\| \frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}} \right\|^p_{\mathrm{F}} \mathrm{d}\mu(x) \leq (c_{15} r)^p.$$

Combining these two inequalities yields $\mathrm{W}^p_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma^p (2^p \mu(A) + (c_{15} r)^p)$. Using the inequality $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$, where $a, b \geq 0$, we deduce the result:

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma \left(2\mu(A)^{\frac{1}{p}} + c_{15} r \right).$$

IV.3.3 Stability of the estimation

In this subsection we study the stability of the operator $\mu \mapsto \overline{\Sigma}_\mu(\cdot)$ with respect to the W_p metric on measures. The results of this subsection only rely on the following hypotheses about μ:

Hypothesis 5. $\exists c_9 > 0, \forall x \in \mathrm{supp}(\mu),\ \forall t \in [0, \frac{1}{2\rho})$,

$$\mu(\overline{\mathcal{B}}(x,t)) \geq c_9 t^d.$$

Hypothesis 6. $\exists c_{10} > 0, \forall x \in \mathrm{supp}(\mu),\ \exists \lambda(x) \geq 0,\ \forall s,t \in [0, \lambda(x) \wedge \frac{1}{2\rho})$ s.t. $s \leq t$,

$$\mu(\overline{\mathcal{B}}(x,t) \setminus \overline{\mathcal{B}}(x,s)) \leq c_{10} t^{d-1}(t-s).$$

Hypothesis 7. $\exists c_{11} > 0, \forall x \in \mathrm{supp}(\mu),\ \forall s,t \in [0, \frac{1}{2\rho})$ s.t. $s \leq t$,

$$\mu(\overline{\mathcal{B}}(x,t) \setminus \overline{\mathcal{B}}(x,s)) \leq c_{11} t^{d-\frac{1}{2}}(t-s)^{\frac{1}{2}}.$$

Note that, as stated in Propositions IV.16 and IV.17, the initial Hypotheses 2 and 3 imply the Hypotheses 5, 6 and 7 with $\lambda(x)$ being the normal reach of $\mathcal{M}$ at x, and with the constants $c_9 = f_{\min} J_{\min} V_d$, $c_{10} = d 2^d f_{\max} J_{\max} V_d$ and $c_{11} = \frac{f_{\max} J_{\max}}{f_{\min} J_{\min}} \left(\frac{\rho}{\sqrt{4-\sqrt{13}}} \right)^d d 2^{2d} \sqrt{3}$.

Let μ and ν be two probability measures, $x \in \mathrm{supp}(\mu) \cap \mathrm{supp}(\nu)$, and consider the Frobenius distance $\left\| \overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x) \right\|_{\mathrm{F}}$ between the normalized local covariance matrices. One shows that this distance is related to the 1-Wasserstein distance between the localized probability measures $\overline{\mu_x}$ and $\overline{\nu_x}$ via the following inequality (see Equation (IV.27) in the proof of Lemma IV.30):

$$\left\| \overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x) \right\|_{\mathrm{F}} \leq \frac{2}{r} \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_x}).$$

Without any assumption on the measures, it is not true that $\mathrm{W}_1\left(\overline{\mu_x}, \overline{\nu_x}\right)$ goes to 0 as $\mathrm{W}_1(\mu,\nu)$ does. However, if we assume that μ satisfies the Hypotheses 5 and 6, that x satisfies $\lambda(x) > 0$ and that r is chosen such that $4 \left(\frac{\mathrm{W}_1(\mu,\nu)}{c_9 \wedge 1} \right)^{\frac{1}{d+1}} \leq r < \lambda(x) \wedge \frac{1}{2\rho}$, then we are able to prove (Lemma IV.48) that

$$\mathrm{W}_1\left(\overline{\mu_x}, \overline{\nu_x}\right) \leq c_{27} \left(\frac{\mathrm{W}_1\left(\mu,\nu\right)}{r^{d-1}} \right)^{\frac{1}{2}}. \tag{IV.25}$$

In Remark IV.50, we show that the exponent $d-1$ on r is optimal. As a consequence of this inequality, estimating local covariance matrices is robust in Wasserstein distance:

$$\left\| \overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x) \right\|_{\mathrm{F}} \leq 2 c_{27} \left(\frac{\mathrm{W}_1(\mu,\nu)}{r^{d+1}} \right)^{\frac{1}{2}}. \tag{IV.26}$$

A stability result of this kind already appears in [MSW19, Theorem 4.3], where μ and ν are two probability measures on a bounded set X, and satisfy the following

condition: $\forall x \in X, \forall s, r \leq 0$ s.t. $s \leq r$, we have $\frac{\mu(\overline{\mathcal{B}}(x,r))}{\mu(\overline{\mathcal{B}}(x,s))} \leq (\frac{r}{s})^d$. The theorem states that, denoting $D = \mathrm{diam(X)}$, for all $x \in X$,

$$\mathrm{W}_1\left(\overline{\mu_x}, \overline{\nu_x}\right) \leq (1+2r)\left[\frac{\mathrm{W}_1\left(\mu,\nu\right)^{\frac{1}{2}}}{1 \wedge (\frac{r}{D})^d} + \left(1 + \frac{\mathrm{W}_1\left(\mu,\nu\right)^{\frac{1}{2}}}{r}\right)^d - 1\right].$$

When $r \leq D$ and $\mathrm{W}_1\left(\mu,\nu\right)$ goes to zero, we obtain that $\mathrm{W}_1\left(\overline{\mu_x}, \overline{\nu_x}\right)$ is of order

$$\mathrm{W}_1\left(\overline{\mu_x}, \overline{\nu_x}\right) \leq (1+2r)D^d\left(\frac{\mathrm{W}_1\left(\mu,\nu\right)}{r^{2d}}\right)^{\frac{1}{2}}.$$

The exponent on r is greater here than in Equation (IV.25).

Another result in [MMM18, Theorem 3] bounds the distance $\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}}$ with the ∞-Wasserstein distance $\mathrm{W}_\infty(\mu,\nu)$. Namely, if μ and ν are fully supported probability measures with densities upper bounded by $l > 0$ and supports included in $X \subset \mathbb{R}^d$, denoting $D = \mathrm{diam(X)}$, we have

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}} \leq lA\mathrm{W}_\infty(\mu,\nu),$$

where $A = \frac{d}{d+2}\frac{(r+D)^{d+1}}{Dr^d} + \frac{(2r+D)(r+D)^d}{r^d} + \frac{2d}{d+2}\frac{(r+D)^{d+2}}{Dr^d}$.

Remark IV.28. Let us show that in general, for $x \in \mathrm{supp}(\mu) \cap \mathrm{supp}(\nu)$, it is not true that $\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(x)\right\|_{\mathrm{F}}$ goes to zero as $\mathrm{W}_1(\mu,\nu)$ goes to zero. Similarly, $\mathrm{W}_{p,\gamma}(\check{\mu},\check{\nu})$ does not have to go to zero. For example, one can consider $\epsilon > 0$, and the measures on $\mathbb{R}$ defined as

$$\mu = \tfrac{1}{2}(\delta_0 + \delta_1) \qquad \text{and} \qquad \nu = \tfrac{1}{2}(\delta_0 + \delta_{1+\epsilon}).$$

Choose the scale parameter $r = 1$. We have $\Sigma_\mu(0) = \Sigma_\mu(1) = \frac{1}{2}1^{\otimes 2}$ and $\Sigma_\nu(0) = \Sigma_\nu(1+\epsilon) = 0$. The measures $\check{\mu}$ and $\check{\nu}$ on $\mathbb{R} \times \mathrm{M}(\mathbb{R})$ can be written

$$\check{\mu} = \tfrac{1}{2}\left(\delta_{\left(0,\frac{1}{2}1^{\otimes 2}\right)} + \delta_{\left(1,\frac{1}{2}1^{\otimes 2}\right)}\right) \qquad \text{and} \qquad \check{\nu} = \tfrac{1}{2}\left(\delta_{(0,0)} + \delta_{(1+\epsilon,0)}\right).$$

A computation shows that

$$\begin{aligned}\mathrm{W}^p_{p,\gamma}(\check{\mu},\check{\nu}) &= \frac{1}{2}\left\|\left(0,\frac{1}{2}1^{\otimes 2}\right) - \left(0,0\right)\right\|^p_\gamma + \frac{1}{2}\left\|\left(1,\frac{1}{2}1^{\otimes 2}\right) - \left(1+\epsilon,0\right)\right\|^p_\gamma \\ &= \frac{1}{2}\left(\left(\frac{\gamma}{2}\right)^p + \left(\epsilon^2 + \gamma^2\frac{1}{4}\right)^{\frac{p}{2}}\right) \geq \left(\frac{\gamma}{2}\right)^p.\end{aligned}$$

Hence $\mathrm{W}_{p,\gamma}(\check{\mu},\check{\nu}) \geq \frac{\gamma}{2} > 0$. Besides, we have $\mathrm{W}_1(\mu,\nu) = \frac{1}{2}\epsilon$. Hence $\mathrm{W}_{p,\gamma}(\check{\mu},\check{\nu})$ does not go to zero as $\mathrm{W}_1(\mu,\nu)$ does. However, under regularity assumptions on μ, the following proposition states that it is the case.

Proposition IV.29. *Let μ and ν be two probability measures on E. Suppose that μ satisfies Hypotheses 5, 6 and 7. Define $w = \mathrm{W}_p(\mu,\nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$\mathrm{W}_{p,\gamma}(\check{\mu},\check{\nu}) \leq 2w + \gamma c_{16}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + \gamma c'_{16}\mu(\lambda^r)^{\frac{1}{p}}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$$

with $c_{16} = 4(1 + c_{17})$ *and* $c'_{16} = 4c_{28}$.

Proof. According to Lemma IV.30 stated below, we have

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} \left(1 + \frac{2\gamma}{r}\right) w + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left(\int \mathrm{W}_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}.$$

Let $\alpha = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}$. Lemma IV.31, also stated below, gives

$$\left(\int \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left(c_{28} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + c_{17}\alpha \right)$$

Combining these inequalities yields

$$\begin{aligned} \mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu}) &\leq 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left(w + 2^{\frac{p-1}{p}} c_{17}\alpha \right) + \left(2^{\frac{p-1}{p}}\right)^2 \frac{2\gamma}{r} c_{28} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} \\ &\leq 2w + 2 \cdot 2\gamma \left(\frac{w}{r} + 2c_{17} \frac{\alpha}{r} \right) + 2^2 \cdot 2\gamma c_{28} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{\alpha}{r}\right)^{\frac{1}{2}}, \end{aligned}$$

where we used $2^{\frac{p-1}{p}} \leq 2$. Since $r \leq 1$, we have $w \leq 1$ and $w = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} r^{\frac{d-1}{2}} w^{\frac{1}{2}} \leq \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} = \alpha$. We get

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} 2\gamma \left(1 + 2^{\frac{p-1}{p}} c_{17}\right) \frac{\alpha}{r} + \left(2^{\frac{p-1}{p}}\right)^2 2\gamma c_{28} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{\alpha}{r}\right)^{\frac{1}{2}}.$$

By replacing $\frac{\alpha}{r}$ with $\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}}$, we obtain the result. ∎

Let us interpret the inequality

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2w + \gamma c_{16} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + \gamma c'_{16} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$$

The first term $2w$ is to be seen as the initial error between the measures μ and ν. The second term $\gamma c_{16} (\frac{w}{r^{d+1}})^{\frac{1}{2}}$ corresponds to the local errors $\mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})$ when comparing the normalized covariance matrices. The third term $\gamma c'_{16} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$ stands for the error on points x such that $\lambda(x) \leq r$, where the stability is weaker.

As a consequence of this proposition, the application $\mu \mapsto \check{\mu}$, seen as an application between spaces of measures endowed with the Wassertein metric, is continuous on the set of measures μ which satisfy Hypotheses 5, 6 and 7 with $\frac{1}{2\rho} \geq r$.

We now state the lemmas used in the proof of this proposition.

Lemma IV.30. *Let* π *be an optimal transport plan for* $\mathrm{W}_p(\mu, \nu)$. *Then*

$$\mathrm{W}_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} \left(1 + \frac{2\gamma}{r}\right) \mathrm{W}_p(\mu, \nu) + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left(\int \mathrm{W}_1^p(\overline{\mu_x}, \overline{\nu_y}) \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}.$$

Proof. We first prove the following fact: for every $x \in \mathrm{supp}(\mu)$ and $y \in \mathrm{supp}(\nu)$,

$$\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathrm{F}} \leq 2r \left(\|x-y\| + \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})\right). \tag{IV.27}$$

Let ρ be any transport plan between $\overline{\mu_x}$ and $\overline{\nu_y}$. We have

$$\begin{aligned}\Sigma_\mu(x) - \Sigma_\nu(y) &= \int (x-y)^{\otimes 2} \mathrm{d}\overline{\mu_x}(x') - \int (y-y')^{\otimes 2} \mathrm{d}\overline{\mu_y}(y') \\ &= \int \left((x-x')^{\otimes 2} - (y-y')^{\otimes 2}\right) \mathrm{d}\rho(x', y'). \end{aligned} \tag{IV.28}$$

For any $x' \in \overline{\mathcal{B}}(x, r)$ and $y' \in \overline{\mathcal{B}}(y, r)$, we can use Lemma IV.44 to get

$$\left\|(x-x')^{\otimes 2} - (y-y')^{\otimes 2}\right\|_{\mathrm{F}} \leq (r+r)(\|x-y\| + \|x'-y'\|).$$

Therefore, Equation (IV.28) yields

$$\begin{aligned}\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathrm{F}} &\leq \int 2r(\|x-y\| + \|x'-y'\|) \mathrm{d}\rho(x', y') \\ &\leq 2r \left(\|x-y\| + \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})\right). \end{aligned}$$

Now, a transport plan π for $\mathrm{W}_p(\mu, \nu)$ begin given, we build a transport plan $\check{\pi}$ for $(\check{\mu}, \check{\nu})$ as follows: for every $\phi\colon (E \times \mathrm{M}(E))^2 \to \mathbb{R}$ with compact support, let $\check{\pi}$ satisfies

$$\int \phi(x, A, y, B) \mathrm{d}\check{\pi}(x, A, y, B) = \int \phi\left(x, \overline{\Sigma}_\mu(x), y, \overline{\Sigma}_\nu(y)\right) \mathrm{d}\pi(x, y).$$

We have the majoration

$$\begin{aligned}\mathrm{W}_{p,\gamma}^p(\check{\mu}, \check{\nu}) &\leq \int \|(x, A) - (y, B)\|_\gamma^p \, \mathrm{d}\check{\pi}(x, A, y, B) \\ &= \int \left(\|x-y\|^2 + \gamma^2 \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}^2\right)^{\frac{p}{2}} \mathrm{d}\pi(x, y) \\ &\leq \int \left(\|x-y\| + \gamma \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}\right)^p \mathrm{d}\pi(x, y) \end{aligned} \tag{IV.29}$$

Besides, Equation (IV.27) gives

$$\left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}} \leq \frac{1}{r^2} \|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathrm{F}} \leq \frac{2}{r} \left(\|x-y\| + \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})\right).$$

We can use the inequality $(a+b)^p \leq 2^{p-1}(a^p + b^p)$, where $a, b \geq 0$, to deduce

$$\begin{aligned}\left(\|x-y\| + \gamma \left\|\overline{\Sigma}_\mu(x) - \overline{\Sigma}_\nu(y)\right\|_{\mathrm{F}}\right)^p &\leq \left(\|x-y\| + \gamma \frac{2}{r} \left(\|x-y\| + \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})\right)\right)^p \\ &\leq 2^{p-1} \left(\left(1 + \frac{2\gamma}{r}\right) \|x-y\|\right)^p + 2^{p-1} \left(\frac{2\gamma}{r} \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y}\right)^p. \end{aligned}$$

By inserting this inequality in Equation (IV.29) we obtain

$$\begin{aligned}\mathrm{W}_{p,\gamma}^p(\check{\mu},\check{\nu}) &\leq 2^{p-1}\int\left(\left(1+\frac{2\gamma}{r}\right)\|x-y\|\right)^p+\left(\frac{2\gamma}{r}\mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\right)^p \mathrm{d}\pi(x,y)\\ &= 2^{p-1}\left(1+\frac{2\gamma}{r}\right)^p\mathrm{W}_p^p(\mu,\nu)+2^{p-1}\left(\frac{2\gamma}{r}\right)^p\int\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y),\end{aligned}$$

which yields the result.

Lemma IV.31. *Let $w=\mathrm{W}_p(\mu,\nu)$ and define $\alpha=(\frac{w}{r^{d-1}})^{\frac{1}{2}}$. Suppose that $r\leq\frac{1}{2\rho}$ and $w\leq(c_9\wedge 1)(\frac{r}{4})^{d+1}$. Let π be an optimal transport plan for $\mathrm{W}_p(\mu,\nu)$. Then*

$$\begin{aligned}&\left(\int\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}}\\ &\qquad\leq 2^{\frac{p-1}{p}}\left(c_{28}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}}+\left(2r^d+c_{26}r^{\frac{d+1}{2}}+c_{27}\right)\alpha+(1+c_{25})w\right).\end{aligned}$$

If we suppose that $r\leq 1$, then

$$\left(\int\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}}\leq 2^{\frac{p-1}{p}}\left(c_{28}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}}+c_{17}\alpha\right)$$

with $c_{17}=3+c_{25}+c_{26}+c_{27}$.

Proof. We denote $w=\mathrm{W}_p(\mu,\nu)$ and $\alpha=(\frac{w}{r^{d-1}})^{\frac{1}{2}}$. Let us cut the integral as follows:

$$\int\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})d\pi(x,y)=\int_A+\int_B+\int_C\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)$$

where $A=\{(x,y),\|x-y\|\geq\alpha\}$, $B=\{(x,y),\|x-y\|<\alpha \text{ and } \lambda(x)>r\}$ and $C=\{(x,y),\|x-y\|<\alpha \text{ and } \lambda(x)\leq r\}$.

Term A: We use the following loose majoration:

$$\begin{aligned}\mathrm{W}_1(\overline{\mu_x},\overline{\nu_y}) &\leq \mathrm{W}_1(\overline{\mu_x},\delta_x)+\mathrm{W}_1(\delta_x,\delta_y)+\mathrm{W}_1(\delta_y,\overline{\nu_y})\\ &\leq r+\|x-y\|+r\end{aligned}$$

to obtain $\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\leq 2^{p-1}\big((2r)^p+\|x-y\|^p\big)$ and

$$\begin{aligned}\int_A\mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y) &\leq \int_A 2^{p-1}\big((2r)^p+\|x-y\|^p\big)\mathrm{d}\pi(x,y)\\ &\leq 2^{p-1}(2r)^p\pi(A)+\int 2^{p-1}\|x-y\|^p\,\mathrm{d}\pi(x,y)\\ &= 2^{p-1}(2r)^p\pi(A)+2^{p-1}w^p.\end{aligned}$$

But $\pi(A) = \pi(\{(x,y), \|x-y\| > \alpha) = \pi\left(\{(x,y), \|x-y\|^p > \alpha^p)\}\right) \leq \left(\frac{w}{a}\right)^p$ by Markov inequality. Therefore,

$$\begin{aligned}\int_A \mathrm{W}_1^p(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x,y) &\leq 2^{p-1}(2r)^p\left(\frac{w}{\alpha}\right)^p + 2^{p-1}w^p \\ &= 2^{p-1}(2r^d\alpha)^p + 2^{p-1}w^p,\end{aligned}$$

where we used $r\frac{w}{\alpha} = r^d\alpha$ on the last line.

Term B: On the event B, we write

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y}) \leq \mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) + \mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}).$$

Since $\lambda(x) > r$, Lemma IV.46 and Lemma IV.48 give $\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{25}\|x-y\|$ and $\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{27}\alpha$. We deduce that

$$\begin{aligned}\int_B \mathrm{W}_1^p(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x,y) &\leq 2^{p-1}\int_B (c_{25}\|x-y\|)^p + (c_{27}\alpha)^p \mathrm{d}\pi(x,y) \\ &\leq 2^{p-1}(c_{25}w)^p + 2^{p-1}(c_{27}\alpha)^p.\end{aligned}$$

Term C: We proceed as for Term B, but using Lemmas IV.47 and IV.49 instead of Lemmas IV.46 and IV.48. This yields

$$\begin{aligned}\mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y}) &\leq \mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) + \mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \\ &\leq c_{26}r^{\frac{1}{2}}\|x-y\|^{\frac{1}{2}} + c_{28}r^{\frac{1}{2}}\alpha^{\frac{1}{2}},\end{aligned}$$

and we deduce that

$$\begin{aligned}\int_C \mathrm{W}_1^p(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x,y) \leq &\int_C 2^{p-1}\left(c_{26}r^{\frac{1}{2}}\|x-y\|^{\frac{1}{2}}\right)^p \mathrm{d}\pi(x,y) \\ &+ 2^{p-1}\pi(C)\left(c_{28}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}\right)^p. \qquad \text{(IV.30)}\end{aligned}$$

On the one hand, we have $\int_C \|x-y\|^{\frac{p}{2}}\,\mathrm{d}\pi(x,y) \leq \int_{E\times E}\|x-y\|^{\frac{p}{2}}\,\mathrm{d}\pi(x,y)$, and by Jensen's inequality,

$$\int_{E\times E}\|x-y\|^{\frac{p}{2}}\,\mathrm{d}\pi(x,y) \leq (w^p)^{\frac{1}{2}}.$$

On the other hand, by definition of C, we have $\pi(C) \leq \mu(\lambda^r)$. Combined with Equation (IV.30), we obtain

$$\int_C \mathrm{W}_1(\overline{\mu_x}, \overline{\nu_y})\mathrm{d}\pi(x,y) \leq 2^{p-1}\left(c_{26}r^{\frac{1}{2}}w^{\frac{1}{2}}\right)^p + 2^{p-1}\mu(\lambda^r)\left(c_{28}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}\right)^p.$$

To conclude the proof, we write

$$\begin{aligned}\int \mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y) &= \int_A + \int_B + \int_C \mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\\ &\leq 2^{p-1}(2r^d\alpha)^p + 2^{p-1}w^p + 2^{p-1}(c_{25}w)^p + 2^{p-1}(c_{27}\alpha)^p\\ &\quad + 2^{p-1}\left(c_{26}r^{\frac{1}{2}}w^{\frac{1}{2}}\right)^p + 2^{p-1}\mu(\lambda^r)\left(c_{28}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}\right)^p.\end{aligned}$$

We use the inequality $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$, where $a, b \geq 0$, to get

$$\begin{aligned}&\left(\int \mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}}\\ &\leq 2^{\frac{p-1}{p}}\left(2r^d\alpha + w + c_{25}w + c_{27}\alpha + c_{26}r^{\frac{1}{2}}w^{\frac{1}{2}} + \mu(\lambda^r)^{\frac{1}{p}}c_{28}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}\right)\\ &\leq 2^{\frac{p-1}{p}}\left(c_{28}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + \left(2r^d + c_{26}r^{\frac{d+1}{2}} + c_{27}\right)\alpha + (1+c_{25})w\right),\end{aligned}$$

where we used $c_{26}r^{\frac{1}{2}}w^{\frac{1}{2}} = c_{26}r^{\frac{d+1}{2}}\alpha$ on the the last line. This proves the first result.

If we suppose $r \leq 1$, we can use the inequalities $r^d \leq r^{\frac{d+1}{2}} \leq 1$ and $w = \alpha r^{\frac{d-1}{2}}w^{\frac{1}{2}} \leq \alpha$ to obtain the simplified expression

$$\left(\int \mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}}\left(c_{28}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + (3 + c_{25} + c_{26} + c_{27})\,\alpha\right).$$

Remark IV.32. On Term C, we could have used the inequality $\mathrm{W}_1(\overline{\mu_x},\overline{\nu_y}) \leq r + \|x-y\| + r$ to obtain

$$\begin{aligned}\int_C \mathrm{W}_1^p(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y) &\leq 2^{p-1}\int_C (2r)^p + \|x-y\|^p\,\mathrm{d}\pi(x,y)\\ &\leq 2^{p-1}(2r)^p\pi(C) + 2^{p-1}w^p.\end{aligned}$$

Following the rest of the proof, and under the assumption $r \leq 1$, we eventually obtain

$$\left(\int \mathrm{W}_1(\overline{\mu_x},\overline{\nu_y})\mathrm{d}\pi(x,y)\right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}}\left(2r\mu(\lambda^r)^{\frac{1}{p}} + c'_{17}\alpha\right)$$

with $c'_{17} = 4 + c_{25} + c_{27}$.

Note that here, in the term $r\mu(\lambda^r)^{\frac{1}{p}}$, the exponent over r is better than in Lemma IV.31, which is $r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}}$. However, we prefer to keep the term $\alpha^{\frac{1}{2}}$, for it goes to zero as w does.

IV.3.4 An approximation theorem

Let us recall the definitions of Subsection IV.3.1: the exact lifted measure is $\check{\mu}_0 = (u_*\mu_0)(x_0)\otimes\left\{\delta_{\frac{1}{d+2}p_{T_x\mathcal{M}}}\right\}$, and the lifted measure associated to ν is $\check{\nu} = \nu(x)\otimes\left\{\delta_{\overline{\Sigma}_\nu(x)}\right\}$.

We are now able to state that $\check{\nu}$ is close to $\check{\mu}_0$, that is, $\check{\nu}$ is a consistent estimator of $\check{\mu}_0$, in Wasserstein distance.

Theorem IV.33. *Assume that $\mathcal{M}_0$ and μ_0 satisfy Hypotheses 1, 2, 3. Let ν be any probability measure. Denote $w = \mathrm{W}_p(\mu, \nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$\mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{19} \mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{15} r + \gamma c_{16} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w$$

where $c_{19} = 2 + \frac{1}{2} c'_{16}$.

Proof. It is a direct consequence of the triangle inequality for $\mathrm{W}_{p,\gamma}$ and Propositions IV.27 and IV.29.

Remark IV.34. In the case where $\mathcal{M}_0$ is embedded, we have seen in Proposition IV.10 that the normal reach λ is bounded below by reach $(\mathcal{M}) > 0$. In particular, $\mu(\lambda^r)$ is zero for r small enough. We deduce an approximation result: if $(\nu_i)_{i \geq 0}$ is a sequence of probability measures such that $w_i = \mathrm{W}_p(\mu, \nu_i)$ goes to zero, and if we choose a sequence of radii $(r_i)_{i \geq 0}$ such that $\left(w_i / r_i^{d+1}\right)_{i \geq 0}$ goes to zero, then $\mathrm{W}_{p,\gamma}(\check{\nu}_i, \check{\mu}_0)$ goes to zero too.

More generally, $\mathrm{W}_{p,\gamma}(\check{\nu}_i, \check{\mu}_0)$ goes to zero if we only assume that $\mathcal{M}_0$ satisfies Hypothesis 4. This is stated in the following corollary.

In order to simplify the results of the following section, we shall use a weaker version of the theorem.

Corollary IV.35. *Let $r > 0$. Assume that $\mathcal{M}_0$ and μ_0 satisfy Hypotheses 1, 2, 3 and Hypothesis 4 with $r_3 \geq r$. Let ν be any probability measure. Denote $w = \mathrm{W}_p(\mu, \nu)$. Suppose that $r \leq \frac{1}{2\rho} \wedge 1$ and $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+2}$. Then*

$$\mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \left(1 + \gamma c_{20}\right) r^{\frac{1}{p}}$$

with $c_{20} = c_{19}(c_3)^{\frac{1}{p}} + c_{16} + c_{15}$.

Proof. According to Theorem IV.33, we have

$$\mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{19} \mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{15} r + \gamma c_{16} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w.$$

Note that the assumption $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+2}$ yields $\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} \leq \frac{r}{4} \leq r$. Besides, $r \leq 1$ yields $w \leq \left(\frac{r}{4}\right)^{d+2} \leq \frac{r}{4} \leq \frac{r}{2}$. Finally, Hypothesis 4 gives $\mu(\lambda^r) \leq c_3 r$, and we deduce the result thanks to the rough majoration $r \leq r^{\frac{1}{p}}$:

$$\begin{aligned} \mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) &\leq \gamma c_{19} (c_3 r)^{\frac{1}{p}} + \gamma c_{15} r + \gamma c_{16} r + r \\ &\leq \left(\gamma c_{19} (c_3)^{\frac{1}{p}} + \gamma c_{15} + \gamma c_{16} + 1\right) r^{\frac{1}{p}}. \end{aligned}$$

IV.4 Topological inference with the lifted measure

Based on the results of the last section, we show how the lifted measure $\check{\nu}$ can be used to infer the homotopy type of $\check{\mathcal{M}}$, or to estimate the persistent homology of $\check{\mu}_0$.

IV.4.1 Overview of the method

Let us recall the results obtained so far. Assume that the immersion $u\colon \mathcal{M}_0 \to \mathcal{M}$ and the measure μ_0 satisfy the Hypotheses 1, 2 and 3. Our goal is to estimate the exact lifted measure $\check{\mu}_0$ on $E \times \mathrm{M}(E)$, since its support is the submanifold $\check{\mathcal{M}}$, which is diffeomorphic to $\mathcal{M}_0$.

To do so, we suppose that we are observing a measure ν on E. No assumptions are made on ν. Our results only depends on the Wasserstein distance

$$w = \mathrm{W}_p(\mu, \nu),$$

where $\mu = u_*\mu_0$. Recall that the measure $\check{\mu}_0$ is defined as (Equation (IV.20)):

$$\check{\mu}_0 = (u_*\mu_0)(x_0) \otimes \left\{\delta_{\frac{1}{d+2}p_{T_x\mathcal{M}}}\right\}.$$

To approximate $\check{\mu}_0$, pick a parameter $r > 0$ and consider the lifted measure $\check{\nu}$ built on ν (Definition IV.23):

$$\check{\nu} = \nu(x) \otimes \left\{\delta_{\overline{\Sigma}_\nu(x)}\right\}.$$

Choose $\gamma > 0$. Endow the space $E \times \mathrm{M}(E)$ with the norm $\|\cdot\|_\gamma$ (Equation (IV.21)), and consider the Wasserstein distance $\mathrm{W}_{p,\gamma}(\cdot,\cdot)$ between measures on $E \times \mathrm{M}(E)$ (Equation (IV.22)). We quantify the quality of the approximation by the Wasserstein distance

$$\mathrm{W}_{p,\gamma}(\check{\mu}_0, \check{\nu}).$$

According to Theorem IV.33, we have

$$\mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{19}\mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{15} r + \gamma c_{16}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w$$

as long as the parameter r satisfies

$$4\left(\frac{w}{c_9 \wedge 1}\right)^{\frac{1}{d+1}} \;\leq\; r \;\leq\; \frac{1}{2\rho} \wedge 1.$$

Under Hypothesis 4, Corollary IV.35 gives a weaker form of this result. We have

$$\mathrm{W}_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \big(1 + \gamma c_{20}\big) r^{\frac{1}{p}}$$

as long as the parameter r satisfies

$$4\left(\frac{w}{c_9 \wedge 1}\right)^{\frac{1}{d+2}} \;\leq\; r \;\leq\; \frac{1}{2\rho} \wedge r_3 \wedge 1.$$

In the following subsections, we show how these results lead to consistent estimations of $\mathcal{M}_0$ and its homology. Namely, we can estimate the homotopy type of $\check{\mathcal{M}}$, and hence of $\mathcal{M}_0$, by considering the sublevel sets of the DTM $\mathrm{d}_{\check{\nu},m,\gamma}$ (Corollary IV.38). The notation $\mathrm{d}_{\check{\nu},m,\gamma}$ corresponds to the DTM, defined in Chapter III, seen in the ambient space $\left(E \times \mathrm{M}(E), \|\cdot\|_\gamma\right)$. Besides, we can estimate the persistent homology of the DTM-filtration $W_\gamma[\check{\mu}_0]$ with the filtration $W_\gamma[\check{\nu}]$ (Corollary IV.42). Here, $W_\gamma[\cdot]$ corresponds to the DTM-filtration in the ambient space $\left(E \times \mathrm{M}(E), \|\cdot\|_\gamma\right)$.

Example IV.36. Let $\mathcal{M}$ be the lemniscate of Bernoulli of diameter 2. It is the immersion of a circle $\mathcal{M}_0$. We observe a 100-sample X of $\mathcal{M}$ (Figure IV.14). Experimentally, we computed the Hausdorff distance $\mathrm{d_H}(\mathcal{M}, X) \approx 0.026$. Let μ be the Hausdorff measure on $\mathcal{M}$ and ν the empirical measure on X. We choose the parameter $p = 2$. Their Wasserstein distance is approximately $\mathrm{W}_2(\mu, \nu) \approx 0.015$.

Figure IV.14: Left: The lemniscate $\mathcal{M}$. **Right:** The set X, a 100-sample of $\mathcal{M}$.

For each point x of X, we compute the matrix $\overline{\Sigma}_\nu(x)$ with parameter $r = 0.5$ and 0.1. This matrix is used as an estimator of the tangent space $T_x\mathcal{M}$. In order to observe the quality of this estimation, we represent on Figure IV.15 (first row) the principal axes of $\overline{\Sigma}_\nu(x)$ for some x. On the second row are represented the distances $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$. One sees that $r = 0.1$ yields a better approximation. However, the estimation is still biased next to the self-intersection points of $\mathcal{M}$.

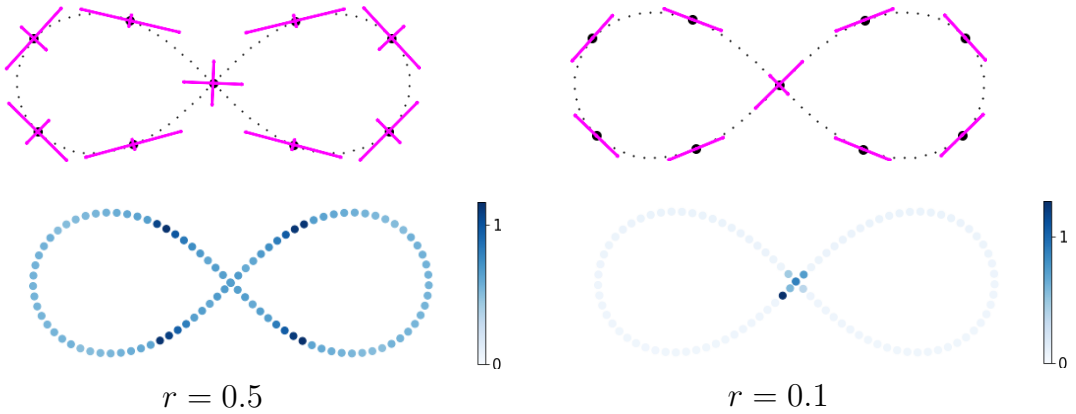


Figure IV.15: First row: The eigenvectors of $\overline{\Sigma}_\nu(x)$ for some $x \in X$, weighted with their corresponding eigeinvalue. **Second row:** color representation of the distances $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$.

Now we choose the parameter $\gamma = 2$. For $r = 0.5$ and 0.1, we consider the lifted measures built on ν, respectively denoted $\check{\nu}^{0.5}$ and $\check{\nu}^{0.1}$. They are measure on the

lift space $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, which is endowed with the norm $\|\cdot\|_\gamma$. We computed the Wasserstein distances:

$$\mathrm{W}_{2,\gamma}\left(\check{\mu}_0, \check{\nu}^{0.5}\right) \approx 0.674 \qquad \text{and} \qquad \mathrm{W}_{2,\gamma}\left(\check{\mu}_0, \check{\nu}^{0.1}\right) \approx 0.200.$$

In comparison, even with a small parameter r, the Hausdorff distance between their support is still large:

$$\mathrm{d_H}\left(\check{\mathcal{M}}, \mathrm{supp}(\check{\nu}^{0.5})\right) \approx 1.142 \qquad \text{and} \qquad \mathrm{d_H}\left(\check{\mathcal{M}}, \mathrm{supp}(\check{\nu}^{0.1})\right) \approx 1.273.$$

These sets are represented in Figure IV.16. Observe that, at the center of the graphs, the measures $\check{\nu}^{0.5}$ and $\check{\nu}^{0.1}$ deviate from the set $\check{\mathcal{M}}$.

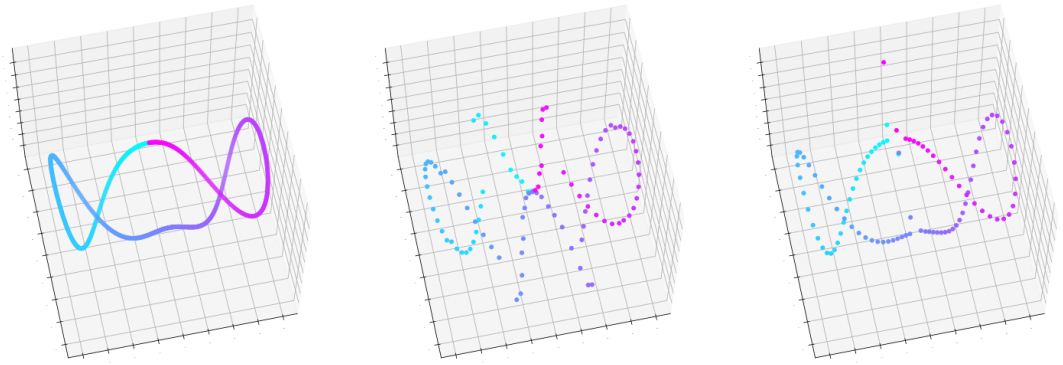


Figure IV.16: Left: The lifted lemniscate $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. **Center:** The set $\mathrm{supp}(\check{\nu}^{0.5})$ projected in the same 3-dimensional subspace. **Right:** Same for $\mathrm{supp}(\check{\nu}^{0.1})$.

Example IV.37. Let $u\colon \mathcal{M}_0 \to \mathcal{M}$ be the figure-8 immersion of the torus in $\mathbb{R}^3$, represented in Figure IV.17. It can be parametrized by rotating a lemniscate around an axis, while forming a full twist. The self-intersection points of this immersion corresponds to the inner circle formed by the center of the lemniscate. These are the points x of $\mathcal{M}$ such that their normal reach $\lambda(x)$ is zero.

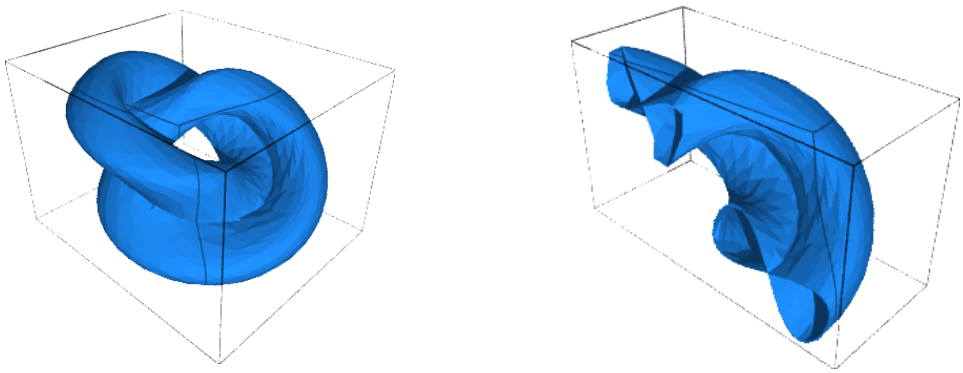


Figure IV.17: Left: The immersion $\mathcal{M}$ of the torus. **Right:** A section of $\mathcal{M}$. One sees the inner lemniscate.

Let $\check{\mathcal{M}}$ be the lift of $\mathcal{M}_0$. It is a submanifold of $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$. One cannot embed $\check{\mathcal{M}}$ in $\mathbb{R}^3$ by performing a PCA. However, we can try to visualize $\check{\mathcal{M}}$ by

considering a small section of it. Figure IV.18 represents a subset of $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. One sees that it does not self-intersect.

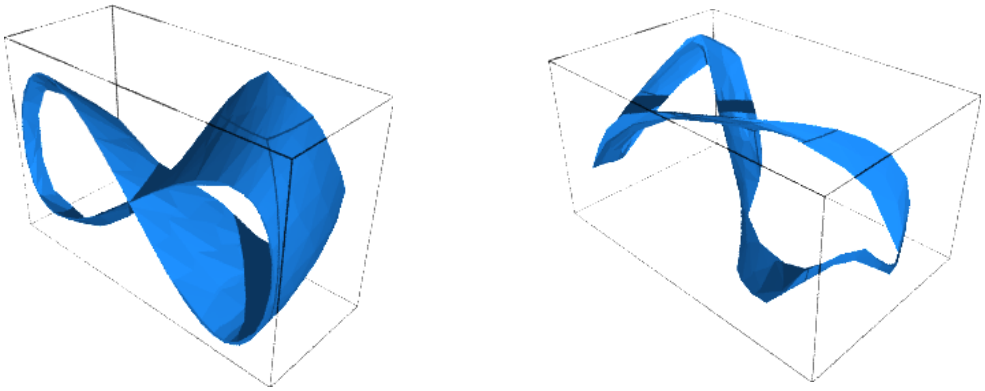


Figure IV.18: Left: A section of $\mathcal{M}$. **Right:** The corresponding section of $\check{\mathcal{M}}$, projected in a 3-dimensional subspace via PCA. Observe that it does not self-intersect.

In order to fit in the context of our study, let μ be the Hausdorff measure on $\mathcal{M}$. We observe a 9000-sample X of $\mathcal{M}$, and consider its empirical measure ν. The set X is depicted in Figure IV.19. Choose the parameter $p = 1$. We compute the Wasserstein distance $\mathrm{W}_1(\mu, \nu) \approx 0.070$ and the Hausdorff distance $\mathrm{d_H}(\mathcal{M}, X) = 0.083$.

Let $r = 0.09$. In order to observe the estimation of tangent spaces by local covariance matrices $\overline{\Sigma}_\nu(x)$ with parameter r, we represent on Figure IV.19 the points x such that the distance $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}}$ is greater than 2. Observe that the estimation is biased next to the self-intersection circle of $\mathcal{M}$. Last, let us choose the parameter $\gamma = 2$, and consider the lifted measure $\check{\nu}$. We have $\mathrm{W}_1(\check{\mu}_0, \check{\nu}) \approx 0.986$. In comparison, the Hausdorff distance between their support is large: $\mathrm{d_H}\left(\check{\mathcal{M}}, \mathrm{supp}(\check{\nu})\right) \approx 2.188$.

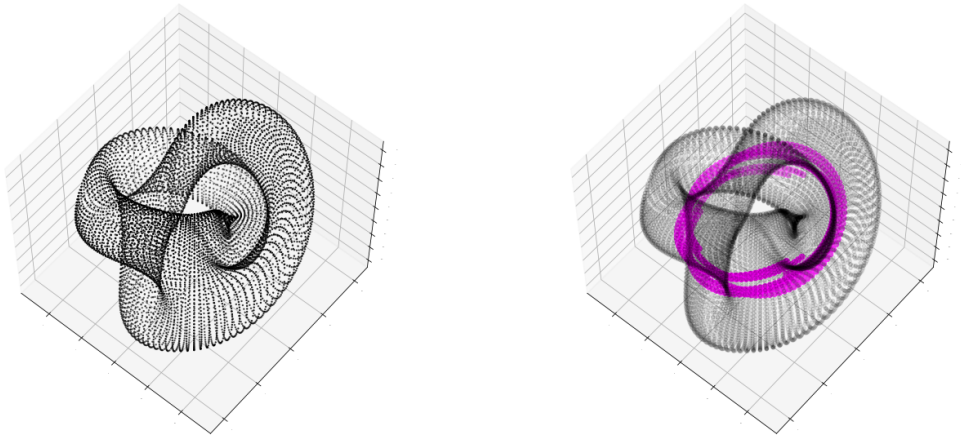


Figure IV.19: Left: The set X, a sample of $\mathcal{M}$. **Right:** The set X, where $x \in X$ is colored in magenta if $\left\|\overline{\Sigma}_\nu(x) - \frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} \geq 2$.

IV.4.2 Homotopy type estimation with the DTM

In this subsection, we use the DTM, as defined in Subsection II.5.2, to infer the homotopy type of $\check{\mathcal{M}}$ from the lifted measure $\check{\nu}$. We shall use the DTM on $\check{\nu}$, which lives in the space $E \times \mathrm{M}(E)$ endowed with the norm $\|\cdot\|_\gamma$. It is denoted $d_{\check{\nu},m,\gamma}$.

In order to apply Theorem II.27 in our setting, we have to consider geometric quantities associated to the submanifold $\check{\mathcal{M}}$. For every $\gamma > 0$, we denote by $\mathrm{reach}_\gamma(\check{\mathcal{M}})$ the reach of $\check{\mathcal{M}}$. Besides, note that the map $\check{u}$ itself satisfies the Hypotheses 2 and 3, as the immersion u does. The corresponding constants are denoted $\check{\rho}_\gamma$, $\check{L}_{0,\gamma}$, $\check{f}_{\min,\gamma}$ and $\check{f}_{\max,\gamma}$. We point out that the constant $\check{\rho}_\gamma$ cannot be deduced from ρ: the first one can be arbitrary large or small compared to the second one, even with γ being fixed. This remark holds for the other constants.

However, we can use the results of Section IV.2 in this context. Proposition IV.17 applied to $\check{\mu}_0$ gives a constant $\check{c}_{9,\gamma}$ such that $\check{\mu}_0(\overline{\mathcal{B}}(\check{x}, r)) \geq \check{c}_{9,\gamma} r^d$ for all $r \leq \frac{1}{2\check{\rho}_\gamma}$. Namely, $\check{c}_{9,\gamma} = \check{f}_{\min,\gamma} J_{\min} V_d$. These constants being given, we propose a way to tune the parameters r, γ, m and t in such a way that the t-sublevel set $d^t_{\check{\nu},m,\gamma}$ of the DTM captures the homotopy type of $\check{\mathcal{M}}$, or equivalently, of $\mathcal{M}_0$.

Corollary IV.38. *Assume that $\mathcal{M}_0$ and μ_0 satisfy Hypotheses 1, 2, 3 and 4. Let ν be any probability measure on E. Denote $w = \mathrm{W}_2(\mu, \nu)$. Choose $r > 0$, $\gamma > 0$ and $m \in (0, 1)$ such that*

- $4\left(\frac{w}{c_9 \wedge 1}\right)^{\frac{1}{d+2}} \leq r \leq \frac{1}{2\rho} \wedge r_3 \wedge 1$
- $m \leq \frac{c_{5,\gamma}}{(2\check{\rho}_\gamma)^d}$ *and*
- $(1 + \gamma c_{20}) r^{\frac{1}{2}} \leq m^{\frac{1}{2}} \left(\frac{\mathrm{reach}_\gamma(\check{\mathcal{M}})}{9} - \left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}} \right)$.

Define ϵ and choose t as follows:

$$\epsilon = \left(\frac{m}{c_{5,\gamma}}\right)^{\frac{1}{d}} + (1 + \gamma c_{20}) \left(\frac{r}{m}\right)^{\frac{1}{2}} \qquad \text{and} \qquad t \in \left[4\epsilon, \mathrm{reach}_\gamma(\check{\mathcal{M}}) - 3\epsilon\right].$$

Then the sublevel set of the DTM $\mathrm{d}^t_{\check{\nu},m,\gamma}$ is homotopy equivalent to $\mathcal{M}_0$.

Proof. In order to fit in the context of Theorem II.27, we have to consider the usual Euclidean norm $\|\cdot\|$ on $E \times \mathrm{M}(E)$. It corresponds to the norm $\|\cdot\|_\gamma$ with $\gamma = 1$. For a general parameter $\gamma > 0$, consider the application $i_\gamma \colon E \times \mathrm{M}(E) \to E \times \mathrm{M}(E)$ defined as

$$i_\gamma \colon (x, A) \mapsto (x, \gamma A).$$

A computation shows that, for every probability measures α, β on $E \times \mathrm{M}(E)$, we have

$$\mathrm{W}_{2,\gamma}(\alpha, \beta) = \mathrm{W}_2\big((i_\gamma)_* \alpha, (i_\gamma)_* \beta\big),$$

where W_2 denotes the 2-Wasserstein distance on $E \times \mathrm{M}(E)$ endowed with the usual Euclidean norm $\|\cdot\|$. Corollary IV.35 then reformulates as

$$\mathrm{W}_2\big((i_\gamma)_*\check{\mu}_0, (i_\gamma)_*\check{\nu}\big) \leq (1+\gamma c_{20})r^{\frac{1}{2}}.$$

Besides, consider the set

$$\check{\mathcal{M}}_\gamma = i_\gamma(\check{\mathcal{M}}) = \{(x, \gamma A), (x, A) \in \check{\mathcal{M}}\}.$$

It is direct to see that

$$\mathrm{reach}_\gamma(\check{\mathcal{M}}) = \mathrm{reach}\left(\check{\mathcal{M}}_\gamma\right),$$

where we recall that $\mathrm{reach}_\gamma(\check{\mathcal{M}})$ is the reach of $\check{\mathcal{M}}$ with respect to the norm $\|\cdot\|_\gamma$, and $\mathrm{reach}\left(\check{\mathcal{M}}_\gamma\right)$ is the reach of $\check{\mathcal{M}}_\gamma$ with respect to the usual norm $\|\cdot\|$ on $E \times \mathrm{M}(E)$. Finally, consider the DTM $\mathrm{d}_{(i_\gamma)_*\check{\nu},m}$ with respect to the usual Euclidean norm. Observe that, for every $t \geq 0$, the sublevel sets of the DTM $\mathrm{d}_{(i_\gamma)_*\check{\nu},m}$ and $\mathrm{d}_{\check{\nu},m,\gamma}$ are linked via

$$\mathrm{d}^t_{\check{\nu},m} = i_\gamma\left(\mathrm{d}^t_{\check{\nu},m,\gamma}\right).$$

In particular, they share the same homotopy type.

Now we obtain the result as a consequence of Theorem II.27 applied to $(i_\gamma)_*\check{\mu}_0$ and $(i_\gamma)_*\check{\nu}$. Let us verify that the assumptions of the theorem are satisfied. Our assumption about m ensures that

$$\left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}} \leq \frac{1}{2\check{\rho}_\gamma},$$

hence by Proposition IV.17 we get $\check{\mu}_0(\mathcal{B}(x,r)) \geq \check{c}_{5,\gamma} r^d$ for all $x \in \mathrm{supp}(\check{\mu}_0)$ and $r < \left(\frac{m}{\check{c}_{5,\gamma}}\right)^{\frac{1}{d}}$. Moreover, the assumption about $(1+\gamma c_{20})r^{\frac{1}{2}}$ ensures that

$$\mathrm{W}_2\big((i_\gamma)_*\check{\mu}_0, (i_\gamma)_*\check{\nu}\big) \leq m^{\frac{1}{2}}\left(\frac{\mathrm{reach}_\gamma(\check{\mathcal{M}})}{9} - \left(\frac{m}{c_{5,\gamma}}\right)^{\frac{1}{d}}\right)$$

is satisfied, since $\mathrm{W}_2\big((i_\gamma)_*\check{\mu}_0, (i_\gamma)_*\check{\nu}\big) \leq (1+\gamma c_{20})r^{\frac{1}{2}}$ by Corollary IV.35.

Example IV.39. Let $\mathcal{M}$ be the lemniscate of Bernoulli, as in Example IV.36. Suppose that μ is the uniform distribution on $\mathcal{M}$, and ν is the empirical measure on a 500-sample of $\mathcal{M}$. We choose the parameters $\gamma = 2$, $r = 0.03$ and $m = 0.01$. Let $\check{\nu}$ be the lifted measure associated to ν.

Figure IV.20 represents set the $\mathrm{supp}(\check{\nu})$, and the values of the DTM $\mathrm{d}_{\check{\nu},m,\gamma}$ on it. Observe that the anomalous points, i.e., points for which the local covariance matrix is not well estimated, have large DTM values.

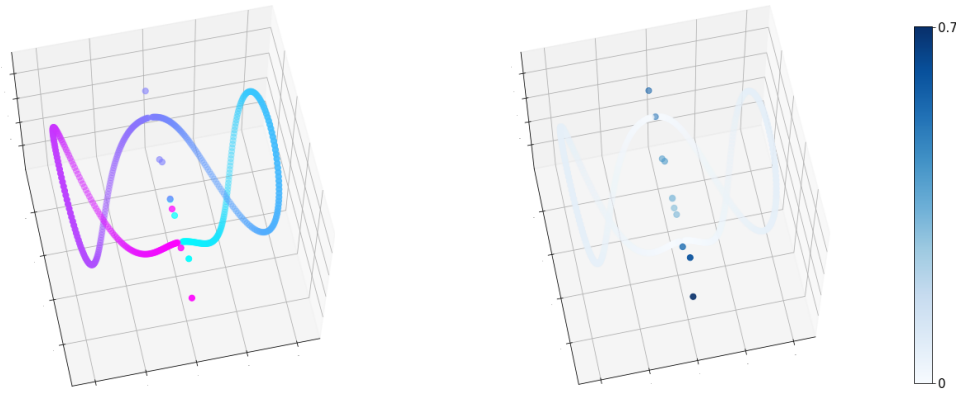


Figure IV.20: Left: The set $\mathrm{supp}(\check{\nu}) \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, projected in a 3-dimensional subspace via PCA. **Right:** The set $\mathrm{supp}(\check{\nu})$ with colors indicating the value of the DTM $\mathrm{d}_{\check{\nu},m,\gamma}$.

IV.4.3 Persistent homology with DTM-filtrations

In this subsection, we aim to estimate the DTM-filtration of $\check{\mu}_0$, as defined in Chapter III, from ν. We shall use the DTM-filtration on $\check{\nu}$, denoted $W_\gamma[\check{\nu}]$, with respect to the ambient norm $\|\cdot\|_\gamma$ on $E \times \mathrm{M}(E)$. We use the notations $\check{\rho}_\gamma$ and $c_{5,\gamma}$ of the previous subsection.

We first recall the definition of the DTM-filtrations on $E = \mathbb{R}^n$, presented in the previous chapter. We only consider the case $p = 1$. Let μ be any measure on E, and $m \in [0, 1)$. For every $t \in T$, consider the set

$$W^t[\mu] = \bigcup_{x \in \mathrm{supp}(\mu)} \overline{\mathcal{B}}\left(x, (t - \mathrm{d}_{\mu,m}(x))^+\right),$$

where $\overline{\mathcal{B}}(x, r^+)$ denotes the closed ball of center x and of radius r if $r \geq 0$, or denotes the empty set if $r < 0$. The family $W[\mu] = (W^t[\mu])_{t \geq 0}$ is a filtration of E. It is called the DTM-filtration with parameters $(\mu, m, 1)$. Define the quantity

$$c(\mu) = \sup_{x \in \mathrm{supp}(\mu)} \mathrm{d}_{\mu,m}(x)$$

The term $c(\mu)$ is to be seen as a quantity controlling the regularity of μ. In particular, if μ is the uniform measure on a submanifold, it goes to 0 as m does, as shown by the following lemma.

Lemma IV.40. *Suppose that μ satisfies the (a, d)-standard assumption (defined in Equation (III.4)). Then $c(\mu) \leq c_{22} m^{\frac{1}{d}}$ with $c_{22} = a^{-\frac{1}{d}}$.*

We restate a stability result we obtained in the previous chapter.

Corollary IV.41. *Let μ, ν with $\mathrm{W}_2(\mu, \nu) = w \leq \frac{1}{4}$. Suppose that μ satisfies the (a, d)-standard assumption (defined in Equation (III.4)). Then*

$$d_i\left(W[\mu], W[\nu]\right) \leq c_{23} \left(\frac{w}{m}\right)^{\frac{1}{2}} + 2c_{22} m^{\frac{1}{d}}$$

with $c_{23} = 8\mathrm{diam}(\mathrm{supp}(\mu)) + 5$.

We now apply these results in our context.

Corollary IV.42. *Assume that* $\mathcal{M}_0$ *and* μ_0 *satisfy Hypotheses 1, 2, 3 and 4. Let* ν *be any probability measure. Denote* $\mathrm{W}_2(\mu, \nu) = w$. *Choose* $r > 0$, $\gamma > 0$ *and* $m \in (0, 1)$ *such that*

- $4\left(\frac{w}{c_9 \wedge 1}\right)^{\frac{1}{d+2}} \leq r \leq \frac{1}{2\rho} \wedge r_3 \wedge 1$,
- $m \leq \frac{c_{5,\gamma}}{(2\check{\rho}_\gamma)^d}$,
- $\big(1 + \gamma c_{20}\big) r^{\frac{1}{2}} \leq \frac{1}{4}$.

Then we have a bound on the interleaving distance between the DTM-filtrations:

$$\mathrm{d_i}\left(W_\gamma[\check{\mu}_0], W_\gamma[\check{\nu}]\right) \leq \check{c}_{23,\gamma}(1 + \gamma c_{20})^{\frac{1}{2}} m^{-\frac{1}{2}} r^{\frac{1}{4}} + 2\check{c}_{22,\gamma} m^{\frac{1}{d}},$$

where $\check{c}_{23,\gamma} = 8\mathrm{diam}(\mathcal{M}) + 8\gamma + 5$ *and* $\check{c}_{22,\gamma} = (c_{5,\gamma})^{-\frac{1}{d}}$.

Proof. As in the proof of Corollary IV.38, let i_γ be the map $i_\gamma \colon (x, A) \mapsto (x, \gamma A)$. Let $W[\cdot]$ denotes the DTM-filtration on $\check{\nu}$ with respect to the usual Euclidean norm. That is, the filtration $W[\cdot]$ corresponds to $W_\gamma[\cdot]$ with $\gamma = 1$. A computation shows that the filtration $W[(i_\gamma)_* \check{\nu}]$ and $W_\gamma[\check{\nu}]$ are linked via

$$W[(i_\gamma)_* \check{\nu}] = i_\gamma\left(W_\gamma[\check{\nu}]\right).$$

Now let $\check{w} = \mathrm{W}_2((i^\gamma)_* \check{\mu}_0, (i^\gamma)_* \check{\nu})$. We have $\check{w} = \mathrm{W}_{2,\gamma}(\check{\mu}_0, \check{\nu})$, hence Corollary IV.35 gives

$$\check{w} \leq \big(1 + \gamma c_{20}\big) r^{\frac{1}{2}}. \tag{IV.31}$$

Moreover, we can apply Corollary IV.41 to $\mu = (i^\gamma)_* \check{\mu}_0$ and $\nu = (i^\gamma)_* \check{\nu}$ to get

$$\mathrm{d_i}\left(W[(i^\gamma)_* \check{\mu}_0], W[(i^\gamma)_* \check{\nu}]\right) \leq \check{c}_{23,\gamma}\left(\frac{\check{w}}{m}\right)^{\frac{1}{2}} + 2\check{c}_{22,\gamma} m^{\frac{1}{d}}, \tag{IV.32}$$

where $\check{c}_{23,\gamma} = \big(8\mathrm{diam}(\check{\mathcal{M}}) + 5\big)$ and $c_{22,\gamma} = (c_{5,\gamma})^{-\frac{1}{d}}$. Note that

$$\mathrm{diam}(\check{\mathcal{M}}) \leq \left(\mathrm{diam}(\mathcal{M})^2 + \gamma^2\left(2\frac{1}{2}\right)^2\right)^{\frac{1}{2}} \leq \mathrm{diam}(\mathcal{M}) + \gamma$$

since the matrices $\frac{1}{d+2} p_{T_x\mathcal{M}}$ have norm $\left\|\frac{1}{d+2} p_{T_x\mathcal{M}}\right\|_{\mathrm{F}} = \frac{\sqrt{d}}{d+2} \leq \frac{1}{2}$. Our assumption $m \leq \frac{c_{5,\gamma}}{(2\check{\rho}_\gamma)^d}$ ensures that the condition $\check{\mu}_0(\mathcal{B}(x, r)) \geq \check{c}_{5,\gamma} r^d$ of the Corollary is satisfied. Similarly, the assumption $\big(1 + \gamma c_{20}\big) r^{\frac{1}{2}} \leq \frac{1}{4}$ yields $\check{w} \leq \frac{1}{4}$.

Combining Equations (IV.31) and (IV.32) we get

$$\mathrm{d_i}\left(W[(i^\gamma)_* \check{\mu}_0], W[(i^\gamma)_* \check{\nu}]\right) \leq \check{c}_{23,\gamma}\big(1 + \gamma c_{20}\big)^{\frac{1}{2}} m^{-\frac{1}{2}} r^{\frac{1}{4}} + 2\check{c}_{22,\gamma} m^{\frac{1}{d}}.$$

Now, by using the definition of an interleaving of filtrations, one proves that

$$\mathrm{d_i}\left(W_\gamma[\check{\mu}_0], W_\gamma[\check{\nu}]\right) = \mathrm{d_i}\left(W[(i^\gamma)_*\check{\mu}_0], W[(i^\gamma)_*\check{\nu}]\right),$$

and we obtain the result.

Example IV.43. Say that μ is the uniform measure on the union of five intersecting circles of radius 1. We observe ν, the empirical measure on the point cloud X drawn in Figure IV.21. It consists in 300 points per circle, and 100 anomalous points. Let $p = 1$. Experimentally, we have $\mathrm{W}_1(\mu, \nu) \approx 0.044$.

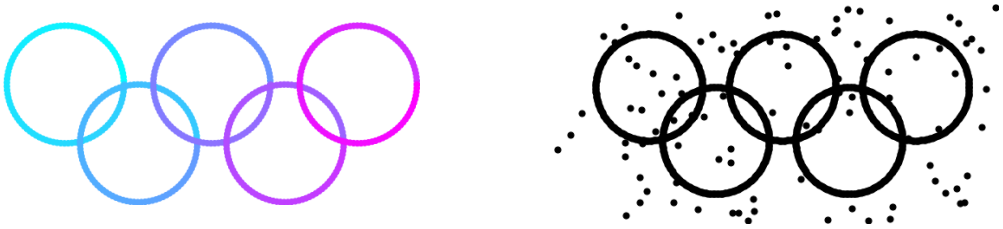


Figure IV.21: Left: the set $\mathcal{M} = \mathrm{supp}(\mu)$. **Right:** The set $X = \mathrm{supp}(\nu)$.

Let $\gamma = 1$. Observe that the barcodes of the DTM-filtration $W[(i^\gamma)_*\check{\mu}_0]$, represented in Figure IV.22, reveal the homology of the disjoint union of five circles—which is the set $\mathcal{M}_0$. Only bars of length larger than 0.1 are displayed. We consider the construction of $\check{\nu}$ with parameter $r = 0.03$, and the DTM-filtration with $m = 0.01$. The barcodes of the DTM-filtration $W[(i^\gamma)_*\check{\nu}]$ are close to the barcodes of $W[(i^\gamma)_*\check{\mu}_0]$. To compare, we also plot the persistence barcodes of the usual Čech filtration on $\mathrm{supp}(\check{\nu})$. Observe that the five connected components do not appear clearly anymore.

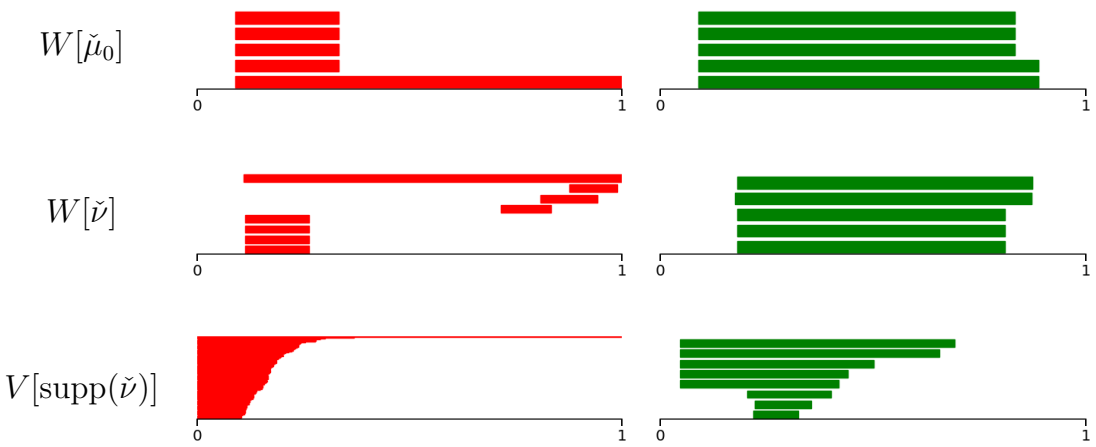


Figure IV.22: First row: Persistence barcode of the 0- and 1-homology of the DTM-filtration on $\check{\mu}_0$. **Second row:** Same for $\check{\nu}$. **Third row:** Persistence barcodes of the usual Čech filtration on $\mathrm{supp}(\check{\nu})$.

IV.5 Conclusion

In this chapter we described a method to estimate the tangent bundle of a manifold $\mathcal{M}_0$ immersed in a Euclidean space, based on a sample of its image. This estimation is stable in Wasserstein distance. Using the DTM, we are able to estimate the homotopy type of $\mathcal{M}_0$. Moreover, via the DTM-filtrations, we can define a filtration of the space $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^n)$, whose persistence module contains information about the homology of $\mathcal{M}_0$.

The robust estimation of tangent bundles of manifolds opens the way to the estimation of other topological invariants than homology groups—such as characteristic classes—a problem that is addressed in the following chapter.

Finally, as we pointed out in Subsection IV.4.2, it would be interesting to understand the geometric quantities associated to the lifted manifold $\check{\mathcal{M}}$ (such as $\check{\rho}_\gamma$, $\check{L}_{0,\gamma}$, $\check{f}_{\min,\gamma}$ and $\check{f}_{\max,\gamma}$) as a function of those associated with the initial manifold $\mathcal{M}_0$ (ρ, L_0, $f_{\min}$ and $f_{\max}$).

IV.A Supplementary material for Section IV.2

Proof of Lemma IV.4 page 137. *Point (1):* We use the triangle inequality, the Pythagorean Theorem and Lemma IV.1 to get

$$\begin{aligned}\|\gamma(t) - x\| &\geq \|(y + tv) - x\| - \|\gamma(t) - (y + tv)\| \\ &\geq \sqrt{\|tv\|^2 + \|y - x\|^2 + 2\langle tv, y - x\rangle} - \frac{\rho}{2}t^2 \\ &\geq \sqrt{t^2 + l^2} - \frac{\rho}{2}t^2.\end{aligned}$$

Now, a computation shows that the function $t \mapsto \sqrt{t^2 + l^2} - \frac{\rho}{2}t^2$ is greater than l on $(0, T_1)$, where $T_1 = \frac{2}{\rho}\sqrt{1 - \rho l}$. Hence for $t \in (0, T_1)$, we have $\phi(t) = \|\gamma(t) - x\|^2 > l^2 = \phi(0)$.

Point (2): Observe that $\dot{\phi}(t) = 2\langle \dot{\gamma}(t), \gamma(t) - x\rangle$, and that

$$\ddot{\phi}(t) = 2\langle \dot{\gamma}(t), \dot{\gamma}(t)\rangle + 2\langle \ddot{\gamma}(t), \gamma(t) - x\rangle .$$

By Cauchy-Schwarz inequality, $\langle \ddot{\gamma}(t), \gamma(t) - x\rangle \geq -\|\ddot{\gamma}(t)\| \, \|\gamma(t) - x\|$. Note that $\langle \dot{\gamma}(t), \dot{\gamma}(t)\rangle = 1$ and $\|\ddot{\gamma}(t)\| \leq \rho$. Hence we get

$$\ddot{\phi}(t) \geq 2(1 - \rho \, \|\gamma(t) - x\|). \tag{IV.33}$$

Now, since $\langle v, y - x\rangle = 0$, we have

$$\begin{aligned}\|\gamma(t) - x\| &\leq \|(y + tv) - x\| + \|\gamma(t) - (y + tv)\| \\ &\leq \sqrt{\|tv\|^2 + \|y - x\|^2} + \frac{\rho}{2}t^2 \\ &= \sqrt{t^2 + l^2} + \frac{\rho}{2}t^2.\end{aligned}$$

A computation shows that the function $t \mapsto \sqrt{t^2+l^2} + \frac{\rho}{2}t^2$ is lower than $\frac{1}{\rho}$ on $(0, T_2)$, where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2-\sqrt{3+\rho^2 l^2}}$. Hence for $t \in (0, T_2)$, we have $\ddot{\phi}(t) \geq 0$. And since $\dot{\phi}(0) = 0$, we have that ϕ is increasing.

Point (3): For all $t \in (0, b)$, it holds that $\|\gamma(t) - x\| \leq r$, hence Equation (IV.33) gives $\ddot{\phi}(t) \geq 2(1-\rho r)$.

Point (4): Assume that $\langle v, y - x\rangle \leq 0$. We still have the inequality

$$\|\gamma(t) - x\| \leq \sqrt{t^2+l^2} + \frac{\rho}{2}t^2. \tag{IV.34}$$

Consider t^*, the first non-negative root of $\sqrt{t^2+l^2} + \frac{\rho}{2}t^2 = r$. According to Equation (IV.34), $b \geq t^*$. Now, a computation gives

$$t^* = \frac{\sqrt{2}}{\rho}\sqrt{1+\rho r - \sqrt{(1+\rho r)^2 - \rho^2(r^2-l^2)}}.$$

Using the inequality $\sqrt{B}-\sqrt{A} = \frac{1}{\sqrt{B}+\sqrt{A}}(B-A) \geq \frac{1}{2\sqrt{B}}(B-A)$, where $A < B$, we get

$$1+\rho r - \sqrt{(1+\rho r)^2 - \rho^2(r^2-l^2)} \geq \frac{1}{2(1+\rho r)}\rho^2(r^2-l^2),$$

and we conclude that $t^* \geq \frac{1}{\sqrt{1+\rho r}}\sqrt{r^2-l^2}$.

Point (5): Assume that $\langle v, y - x\rangle \geq 0$ In the same vein as Point 4, we have $\|\gamma(t) - x\| \geq \sqrt{t^2+l^2} - \frac{\rho}{2}t^2$, and we deduce $b \leq t^*$, where t^* is the first positive root of $\sqrt{t^2+l^2} - \frac{\rho}{2}t^2 = r$. Solving this equation leads to

$$t^* = \frac{\sqrt{2}}{\rho}\sqrt{1-\rho r - \sqrt{(1-\rho r)^2 - \rho^2(r^2-l^2)}}.$$

We use the inequality $\sqrt{B} - \sqrt{A} = \frac{1}{\sqrt{A}+\sqrt{B}}(B-A) \leq \frac{1}{\sqrt{B}}(B-A)$, where $A < B$, to get

$$1-\rho r - \sqrt{(1-\rho r)^2 - \rho^2(r^2-l^2)} \leq \frac{1}{1-\rho r}\rho^2(r^2-l^2)$$

and we conclude that $t^* \leq \frac{\sqrt{2}}{\sqrt{1-\rho r}}\sqrt{r^2-l^2}$.

Proof of Proposition IV.17 page 150. Let $\mathcal{M}^x = \mathcal{M} \cap \overline{\mathcal{B}}(x, r)$ and $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$. Lemma IV.9 does not apply: it is not true that $\mathcal{M}_0^x \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_4(\rho r)r)$. However, we can decompose $\mathcal{M}_0^x$ in connected components $C_0^i, i \in I$.

For every $i \in I$, let z_0^i be a minimizer of $z_0 \mapsto \|z - x\|$ on C_0^i. We have $x - z^i \perp T_{z^i}\mathcal{M}$, hence according to Lemma IV.4 Point 5, $C_0^i \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(z_0^i, \frac{1}{\rho})$. For all $i \in I$, consider μ_0^i, the measure μ_0 restricted to C_0^i, and define $\nu_0^i =$

$(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}_* \mu_0^i$, as in Remark IV.15. The measure ν_0^i admits g_0^i as a density over the d-dimensional Hausdorff measure on $T_{z_0^i}\mathcal{M}_0$, where

$$g_0^i(v) = f_0\big(\overline{\exp}_{z_0}^{\mathcal{M}_0}(v)\big) \cdot J_v \cdot 1_{(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)}(v).$$

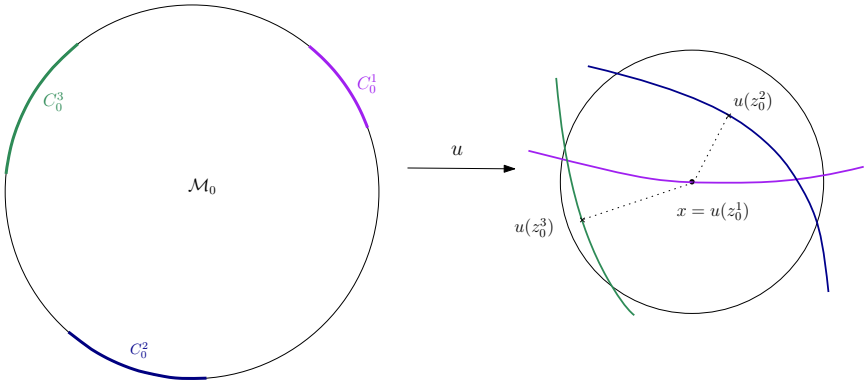


Figure IV.23: The connected components C_0^i.

Point (1): We can write

$$\mu(\overline{\mathcal{B}}(x,r)) = \mu_0(u^{-1}(\overline{\mathcal{B}}(x,r))) = \sum_{i\in I} \mu_0(C_0^i).$$

Let $i_* \in I$ be the index of the connected component of $\mathcal{M}_0^x$ which contains x_0. We have $C_0^{i_*} \supset \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, r)$, and we deduce that

$$\begin{aligned}\mu_0(C_0^{i_*}) &\geq \int_{(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)} g_0^{i_*} \mathrm{d}\mathcal{H}^d \\ &\geq f_{\min} J_{\min} \mathcal{H}^d\big((\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^i)\big) = f_{\min} J_{\min} V_d r^d.\end{aligned}$$

Therefore, $\mu(\overline{\mathcal{B}}(x,r)) \geq f_{\min} J_{\min} V_d r^d$.

Point (2): We now prove the second point.

Step 1: Let us show that the cardinal of I is lower than $\frac{1}{f_{\min} J_{\min} V_d}(\frac{2\rho}{\alpha})^d$, with $\alpha = \sqrt{4-\sqrt{13}}$. We shall prove that for every $i, j \in I$ such that $i \neq j$, $\mathrm{d}_{\mathcal{M}_0}(z_0^i, z_0^j) \geq \frac{\alpha}{\rho}$.

Let γ_0 be a geodesic from z_0^i to z_0^j, with $\gamma(0) = z^i$, $\gamma(T) = z^j$, and $\dot\gamma_0(0) = v_0$. Consider the application $\phi\colon t \mapsto \|\gamma(t) - x\|^2$. Since C_0^i and C_0^j are disjoint connected components, there must be a $t^* < T$ such that $\|\gamma(t^*) - x_0\| > r$. Moreover, according to Lemma IV.4 Point 2, ϕ is increasing on $[0, T_2]$ where $T_2 = \frac{\sqrt{2}}{\rho}\sqrt{2-\sqrt{3+\rho^2 l^2}}$. Since $\phi(T) \leq r$, we deduce that T is greater than T_2. Note that the assumption $r \leq \frac{1}{2\rho}$ yields $T_2 \geq \frac{\alpha}{\rho}$.

This implies that the geodesic balls $\mathcal{B}_{\mathcal{M}_0}(z_0^i, \frac{\alpha}{2\rho})$ are disjoint. Therefore,

$$1 \geq \mu_0\big(\bigcup_i \mathcal{B}_{\mathcal{M}_0}(z_0^i, \frac{\alpha}{2\rho})\big) \geq |I| f_{\min} J_{\min} V_d (\frac{\alpha}{2\rho})^d,$$

and we deduce $|I| \leq \frac{1}{f_{\min} J_{\min} V_d} (\frac{2\rho}{\alpha})^d$.

Step 2: Let $i \in I$, and define $D_0^i = C_0^i \cap u^{-1}(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s))$. Let us show that

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1}\sqrt{6} d V_d \cdot r^{d-1}\sqrt{r^2 - s^2}.$$

Let us distinguish two cases: $l_i \geq s$ or $l_i < s$.

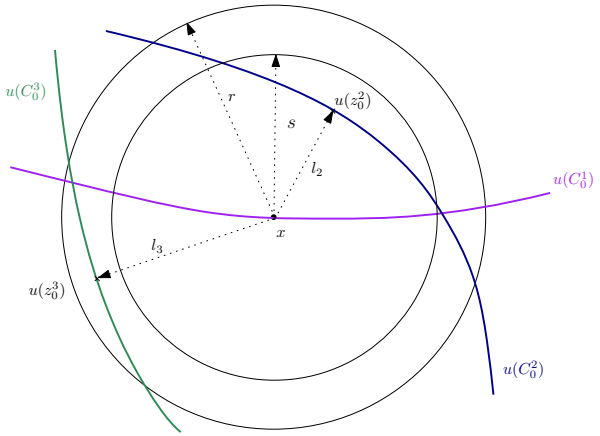


Figure IV.24: Illustration of the cases $l_i \geq s$ and $l_i < s$.

First, assume that $l_i < s$. Let γ be a geodesic starting from z_0^i, denote $v = \dot{\gamma}(0)$ and consider the application $\phi\colon t \mapsto \|\gamma(t) - x\|^2$. Let $a(v), b(v)$ be the first values of $t \geq 0$ such that $\|\gamma(t) - x\| = s$ and $\|\gamma(t) - x\| = r$. As in the proof of Proposition IV.16 Point 3, we still have Equation (IV.7):

$$r^2 - s^2 \geq (1 - \rho r)(b(v)^2 - a(v)^2),$$

from which we deduce $b(v) - a(v) \leq \frac{1}{1-\rho r} \frac{1}{b(v)+a(v)} (r^2 - s^2)$. According to Lemma IV.4 Point 4, $b(v) + a(v) \geq b(v) \geq (1+\rho r)^{-\frac{1}{2}} \sqrt{r^2 - l_i^2} \geq (1+\rho r)^{-\frac{1}{2}} \sqrt{r^2 - s^2}$, and we obtain

$$b(v) - a(v) \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1 - \rho r} \sqrt{r^2 - s^2}. \tag{IV.35}$$

Now, we write

$$\mu_0(D_0^i) = \nu_0^i\big((\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)\big).$$

In spherical coordinates, this measure reads

$$\int_{(\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)} g_0^i(y) \mathrm{d}\mathcal{H}^d(y) = \int_{v \in \partial\mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g_0^i(tv) t^{d-1} \mathrm{d}t \mathrm{d}v.$$

We can now conclude as in the proof of Proposition IV.16 Point 3. We still have $b(v) \leq 2r$, and we write

$$\int_{t=a(v)}^{b(v)} g_0^i(tv)t^{d-1}\mathrm{d}t \leq \int_{t=a(v)}^{b(v)} f_{\max}J_{\max}(2r)^{d-1}\mathrm{d}t.$$

Using Equation (IV.35), we obtain

$$\int_{t=a(v)}^{b(v)} f_{\max}J_{\max}(2r)^{d-1}\mathrm{d}t \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r}\sqrt{r^2-s^2}f_{\max}J_{\max}(2r)^{d-1}.$$

Therefore,

$$\int_{v\in\partial\mathcal{B}(0,1)}\int_{t=a(v)}^{b(v)} g_0^i(tv)t^{d-1}\mathrm{d}t\mathrm{d}v \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r}\sqrt{r^2-s^2}f_{\max}J_{\max}(2r)^{d-1}dV_d.$$

The assumption $r < \frac{1}{2\rho}$ yields $\frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r} < \sqrt{6}$, and we finally obtain

$$\mu_0(D_0^i) \leq f_{\max}J_{\max}2^{d-1}\sqrt{6}dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.$$

Now, assume that $l_i \geq s$. This case is similar to the first one. One has

$$\mu_0(D_0^i) \leq \int_{(\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)} g_0^i(y)\mathrm{d}\mathcal{H}^d(y) = \int_{v\in\partial\mathcal{B}(0,1)}\int_{t=0}^{b(v)} g_0(tv)t^{d-1}\mathrm{d}t\mathrm{d}v.$$

and Lemma IV.4 Point 5 gives $b(v) \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}}\sqrt{r^2-l^2} \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}}\sqrt{r^2-s^2}$. Note that $(\frac{1-\rho r}{2})^{-\frac{1}{2}}$ is not greater than 2 when $r < \frac{1}{2\rho}$. One deduces that

$$\mu_0(D_0^i) \leq f_{\max}J_{\max}2^{d-1}2dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.$$

Step 3: We conclude: since $u^{-1}(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) = \bigcup_i D_0^i$, Step 1 and 2 yield

$$\begin{aligned}\mu(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) = \sum_{i\in I}\mu_0(D_i) &\leq |I|f_{\max}J_{\max}2^{d-1}\sqrt{6}dV_d \cdot r^{d-1}\sqrt{r^2-s^2} \\ &\leq \frac{1}{f_{\min}J_{\min}V_d}\left(\frac{2\rho}{\alpha}\right)^d f_{\max}J_{\max}2^{d-1}\sqrt{6}dV_d \cdot r^{d-1}\sqrt{r^2-s^2}.\end{aligned}$$

Finally, the inequality $\sqrt{r^2-s^2} \leq \sqrt{2r}\sqrt{r-s}$ yields

$$\mu(\overline{\mathcal{B}}(x,r) \setminus \overline{\mathcal{B}}(x,s)) \leq \frac{f_{\max}J_{\max}}{f_{\min}J_{\min}}\left(\frac{\rho}{\alpha}\right)^d d2^{2d}\sqrt{3}r^{d-\frac{1}{2}}\sqrt{r-s}.$$

IV.B Supplementary material for Section IV.3

In this subsection, we suppose that μ and ν are probability measures on E.

Lemma IV.44. *For every $x, y \in E$, we have $\|x^{\otimes 2} - y^{\otimes 2}\|_{\mathrm{F}} \leq (\|x\|+\|y\|)\,\|x-y\|$.*

Proof. We apply the triangular inequality to $x^t x - y^t y = (x-y)^t x + y^t (x-y)$:

$$\begin{aligned}\left\|x^t x - y^t y\right\|_{\mathrm{F}} \leq \left\|(x-y)^t x\right\|_{\mathrm{F}} + \left\|y^t (x-y)\right\|_{\mathrm{F}} &\leq \|x-y\|\,\|x\| + \|y\|\,\|x-y\| \\ &= (\|x\| + \|y\|)\,\|x-y\|\,.\end{aligned}$$

Lemma IV.45. *Let μ' be a submeasure of μ with $|\mu'| > 0$, and consider the corresponding probability measure $\overline{\mu'}$. Suppose that* $\mathrm{supp}(\mu)$ *is included in a ball $\overline{\mathcal{B}}(x, r)$. Then*

$$\mathrm{W}_p\left(\mu, \overline{\mu'}\right) \leq 2(1 - |\mu'|)^{\frac{1}{p}} r.$$

More generally, let μ be any measure of positive mass (potentially with $|\mu| \neq 1$), and let μ' be a submeasure of μ with $|\mu'| > 0$. Suppose that $\mathrm{supp}(\mu)$ *is included in a ball $\overline{\mathcal{B}}(x, r)$. Then*

$$\mathrm{W}_p\left(\overline{\mu}, \overline{\mu'}\right) \leq 2\left(1 - \frac{|\mu'|}{|\mu|}\right)^{\frac{1}{p}} r.$$

Proof. We start with the first inequality. Consider the intermediate probability measure $\omega = \mu' + (1-|\mu'|)\delta_x$. We shall use the triangular inequality $\mathrm{W}_p(\mu, \overline{\mu'}) \leq \mathrm{W}_p(\mu, \omega) + \mathrm{W}_p(\omega, \overline{\mu'})$. We can write

- $\mu = \mu' + (\mu - \mu')$,
- $\omega = \mu' + (1 - |\mu'|)\delta_x$,
- $\overline{\mu'} = \mu' + (\overline{\mu'} - \mu')$.

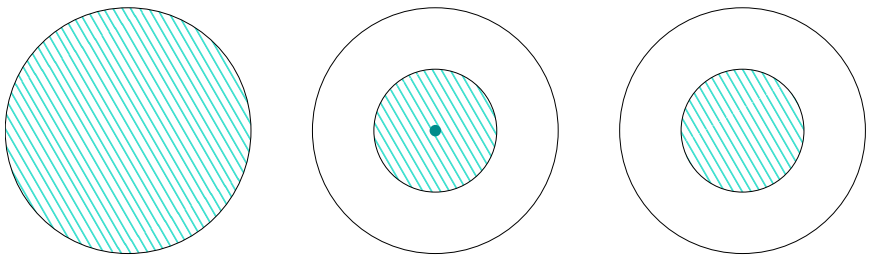


Figure IV.25: The measures involved in the proof of Lemma IV.45. A hatched area represents the support of the measure, and a point represents a Dirac mass.

Observe that μ and ω admits μ' as a common submeasure of mass $|\mu'|$. Therefore we can build a transport plan between μ and ω where only a mass

$1 - |\mu'|$ of μ is moved to x. In other words,

$$\mathrm{W}_p(\mu, \omega) \leq (1 - |\mu'|)^{\frac{1}{p}} r.$$

Similarly, one shows that $\mathrm{W}_p\left(\omega, \overline{\mu'}\right) \leq (1 - |\mu'|)^{\frac{1}{p}} r$.

Now let us prove the second inequality. Since μ' is a submeasure of μ of mass $|\mu'|$, then $\frac{1}{|\mu|}\mu'$ is a submeasure of $\overline{\mu} = \frac{1}{|\mu|}\mu$ of mass $\frac{1}{|\mu|}|\mu'|$. We then apply the first inequality.

Lemma IV.46. *Let $x \in \mathrm{supp}(\mu)$. Suppose that x satisfies Hypotheses 5 and 6 with $\lambda(x) \wedge \frac{1}{2\rho} > r$. Let $y \in E$ such that $\|x - y\| < \frac{r}{4}$. Then $|\mu_x|, |\mu_y| > 0$, and*

$$\mathrm{W}_1\left(\overline{\mu_x}, \overline{\mu_y}\right) \leq c_{25} \|x - y\|$$

with $c_{25} = 2\left(1 + 4\frac{5^{d-1}}{3^d}\right)\frac{c_{10}}{c_9}$.

Proof. It is clear that $|\mu_y| > 0$ since $\mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|))$ and $x \in \mathrm{supp}(\mu)$. Let us show the inequality $\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{25} \|x - y\|$ by studying the measure μ on the intersection $\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)$. Let $\mu_{x,y}$ be the restriction of μ to $\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)$, and $\overline{\mu_{x,y}}$ the corresponding probability measure. The triangular inequality gives:

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) \leq \underbrace{\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}})}_{(1)} + \underbrace{\mathrm{W}_1(\overline{\mu_{x,y}}, \overline{\mu_y})}_{(2)}.$$

Term (1): Let us show that $\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2\frac{c_{10}}{c_9} \|x - y\|$. Note that $\overline{\mu_{x,y}}$ is a submeasure of $\overline{\mu_x}$. According to Lemma IV.45, we have

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2\left(1 - \frac{|\mu_{x,y}|}{|\mu_x|}\right) r = 2\frac{|\mu_x| - |\mu_{x,y}|}{|\mu_x|} r.$$

We know from Hypothesis 5 that $|\mu_x| \geq c_9 r^d$. On the other hand,

$$\begin{aligned} |\mu_x| - |\mu_{x,y}| &= \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)) \\ &\leq \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)), \end{aligned}$$

hence we can apply Hypothesis 6 to get $|\mu_x| - |\mu_{x,y}| \leq c_{10} r^{d-1} \|x - y\|$. We finally obtain

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2\frac{c_{10} r^{d-1} \|x - y\|}{c_9 r^d} r = 2\frac{c_{10}}{c_9} \|x - y\|.$$

Term (2): Similarly, Lemma IV.45 yields

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{|\mu_y| - |\mu_{x,y}|}{|\mu_y|} r.$$

Let us show that we still have $|\mu_y| \geq a' r^d$ and $|\mu_y| - |\mu_{x,y}| \leq b' r^{d-1} \|x - y\|$ with the constants $a' = (\frac{3}{4})^d c_9$ and $b' = 2(\frac{5}{4})^{d-1} c_{10}$. The first inequality comes from Hypothesis 5:

$$\mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \geq c_9 (r - \|x - y\|)^d$$

and $\|x - y\| \leq \frac{r}{4}$. The second inequality comes from Hypothesis 6:

$$\begin{aligned}\mu(\overline{\mathcal{B}}(y, r)) - \mu(\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)) &\leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\leq c_{10} (r + \|x - y\|)^{d-1} 2 \|x - y\|\end{aligned}$$

and $\|x - y\| \leq \frac{r}{4}$. To conclude,

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{2(\frac{5}{4})^{d-1} r^{d-1} c_9 \|x - y\|}{2(\frac{3}{4})^d c_{10} r^d} r = 8\frac{5^{d-1}}{3^d}\frac{c_{10}}{c_9} \|x - y\|.$$

Lemma IV.47. *Let $x \in \mathrm{supp}(\mu)$. Suppose that x satisfies Hypotheses 5 and 7 at x with $\frac{1}{2\rho} > r$. Let $y \in E$ such that $\|x - y\| < \frac{r}{4}$. Then $|\mu_x|, |\mu_y| > 0$, and*

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{26} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}}$$

with $c_{26} = \left(2 + \frac{2^{\frac{5}{2}} 5^{d-\frac{1}{2}}}{3^d}\right) \frac{c_{11}}{c_9}$.

Proof. The proof is similar to Lemma IV.46 with slight modifications. We still consider

$$\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_y}) \leq \underbrace{\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}})}_{(1)} + \underbrace{\mathrm{W}_1(\overline{\mu_{x,y}}, \overline{\mu_y})}_{(2)}.$$

Term (1): We have $\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2\frac{|\mu_x| - |\mu_{x,y}|}{|\mu_x|} r$. Hypothesis 5 still gives $|\mu_x| \geq c_9 r^d$. But Hypothesis 7 now yields

$$\begin{aligned}|\mu_x| - |\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\leq c_{11} r^{d-\frac{1}{2}} \|x - y\|^{\frac{1}{2}}.\end{aligned}$$

We finally obtain $\mathrm{W}_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2\frac{c_{11}}{c_9} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}}$.

Term (2): In order to bound $\mathrm{W}_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{|\mu_y| - |\mu_{x,y}|}{|\mu_y|} r$, Hypothesis 5 still

gives $|\mu_x| \geq (\frac{3}{4})^d c_9 r^d$, and Hypothesis 7 yields

$$\begin{aligned}|\mu_y| - |\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\ &\leq c_{11}(r + \|x-y\|)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}},\end{aligned}$$

which is not greater than $c_{11}(\frac{5}{4}r)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}}$.

We finally get $\mathrm{W}_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2\frac{c_{11}(\frac{5}{4}r)^{d-\frac{1}{2}}(2\|x-y\|)^{\frac{1}{2}}}{(\frac{3}{4})^d c_9 r^d} r \leq \frac{2^{\frac{5}{2}}5^{d-\frac{1}{2}}c_{11}}{3^d c_9} r^{\frac{1}{2}} \|x-y\|^{\frac{1}{2}}$.

Lemma IV.48. *Let $w = \mathrm{W}_p(\mu, \nu)$. Let $y \in E$. Suppose that there exists $x \in \mathrm{supp}(\mu)$ such that $\|x-y\| \leq \alpha$ with $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$, and that μ satisfies Hypotheses 5 and 6 at x with $\lambda(x) \wedge \frac{1}{2\rho} > r$. Assume that $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{27}\alpha$$

with $c_{27} = \frac{2^{d-1}}{c_9} + 2\frac{12 \cdot 5^{d-1}c_{10}+1}{3^d c_9} + 2^{d+3}\frac{(\frac{3}{2})^{d-1}c_{10}+1}{c_9}$.

Proof. Let π be an optimal transport for $\mathrm{W}_p(\mu, \nu)$. Define π_y to be the restriction of the measure π to the set $\overline{\mathcal{B}}(y, r) \times \overline{\mathcal{B}}(y, r) \subset E \times E$. Its marginals $p_{1*}\pi_y$ and $p_{2*}\pi_y$ are submeasures of μ_y and ν_y. We shall use the triangular inequality:

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{\mathrm{W}_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

Before examinating each of these terms, note that we have

$$|\pi_y| = |p_{1*}\pi_y| = |p_{2*}\pi_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \tag{IV.36}$$

$$|\nu_y| \leq \mu(\overline{\mathcal{B}}(y, r+\alpha)) + \frac{w}{\alpha} \tag{IV.37}$$

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \tag{IV.38}$$

The first equation can be proven as follows:

$$\begin{aligned}\mu(\overline{\mathcal{B}}(y, r-\alpha)) &= \pi(\overline{\mathcal{B}}(y, r-\alpha) \times E) \\ &= \pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y, r)) + \pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y, r)^c)\end{aligned}$$

On the one hand, $\pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y, r)) \leq \pi(\overline{\mathcal{B}}(y, r) \times \overline{\mathcal{B}}(y, r)) \leq |\pi_y|$. On the other hand, Markov inequality yields

$$\pi(\overline{\mathcal{B}}(y, r-\alpha) \times \overline{\mathcal{B}}(y, r)^c) \leq \pi(\{(z, z'), \|z-z'\| \geq \alpha\}) \leq \frac{1}{\alpha}\int \|z-z'\| \, \mathrm{d}\pi(z, z'),$$

and Jensen inequality gives

$$\frac{1}{\alpha}\int \|z-z'\|\,\mathrm{d}\pi(z,z') \le \frac{1}{\alpha}\bigg(\int \|z-z'\|^p\,\mathrm{d}\pi(z,z')\bigg)^{\frac{1}{p}} = \frac{w}{\alpha}.$$

We deduce that $\mu(\overline{\mathcal{B}}(y,r-\alpha)) \le |\pi_y| + \frac{w}{\alpha}$, which gives Equation (IV.36). Equations (IV.37) and (IV.38) can be proven similarly.

In addition, note that the assumption $w \le (c_9 \wedge 1)(\frac{r}{4})^{d+1}$ yields

$$\alpha \le \frac{r}{4} \tag{IV.39}$$

$$\frac{w}{\alpha} \le \frac{c_9}{2}\left(\frac{r}{2}\right)^d \tag{IV.40}$$

We now study the terms (2), (1) and (3).

Term (2): Since $\overline{\pi_y} = \frac{\pi_y}{|\pi_y|}$ is a transport plan between $\overline{p_{1*}\pi_y}$ and $\overline{p_{2*}\pi_y}$, we have

$$\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \le \int \|z-z'\|\,\frac{\mathrm{d}\pi_y(z,z')}{|\pi_y|} \le \frac{1}{|\pi_y|}\int \|z-z'\|\,\mathrm{d}\pi(z,z').$$

Moreover, Jensen inequality yields $\int \|z-z'\|\,\mathrm{d}\pi(z,z') \le w$. Hence

$$\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \le \frac{w}{|\pi_y|}.$$

Let us prove that $|\pi_y| \ge \frac{c_9}{2}(\frac{r}{2})^d$. According to Equation (IV.36), $|\pi_y| \ge \mu(\overline{\mathcal{B}}(y,r-\alpha)) - \frac{w}{\alpha}$. Now, remark that $\mu(\overline{\mathcal{B}}(y,r-\alpha)) \ge \frac{c_9}{2^d} r^d$. Indeed, using Hypothesis 5,

$$\mu(\overline{\mathcal{B}}(y,r-\alpha)) \ge \mu(\overline{\mathcal{B}}(x,r-\alpha-\|x-y\|)) \ge c_9(r-\alpha-\|x-y\|)^d,$$

and we conclude with $\|x-y\| \le \alpha \le \frac{r}{4}$. Now, using Equation (IV.40), we get

$$\begin{aligned}|\pi_y| &\ge \mu(\overline{\mathcal{B}}(y,r-\alpha)) - \frac{w}{\alpha}\\ &\ge c_9\left(\frac{r}{2}\right)^d - \frac{c_9}{2}\left(\frac{r}{2}\right)^d \ge \frac{c_9}{2}\left(\frac{r}{2}\right)^d.\end{aligned}$$

Finally, since $\alpha = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}$ and $\alpha \le \frac{r}{4}$, we obtain

$$\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \le \frac{w}{|\pi_y|} \le \frac{w}{\frac{c_9}{2}(\frac{r}{2})^d} = \frac{2^{d+1}}{c_9}\alpha^2\frac{1}{r} \le \frac{2^{d-1}}{c_9}\alpha.$$

Term (1): According to Lemma IV.45, we have

$$\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \le 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|}r.$$

We can use Equation (IV.36) to get

$$\begin{aligned}|\mu_y| - |p_{1*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y,r)) - \mu(\overline{\mathcal{B}}(y,r-\alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x,r+\|x-y\|)) - \mu(\overline{\mathcal{B}}(x,r-\alpha-\|x-y\|)) + \frac{w}{\alpha}.\end{aligned}$$

By Hypothesis 6,

$$\mu(\overline{\mathcal{B}}(x,r+\|x-y\|))-\mu(\overline{\mathcal{B}}(x,r-\alpha-\|x-y\|)) \leq c_{10}(r+\|x-y\|)^{d-1}(2\|x-y\|+\alpha),$$

which is not greater than $c_{10}(\frac{5}{4}r)^{d-1}3\alpha$ since $\|x-y\| \leq \alpha \leq \frac{r}{4}$. Moreover, $\frac{w}{\alpha} = r^{d-1}\alpha$, and we obtain

$$|\mu_y| - |p_{1*}\pi_y| \leq \left(3\left(\frac{5}{4}\right)^{d-1} c_{10} + 1\right) r^{d-1}\alpha,$$

Finally, thanks to Hypothesis 5, we write

$$\begin{aligned}|\mu_y| = \mu(\overline{\mathcal{B}}(y,r)) &\geq \mu(\overline{\mathcal{B}}(x,r-\|x-y\|)) \\ &\geq c_9(r-\|x-y\|)^d \geq c_9\left(\frac{3}{4}\right)^d r^d\end{aligned}$$

and we obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((3(\frac{5}{4})^{d-1}c_{10}+1)r^{d-1}}{c_9(\frac{3}{4})^d r^d}\alpha = \frac{1}{r}\cdot\frac{12\cdot 5^{d-1}c_{10}+1}{3^d c_9}\alpha.$$

We deduce

$$\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{12\cdot 5^{d-1}c_{10}+1}{3^d c_9}\alpha.$$

Term (3): It is similar to Term (1). First, one shows that

$$\mathrm{W}_1(\overline{\nu_y}, \overline{p_{2*}\pi_y}) \leq 2\frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|}r.$$

Using Equations (IV.36) and (IV.37) we get

$$\begin{aligned}|\nu_y| - |p_{2*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y,r+\alpha)) + \frac{w}{\alpha} - \mu(\overline{\mathcal{B}}(y,r-\alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x,r+\|x-y\|+\alpha)) - \mu(\overline{\mathcal{B}}(x,r-\alpha-\|x-y\|)) + 2\frac{w}{\alpha}.\end{aligned}$$

By Hypothesis 6, we have

$$\begin{aligned}&\mu(\overline{\mathcal{B}}(x,r+\|x-y\|+\alpha)) - \mu(\overline{\mathcal{B}}(x,r-\alpha-\|x-y\|)) \\ &\leq c_{10}(r+\|x-y\|+\alpha)^{d-1}(2\|x-y\|+2\alpha)\end{aligned}$$

which is not greater than $c_{10}(\frac{3}{2}r)^{d-1}4\alpha$ since $\|x-y\| \leq \alpha \leq \frac{r}{4}$. Moreover, $\frac{w}{\alpha} = r^{d-1}\alpha$, and we obtain

$$|\nu_y| - |p_{2*}\pi_y| \leq (4(\frac{3}{2})^{d-1}c_{10} + 2)r^{d-1}\alpha.$$

We have seen that

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r-\alpha)) - \frac{w}{\alpha} \geq \frac{c_9}{2}\left(\frac{r}{2}\right)^d.$$

Hence

$$\frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|} \leq \frac{(4(\frac{3}{2})^{d-1}c_{10} + 2)r^{d-1}}{\frac{c_9}{2}(\frac{r}{2})^d}\alpha = \frac{1}{r}\cdot 2^{d+2}\frac{(\frac{3}{2})^{d-1}b + 1}{c_9}\alpha,$$

and we finally obtain

$$\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2^{d+3}\frac{(\frac{3}{2})^{d-1}c_{10} + 1}{c_9}\alpha.$$

To conclude, summing up these three terms gives $\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{27}\alpha$ with

$$c_{27} = \frac{2^{d-1}}{c_9} + 2\frac{12\cdot 5^{d-1}c_{10} + 1}{3^d c_9} + 2^{d+3}\frac{(\frac{3}{2})^{d-1}c_{10} + 1}{c_9}.$$

Lemma IV.49. *Let $w = \mathrm{W}_p(\mu,\nu)$. Let $y \in E$. Suppose that there exists $x \in \mathrm{supp}(\mu)$ such that $\|x-y\| \leq \alpha$ with $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$, and that μ satisfies Hypotheses 5 and 7 at x with $\frac{1}{2\rho} > r$. Assume that $w \leq (c_9 \wedge 1)(\frac{r}{4})^{d+1}$. Then*

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{28} r^{\frac{1}{2}}\alpha^{\frac{1}{2}}$$

with $c_{28} = \frac{2^{d-2}}{c_9} + \frac{4\cdot 3^{\frac{1}{2}}5^{d-\frac{1}{2}}c_{11} + 4^{d-\frac{1}{2}}}{3^d c_9} + 2\cdot 4^d\frac{2c_{11}(\frac{3}{2})^{d-\frac{1}{2}} + 1}{3^d c_9}$.

Proof. The proof is similar as Lemma IV.48. Let us highlight the modifications. Since $\alpha \leq \frac{r}{4}$ and $\frac{w}{\alpha} = r^{d-1}\alpha$, we have the inequalities

$$\alpha^{\frac{1}{2}} \leq \frac{1}{2}r^{\frac{1}{2}}$$

$$\frac{w}{\alpha} \leq \frac{1}{2}r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}$$

We still write the triangular inequality:

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{\mathrm{W}_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

where π is an optimal transport plan for $\mathrm{W}_p(\mu,\nu)$.

Term (2): The argument to obtain $\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-1}}{c_9}\alpha$ is unchanged, and we use $\alpha^{\frac{1}{2}} \leq \frac{1}{2}r^{\frac{1}{2}}$ to get

$$\mathrm{W}_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-2}}{c_9}\alpha^{\frac{1}{2}}r^{\frac{1}{2}}.$$

Term (1): Using Hypothesis 7, we have

$$\begin{aligned}
&\mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r-\alpha-\|x-y\|)) \\
&\leq c_{11}(r + \|x-y\|)^{d-\frac{1}{2}}(2\|x-y\| + \alpha))^{\frac{1}{2}} \\
&\leq c_{11}\left(\frac{5}{4}r\right)^{d-\frac{1}{2}} 3^{\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

And since $\frac{w}{\alpha} \leq \frac{1}{2}r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}$, we get

$$\begin{aligned}
|\mu_y| - |p_{1*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\|)) - \mu(\overline{\mathcal{B}}(x, r-\alpha-\|x-y\|)) + \frac{w}{\alpha} \\
&\leq \left(c_{11}\left(\frac{5}{4}\right)^{d-\frac{1}{2}} 3^{\frac{1}{2}} + \frac{1}{2}\right) r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}.
\end{aligned}$$

Finally, we use

$$\begin{aligned}
|\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) &\geq \mu(\overline{\mathcal{B}}(x, r - \|x-y\|)) \\
&\geq c_9(r - \|x-y\|)^d \geq c_9\left(\frac{3}{4}\right)^d r^d
\end{aligned}$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((c_{11}(\frac{5}{4})^{d-\frac{1}{2}}3^{\frac{1}{2}} + \frac{1}{2})r^{d-\frac{1}{2}}}{c_9(\frac{3}{4})^d r^d}\alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}} \cdot \frac{2 \cdot 3^{\frac{1}{2}}5^{d-\frac{1}{2}}c_{11} + 4^{d-\frac{1}{2}}}{3^d c_9}\alpha^{\frac{1}{2}}$$

and we deduce

$$\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|}r \leq \frac{4 \cdot 3^{\frac{1}{2}}5^{d-\frac{1}{2}}c_{11} + 4^{d-\frac{1}{2}}}{3^d c_9}r^{\frac{1}{2}}\alpha^{\frac{1}{2}}.$$

Term (3): We use Hypothesis 7 to get

$$\begin{aligned}
&\mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r-\alpha-\|x-y\|)) \\
&\leq c_{11}(r + \|x-y\| + \alpha)^{d-\frac{1}{2}}(2\|x-y\| + 2\alpha)^{\frac{1}{2}} \\
&\leq 2c_{11}\left(\frac{3}{2}r\right)^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}.
\end{aligned}$$

And since $\frac{w}{\alpha} \leq \frac{1}{2} r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}$, we get

$$|\nu_y| - |p_{2*}\pi_y| \leq \mu(\overline{\mathcal{B}}(x, r + \|x-y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x-y\|)) + 2\frac{w}{\alpha}$$
$$\leq \left(2c_{11}\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1\right) r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}.$$

Finally, we use

$$|\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x-y\|))$$
$$\geq c_9 (r - \|x-y\|)^d \geq c_9 \left(\frac{3}{4}\right)^d r^d$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{(2c_{11}(\frac{3}{2})^{d-\frac{1}{2}} + 1) r^{d-\frac{1}{2}}}{c_9 (\frac{3}{4})^d r^d} \alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}} \cdot 4^d \frac{2c_{11}(\frac{3}{2})^{d-\frac{1}{2}} + 1}{3^d c_9} \alpha^{\frac{1}{2}}$$

and we deduce

$$\mathrm{W}_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r \leq 2 \cdot 4^d \frac{2c_{11}(\frac{3}{2})^{d-\frac{1}{2}} + 1}{3^d c_9} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}.$$

Remark IV.50. Let us comment the inequality of Lemma IV.48 with $p = 1$, valid for all r such that $w \leq (a \wedge 1)(\frac{r}{4})^{d+1}$:

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{27} \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}.$$

If r is assumed to be constant, the behavior of $\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y})$, when w goes to 0, is

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \lesssim w^{\frac{1}{2}}.$$

On the other hand, if r is supposed to follow the worst case, i.e. r is of order $w^{\frac{1}{d+1}}$, then $\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y})$ is of order

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \lesssim \left(\frac{w}{w^{\frac{d-1}{d+1}}}\right)^{\frac{1}{2}} = w^{\frac{1}{d+1}}.$$

Now, let us show that the order $(\frac{w}{r^{d-1}})^{\frac{1}{2}}$ is optimal. More precisely, we show that, for every $d \geq 1$, $r > 0$ and $\epsilon > 0$ fixed, there exists measures μ and ν on $\mathbb{R}^d$ that satisfies the assumptions of Lemma IV.48, but such that

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \geq c_d \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}} - \epsilon$$

with $c_d = \frac{1}{d+1}\left(\frac{2d}{V_d}\right)^{\frac{1}{2}}$. We consider the following example. Let $\mu = \mathcal{H}^d_{[0,1]^d}$ be the Lebesgue measure on the hypercube $[0,1]^d$. Denote $y = \left(\frac{1}{2}, ..., \frac{1}{2}\right)$ its center, $B = \mathcal{B}(y, r)$ the open ball, and A the annulus defined as

$$A = \mathcal{B}(y, r + \epsilon) \setminus \mathcal{B}(y, r)$$

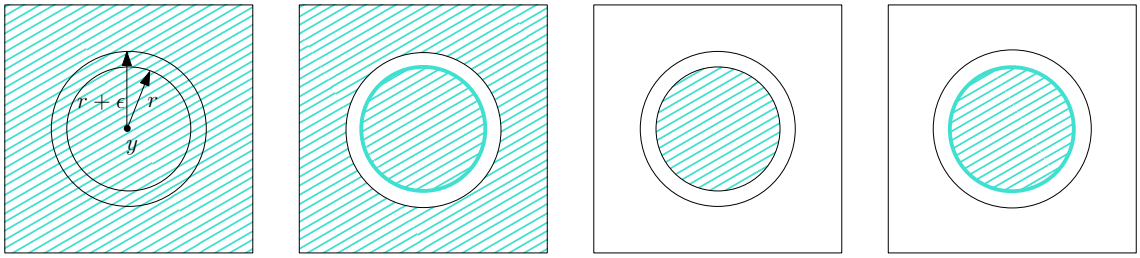
where $0 < \epsilon < r < \frac{1}{4}$. In the following, r stays fixed, and ϵ will go to zero. Consider the probability measure

$$\nu = \mathcal{H}^d_{[0,1]^d \setminus A} + \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1} r^{d-1}} \mathcal{H}^{d-1}_{\partial \mathcal{B}(y,r)}.$$

Let $\overline{\mu_y}$ and $\overline{\nu_y}$ be the localized probability measures associated to μ and ν with parameter r. We shall show that

$$\mathrm{W}_1(\mu, \nu) \text{ is of order } r^{d-1}\epsilon^2 \qquad \text{and} \qquad \mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) \text{ is of order } \epsilon$$

when $\epsilon \to 0$.



μ ν $\overline{\mu_y}$ $\overline{\nu_y}$

Figure IV.26: The measures involved in the example. A hatched area represents the d-dimensional Hausdorff measure $\mathcal{H}^d$, and a bold circle represents the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$.

Step 1: Study of $\mathrm{W}_1(\mu, \nu)$. An optimal transport plan between μ and ν is given by transporting the submeasure $\mathcal{H}^d_A$ of μ onto the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1} r^{d-1}} \mathcal{H}^{d-1}_{\partial \mathcal{B}(y,r)}$ of ν via the application

$$\begin{aligned} A &\longrightarrow \partial \mathcal{B}(y,r) \\ x &\longmapsto \frac{r}{\|x\|} x. \end{aligned}$$

Consequently, the Wasserstein distance is

$$\mathrm{W}_1(\mu, \nu) = \int_A \left\| x - \frac{r}{\|x\|} x \right\| \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1} r^{d-1}} \mathrm{d}\mathcal{H}^d(x).$$

A change of coordinates shows that

$$\int_A \left\| x - \frac{r}{\|x\|} x \right\| \mathrm{d}\mathcal{H}^d(x) = \int_{\partial \mathcal{B}(0,1)} \int_r^{r+\epsilon} (t-r) t^{d-1} \mathrm{d}\mathcal{H}^1(t) \mathrm{d}\mathcal{H}^{d-1}(v).$$

Let us write $\int_r^{r+\epsilon} (t-r) t^{d-1} \mathrm{d}\mathcal{H}^1(t) = \int_r^{r+\epsilon} t^d \mathrm{d}\mathcal{H}^1(t) - \int_r^{r+\epsilon} r t^{d-1} \mathrm{d}\mathcal{H}^1(t)$. We have

$$\begin{aligned} \int_r^{r+\epsilon} t^d \mathrm{d}\mathcal{H}^1(t) &= \frac{1}{d+1} \left((r+\epsilon)^{d+1} - r^{d+1} \right) \\ &= r^d \epsilon + \frac{d}{2} r^{d-1} \epsilon^2 + o(\epsilon^2), \end{aligned}$$

where the Little-O notation refers to $\epsilon \to 0$. Moreover,

$$\int_r^{r+\epsilon} rt^{d-1}\mathrm{d}\mathcal{H}^1(t) = r\left(r^{d-1}\epsilon + \frac{d-1}{2}r^{d-2}\epsilon^2 + o(\epsilon)^2\right)$$
$$= r^d\epsilon + \frac{d-1}{2}r^{d-1}\epsilon^2 + o(\epsilon^2).$$

We deduce that $\int_r^{r+\epsilon}(t-r)t^{d-1}\mathrm{d}\mathcal{H}^1(t) = \frac{1}{2}r^{d-1}\epsilon^2 + o(\epsilon^2)$, and

$$\int_A \left\| x - \frac{r}{\|x\|}x \right\| \mathrm{d}\mathcal{H}^d(x) = \frac{S_{d-1}}{2}r^{d-1}\epsilon^2 + o(\epsilon^2).$$

In other words,

$$\mathrm{W}_1(\mu,\nu) = \frac{dV_d}{2}r^{d-1}\epsilon^2 + o(\epsilon^2).$$

Step 2: Study of $\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y})$. Consider the measures

$$\overline{\mu_x} = \frac{1}{V_d r^d}\mathcal{H}^d_B = \left(\frac{1}{V_d(r+\epsilon)^d} + \frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d}\right)\mathcal{H}^d_B$$

and

$$\overline{\nu_x} = \frac{1}{V_d(r+\epsilon)^d}\left(\mathcal{H}^d_B + \frac{V_d(r+\epsilon)^d - V_d r^d}{S_{d-1}r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}\right).$$

Consider the Wasserstein distance $\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y})$. As before, an optimal transport plan is given by transporting the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d}\mathcal{H}^d_B$ of $\overline{\mu_x}$ onto the submeasure $\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d S_{d-1}r^{d-1}}\mathcal{H}^{d-1}_{\partial\mathcal{B}(y,r)}$ of $\overline{\nu_x}$. We have:

$$\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y}) = \int_B \left\| x - \frac{r}{\|x\|}x \right\| \frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d}\mathrm{d}\mathcal{H}^d(x)$$

A change of coordinates yields

$$\int_B \left\| x - \frac{r}{\|x\|}x \right\| \mathrm{d}\mathcal{H}^d(x) = \frac{S_{d-1}}{d(d+1)}r^{d+1}.$$

Besides, we have

$$\frac{V_d(r+\epsilon)^d - V_d r^d}{V_d(r+\epsilon)^d V_d r^d} = \frac{dV_d r^{d-1}\epsilon + O(\epsilon^2)}{V_d(r+\epsilon)^d V_d r^d} = \frac{d}{V_d}\frac{\epsilon}{r^{d+1}} + O(\epsilon^2).$$

We deduce that

$$\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y}) = \frac{S_{d-1}}{d(d+1)}\frac{d}{V_d}\epsilon + O(\epsilon^2) = \frac{d}{d+1}\epsilon + O(\epsilon^2).$$

Step 3. Using $\mathrm{W}_1(\mu,\nu) = \frac{dV_d}{2}r^{d-1}\epsilon^2 + o(\epsilon^2)$ and $\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y}) = \frac{d}{d+1}\epsilon + O(\epsilon^2)$, we get

$$\frac{\mathrm{W}_1(\overline{\mu_y},\overline{\nu_y})^2}{\mathrm{W}_1(\mu,\nu)} = c\frac{1}{r^{d-1}} + O(\epsilon)$$

with $c = \frac{\left(\frac{d}{d+1}\right)^2}{\frac{dV_d}{2}} = \frac{2d}{(d+1)^2 V_d}$. In conclusion,

$$\frac{\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y})}{\mathrm{W}_1(\mu, \nu)^{\frac{1}{2}}} = c^{\frac{1}{2}} \left(\frac{1}{r^{d-1}}\right)^{\frac{1}{2}} + O(\epsilon),$$

and since $\mathrm{W}_1(\mu, \nu)^{\frac{1}{2}} = O(\epsilon)$, we deduce

$$\mathrm{W}_1(\overline{\mu_y}, \overline{\nu_y}) = c^{\frac{1}{2}} \left(\frac{\mathrm{W}_1(\mu, \nu)}{r^{d-1}}\right)^{\frac{1}{2}} + O(\epsilon^2).$$

V Persistent Stiefel-Whitney classes

Abstract. We propose a definition of persistent Stiefel-Whitney classes of vector bundle filtrations. It relies on seeing vector bundles as subsets of some Euclidean spaces. The usual Čech filtration of such a subset can be endowed with a vector bundle structure, that we call a Čech bundle filtration. We show that this construction is stable and consistent. When the dataset is a finite sample of a line bundle, we implement an effective algorithm to compute its persistent Stiefel-Whitney classes. In order to use simplicial approximation techniques in practice, we develop a notion of weak simplicial approximation. As a theoretical example, we give an in-depth study of the normal bundle of the circle, which reduces to understanding the persistent cohomology of the torus knot (1,2).

V.1 Persistent Stiefel-Whitney classes . . . 198
V.1.1 Definition . . . 198
V.1.2 Čech bundle filtrations . . . 199
V.1.3 Stability . . . 203
V.1.4 Consistency . . . 206
V.2 Computation of persistent Stiefel-Whitney classes . . . 209
V.2.1 Simplicial approximation to Čech bundle filtrations . . . 209
V.2.2 A sketch of algorithm . . . 211
V.3 An algorithm when $d = 1$ . . . 213
V.3.1 The star condition in practice . . . 213
V.3.2 Triangulating the projective spaces . . . 216
V.3.3 Vietoris-Rips version of the Čech bundle filtration . . . 219
V.3.4 Choice of the parameter γ . . . 219
V.4 Conclusion . . . 222
V.A Supplementary material for Section V.1 . . . 223
V.B Supplementary material for Section V.3 . . . 224
V.B.1 Study of Example V.24 . . . 224
V.B.2 Study of Example V.25 . . . 226

Numerical experiments. A `Python` notebook can be found at `https://github.com/raphaeltinarrage/PersistentCharacteristicClasses/blob/master/Demo.ipynb`.

Publication. A journal-formatted version is available at [Tin20].

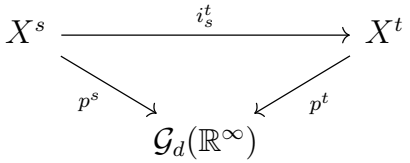
Organisation of the chapter. The definitions of fiber bundle filtrations are given in Section V.1, where their stability and consistency properties are established. In Section V.2, we propose a sketch of algorithm to compute these classes, based on simplicial approximation techniques. In Section V.3 we give a particular attention to some technical details needed to implement this algorithm. For the clarity of the chapter, the proofs of several results have been postponed to the appendices. Please refer to Subsection I.2.4 for an introduction to this chapter.

V.1 Persistent Stiefel-Whitney classes

V.1.1 Definition

Let $E = \mathbb{R}^n$ be a Euclidean space, and $\mathbb{X} = (X^t)_{t\in T}$ a set filtration of E (see Subsection II.4.2 for definitions). Let us denote by i_s^t the inclusion map from X^s to X^t. In order to define persistent Stiefel-Whitney classes, we have to give such a filtration a vector bundle structure. Basic notions of vector bundles and Stiefel-Whitney classes are recalled in Subsections II.2.3 and II.2.4. The infinite Grassmann manifold is denoted $\mathcal{G}_d(\mathbb{R}^\infty)$.

Definition V.1 (Vector bundle filtrations). A vector bundle filtration of dimension d on E is a couple $(\mathbb{X}, \mathbb{p})$ where $\mathbb{X} = (X^t)_{t\in T}$ is a set filtration of E and $\mathbb{p} = (p^t)_{t\in T}$ a family of continuous maps $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^\infty)$ such that, for every $s, t \in T$ with $s \leq t$, we have $p^t \circ i_s^t = p^s$. In other words, the following diagram commutes:



Note that for any $m \in \mathbb{N}$, and by using the inclusion $\mathcal{G}_d(\mathbb{R}^m) \hookrightarrow \mathcal{G}_d(\mathbb{R}^\infty)$, one may define a vector bundle filtration by considering maps $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$.

Let us fix a $t \in T$. The map $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^\infty)$ gives the topological space X^t a vector bundle structure, as discussed in Subsection II.2.3. Following Subsection II.2.4, the induced map in cohomology, $(p^t)^*$, allows to define the Stiefel-Whitney classes of this vector bundle. Let us introduce some notations. The Stiefel-Whitney classes of $\mathcal{G}_d(\mathbb{R}^\infty)$ are denoted $w_1, ..., w_d$. The Stiefel-Whitney classes of the vector bundle (X^t, p^t) are denoted $w_1^t(\mathbb{p}), ..., w_d^t(\mathbb{p})$, and can be defined as $w_i^t(\mathbb{p}) = (p^t)^*(w_i)$ (as in Theorem II.6).

$$(p^t)^*\colon H^*\big(X^t\big) \longleftarrow H^*\big(\mathcal{G}_d(\mathbb{R}^\infty)\big)$$

$$w_1^t(\mathbb{p}) \longleftarrow\!\shortmid\ w_1$$

$$\vdots$$

$$w_d^t(\mathbb{p}) \longleftarrow\!\shortmid\ w_d$$

Let $(\mathbb{V}, \mathbb{v})$ denote the persistence module associated to the filtration $\mathbb{X}$, with $\mathbb{V} = (V^t)_{t\in T}$ and $\mathbb{v} = (v_s^t)_{s\leq t\in T}$. Explicitly, V^t is the cohomology ring $H^*(X^t)$, and v_s^t is

the induced map $H^*(X^s) \leftarrow H^*(X^t)$. For every $t \in T$, the classes $w_1^t(\mathbb{p}), \cdots, w_d^t(\mathbb{p})$ belong to the vector space V^t. The persistent Stiefel-Whitney classes are defined to be the collection of such classes over t.

Definition V.2 (Persistent Stiefel-Whitney classes). Let $(\mathbb{X}, \mathbb{p})$ be a vector bundle filtration. The persistent Stiefel-Whitney classes of $(\mathbb{X}, \mathbb{p})$ are the families of classes

$$\begin{aligned} w_1(\mathbb{p}) &= \big(w_1^t(\mathbb{p})\big)_{t\in T} \\ &\vdots \\ w_d(\mathbb{p}) &= \big(w_d^t(\mathbb{p})\big)_{t\in T}. \end{aligned}$$

Let $i \in [1, d]$, and consider a persistent Stiefel-Whitney class $w_i(\mathbb{p})$. Note that it satisfies the following property: for all $s, t \in T$ such that $s \leq t$, $w_i^s(\mathbb{p}) = v_s^t\big(w_i^t(\mathbb{p})\big)$. As a consequence, if a class $w_i^t(\mathbb{p})$ is given for a $t \in T$, one obtains all the others $w_i^s(\mathbb{p})$, with $s \leq t$, by applying the maps v_s^t. In particular, if $w_i^t(\mathbb{p}) = 0$, then $w_i^s(\mathbb{p}) = 0$ for all $s \in T$ such that $s \leq t$.

Lifebar. In order to visualize the evolution of a persistent Stiefel-Whitney class through the persistence module $(\mathbb{V}, \mathbb{v})$, we propose the following bar representation: the lifebar of $w_i(\mathbb{p})$ is the set

$$\left\{t \in T, w_i^t(\mathbb{p}) \neq 0\right\}.$$

According to the last paragraph, the lifebar of a persistent class is an interval of T, of the form $[t^\dagger, \sup(T))$ or $(t^\dagger, \sup(T))$, where

$$t^\dagger = \inf\left\{t \in T, w_i^t(\mathbb{p}) \neq 0\right\},$$

with the convention $\inf(\emptyset) = \inf(T)$. In order to distinguish the lifebar of a persistent Stiefel-Whitney class from the bars of the persistence barcodes, we draw the rest of the interval hatched.

Figure V.1: Example of a lifebar of a persistent Stiefel-Whitney class with $t^\dagger = 0.2$ and $\max(T) = 1$.

V.1.2 Čech bundle filtrations

In this subsection, we propose a particular construction of vector bundle filtration, called the Čech bundle filtration. We shall work in the ambient space $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Let $\|\cdot\|$ be the usual Euclidean norm on the space $\mathbb{R}^n$, and $\|\cdot\|_{\mathrm{F}}$ the

Frobenius norm on $\mathrm{M}(\mathbb{R}^m)$, the space of $m \times m$ matrices. Let $\gamma > 0$. We endow the vector space E with the Euclidean norm $\|\cdot\|_\gamma$ defined for every $(x, A) \in E$ as

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_{\mathrm{F}}^2 . \tag{V.1}$$

See Subsection V.3.4 for a discussion about the parameter γ.

In order to define the Čech bundle filtration, we shall first study the usual embedding of the Grassmann manifold $\mathcal{G}_d(\mathbb{R}^m)$ into the matrix space $\mathrm{M}(\mathbb{R}^m)$.

Embedding of $\mathcal{G}_d(\mathbb{R}^m)$. We embed the Grassmannian $\mathcal{G}_d(\mathbb{R}^m)$ into $\mathrm{M}(\mathbb{R}^m)$ via the application which sends a d-dimensional subspace $T \subset \mathbb{R}^m$ to its orthogonal projection matrix P_T. We can now see $\mathcal{G}_d(\mathbb{R}^m)$ as a submanifold of $\mathrm{M}(\mathbb{R}^m)$. Recall that $\mathrm{M}(\mathbb{R}^m)$ is endowed with the Frobenius norm. According to this metric, $\mathcal{G}_d(\mathbb{R}^m)$ is included in the sphere of center 0 and radius $\sqrt{d}$ of $\mathrm{M}(\mathbb{R}^m)$.

In the metric space $(\mathrm{M}(\mathbb{R}^m), \|\cdot\|_{\mathrm{F}})$, consider the distance function to $\mathcal{G}_d(\mathbb{R}^m)$, denoted $\mathrm{dist}\,(\cdot, \mathcal{G}_d(\mathbb{R}^m))$. Let $\mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m))$ denote the medial axis of $\mathcal{G}_d(\mathbb{R}^m)$. It consists in the points $A \in \mathrm{M}(\mathbb{R}^m)$ which admit at least two projections on $\mathcal{G}_d(\mathbb{R}^m)$:

$$\mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m)) = \{A \in \mathrm{M}(\mathbb{R}^m), \exists P, P' \in \mathcal{G}_d(\mathbb{R}^m), P \neq P', \\ \|A - P\|_{\mathrm{F}} = \|A - P\|_{\mathrm{F}} = \mathrm{dist}\,(A, \mathcal{G}_d(\mathbb{R}^m))\}.$$

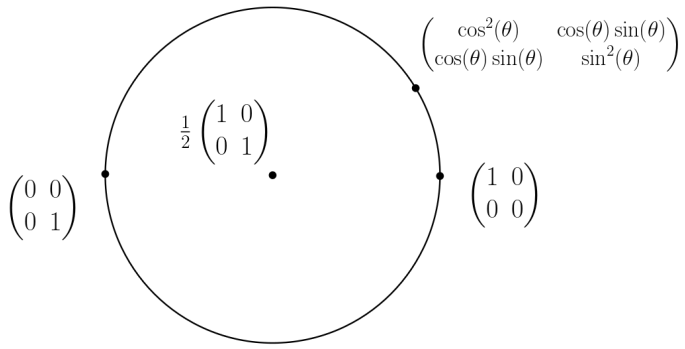


Figure V.2: Representation of the Grassmannian $\mathcal{G}_1(\mathbb{R}^2) \subset \mathrm{M}(\mathbb{R}^2) \simeq \mathbb{R}^4$. It is equal to the circle of radius $\frac{\sqrt{2}}{2}$, in the 2-affine space generated by $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and with origin $\frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The matrix $\frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is an element of $\mathrm{med}\,(\mathcal{G}_1(\mathbb{R}^2))$.

On the set $\mathrm{M}(\mathbb{R}^m) \setminus \mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m))$, the projection on $\mathcal{G}_d(\mathbb{R}^m)$ is well-defined:

$$\mathrm{proj}\,(\cdot, \mathcal{G}_d(\mathbb{R}^m)) : \mathrm{M}(\mathbb{R}^m) \setminus \mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m)) \longrightarrow \mathcal{G}_d(\mathbb{R}^m) \subset \mathrm{M}(\mathbb{R}^m) \\ A \longmapsto P \ \text{ s.t. } \ \|P - A\|_{\mathrm{F}} = \mathrm{dist}\,(A, \mathcal{G}_d(\mathbb{R}^m)) .$$

The following lemma describes this projection explicitly. We defer its proof to Appendix V.A (page 223).

Lemma V.3. *For any $A \in \mathrm{M}(\mathbb{R}^m)$, let A^s denote the matrix $A^s = \frac{1}{2}(A + {}^tA)$, and let $\lambda_1(A^s), ..., \lambda_n(A^s)$ be the eigenvalues of A^s in decreasing order. The distance from A to* $\mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m))$ *is*

$$\mathrm{dist}\,(A, \mathrm{med}\,(\mathcal{G}_d(\mathbb{R}^m))) = \frac{\sqrt{2}}{2}\big|\lambda_d(A^s) - \lambda_{d+1}(A^s)\big|.$$

If this distance is positive, the projection of A on $\mathcal{G}_d(\mathbb{R}^m)$ can be described as follows: consider the symmetric matrix A^s, and let $A^s = OD^tO$, with O an orthogonal matrix, and D the diagonal matrix containing the eigenvalues of A^s in decreasing order. Let J_d be the diagonal matrix whose first d terms are 1, and the other ones are zero. We have

$$\operatorname{proj}(A, \mathcal{G}_d(\mathbb{R}^m)) = OJ_d{}^tO.$$

Observe that, as a consequence of this lemma, every point of $\mathcal{G}_d(\mathbb{R}^m)$ is at equal distance from $\operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))$, and this distance is equal to $\frac{\sqrt{2}}{2}$. Therefore the reach of the subset $\mathcal{G}_d(\mathbb{R}^m) \subset \mathrm{M}(\mathbb{R}^m)$ is

$$\operatorname{reach}(\mathcal{G}_d(\mathbb{R}^m)) = \frac{\sqrt{2}}{2}.$$

Čech bundle filtration. Let X be a subset of $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Consider the usual Čech filtration $\mathbb{X} = (X^t)_{t\geq 0}$, where X^t denotes the t-thickening of $\check{X}$ in the metric space $(E, \|\cdot\|_\gamma)$. In order to give this filtration a vector bundle structure, consider the map p^t defined as the composition

$$X^t \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m) \xrightarrow{\ \mathrm{proj}_2\ } \mathrm{M}(\mathbb{R}^m) \setminus \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m)) \xrightarrow{\ \mathrm{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))\ } \mathcal{G}_d(\mathbb{R}^m), \tag{V.2}$$

where proj_2 represents the projection on the second coordinate of $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$, and $\operatorname{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))$ the projection on $\mathcal{G}_d(\mathbb{R}^m) \subset \mathrm{M}(\mathbb{R}^m)$. Note that p^t is well-defined only when X^t does not intersect $\mathbb{R}^n \times \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))$. The supremum of such t's is denoted $t_\gamma^{\max}(X)$. We have

$$t_\gamma^{\max}(X) = \inf\left\{\operatorname{dist}_\gamma(x, \mathbb{R}^n \times \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))), x \in X\right\}, \tag{V.3}$$

where $\operatorname{dist}_\gamma(x, \mathbb{R}^n \times \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m)))$ is the distance between the point $x \in \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$ and the subspace $\mathbb{R}^n \times \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))$, with respect to the norm $\|\cdot\|_\gamma$. By definition of $\|\cdot\|_\gamma$, Equation (V.3) rewrites as

$$t_\gamma^{\max}(X) = \gamma \cdot \inf\{\operatorname{dist}(A, \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))), (y, A) \in X\},$$

where $\operatorname{dist}(A, \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m)))$ represents the distance between the matrix A and the subspace $\operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))$ with respect to the Frobenius norm $\|\cdot\|_{\mathrm{F}}$. Denoting $t^{\max}(X)$ the value $t_\gamma^{\max}(X)$ for $\gamma = 1$, we obtain

$$\begin{aligned} & t_\gamma^{\max}(X) = \gamma \cdot t^{\max}(X) \\ \text{and} \quad & t^{\max}(X) = \inf\{\operatorname{dist}(A, \operatorname{med}(\mathcal{G}_d(\mathbb{R}^m))), (y, A) \in X\}. \end{aligned} \tag{V.4}$$

Note that the values $t^{\max}(X)$ can be computed explicitly thanks to Lemma V.3. In particular, if X is a subset of $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$, then $t^{\max}(X) = \frac{\sqrt{2}}{2}$. Accordingly,

$$t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2}\gamma. \tag{V.5}$$

Definition V.4 (Čech bundle filtration). Consider a subset X of $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$, and suppose that $t^{\max}(X) > 0$. The Čech bundle filtration associated to

X in the ambient space $(E, \|\cdot\|_\gamma)$ is the vector bundle filtration $(\mathbb{X}, \mathbb{p})$ consisting of the Čech filtration $\mathbb{X} = (X^t)_{t\in T}$, and the maps $\mathbb{p} = (p^t)_{t\in T}$ as defined in Equation (V.2). This vector bundle filtration is defined on the index set $T = \left[0, t_\gamma^{\max}(X)\right)$, where $t_\gamma^{\max}(X)$ is defined in Equation (V.4).

The i^{th} persistent Stiefel-Whitney class of the Čech bundle filtration $(\mathbb{X}, \mathbb{p})$, as in Definition V.2, shall be denoted $w_i(X)$ instead of $w_i(\mathbb{p})$.

Example V.5. Let $E = \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$. Let X and Y be the subsets of E defined as:

$$X = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}$$

$$Y = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}$$

The set X is to be seen as the normal bundle of the circle, and Y as the universal bundle of the circle, known as the Mobius band. We have $t^{\max}(X) = t^{\max}(Y) = \frac{\sqrt{2}}{2}$ as in Lemma V.3. Let $\gamma = 1$.

Figure V.3: Representation of the sets X and $Y \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$: the black points correspond to the $\mathbb{R}^2$-coordinate, and the pink segments over them correspond to the orientation of the $\mathrm{M}(\mathbb{R}^2)$-coordinate.

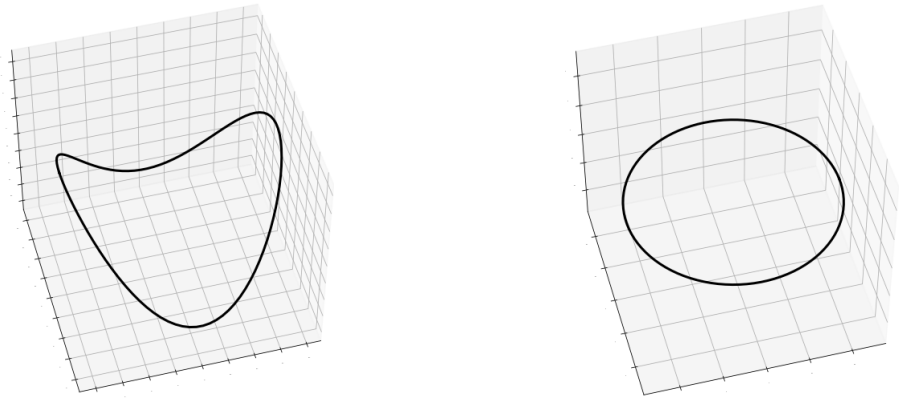


Figure V.4: The sets X and $Y \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, projected in a 3-dimensional subspace of $\mathbb{R}^3$ via PCA.

We now compute the persistence barcodes of the Čech filtrations of X and Y in the ambient space E.

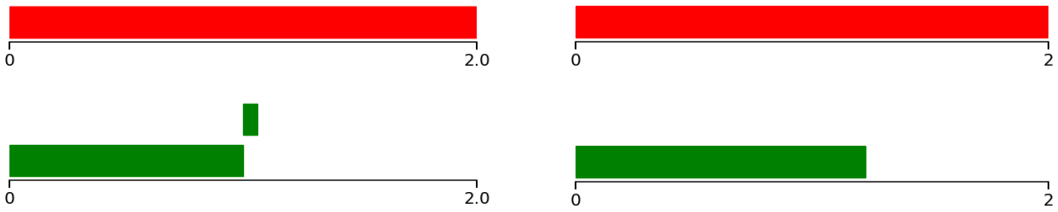


Figure V.5: H^0 and H^1 persistence barcode of the Čech filtration of X **(left)** and Y **(right)**.

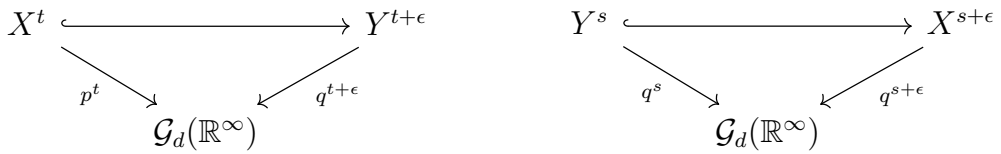
Consider the first persistent Stiefel-Whitney classes $w_1(X)$ and $w_1(Y)$ of the corresponding Čech bundle filtrations. We compute that their lifebars are $\emptyset$ for $w_1(X)$, and $[0, t^{\max}(Y))$ for $w_1(Y)$. This is illustrated in Figure V.6. One reads these bars as follows: $w_1^t(X)$ is zero for every $t \in \left[0, \frac{\sqrt{2}}{2}\right)$, while $w_1^t(Y)$ is nonzero.

Figure V.6: Lifebars of the first persistent Stiefel-Whitney classes $w_1(X)$ and $w_1(Y)$.

V.1.3 Stability

In this subsection we derive a straightforward stability result for persistent Stiefel-Whitney classes. We start by defining a notion of interleavings for vector bundle filtrations, in the same vein as the usual interleavings of set filtrations.

Definition V.6 (Interleavings of vector bundle filtrations). Let $\epsilon \geq 0$, and consider two vector bundle filtrations $(\mathbb{X}, \mathbb{p})$, $(\mathbb{Y}, \mathbb{q})$ of dimension d on E with respective index sets T and U. They are ϵ-interleaved if the underlying filtrations $\mathbb{X} = (X^t)_{t\in T}$ and $\mathbb{Y} = (Y^t)_{t\in U}$ are ϵ-interleaved, and if the following diagrams commute for every $t \in T \cap (U - \epsilon)$ and $s \in U \cap (T - \epsilon)$:



The following theorem shows that interleavings of vector bundle filtrations give rise to interleavings of persistence modules which respect the persistent Stiefel-Whitney classes.

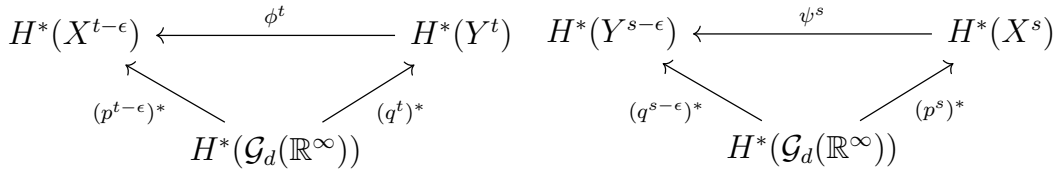
Theorem V.7. *Consider two vector bundle filtrations $(\mathbb{X}, \mathbb{p})$, $(\mathbb{Y}, \mathbb{q})$ of dimension d with respective index sets T and U. Suppose that they are ϵ-interleaved. Then there exists an ϵ-interleaving (ϕ, ψ) between their corresponding persistent*

cohomology modules which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes. In other words, for every $i \in [1,d]$, and for every $t \in (T+\epsilon)\cap U$ and $s \in U \cap (T+\epsilon)$, we have

$$\begin{aligned} & \phi^t(w_i^t(\mathbb{p})) = w_i^{t-\epsilon}(\mathbb{q}) \\ \textit{and} \quad & \psi^s(w_i^s(\mathbb{p})) = w_i^{s-\epsilon}(\mathbb{q}). \end{aligned}$$

Proof. Define (ϕ,ψ) to be the ϵ-interleaving between the cohomology persistence modules $\mathbb{V}(\mathbb{X})$ and $\mathbb{V}(\mathbb{Y})$ given by the ϵ-interleaving between the filtrations $\mathbb{X}$ and $\mathbb{Y}$. Explicitly, if $i_t^{t+\epsilon}$ denotes the inclusion $X^t \hookrightarrow Y^{t+\epsilon}$ and $j_s^{s+\epsilon}$ denotes the inclusion $Y^s \hookrightarrow X^{s+\epsilon}$, then $\phi = (\phi^t)_{t\in(T+\epsilon)\cap U}$ is given by the induced maps in cohomology $\phi^t = (i_{t-\epsilon}^t)^*$, and $\psi = (\psi^s)_{s\in(U+\epsilon)\cap T}$ is given by $\psi^s = (j_{s-\epsilon}^s)^*$.

Now, by fonctoriality, the diagrams of Definition V.6 give rise to commutative diagrams in cohomology:



Let $i \in [1,d]$. By definition, the persistent Stiefel-Whitney classes $w_i(\mathbb{p}) = (w_i^t(\mathbb{p}))_{t\in T}$ and $w_i(\mathbb{q}) = (w_i^s(\mathbb{q}))_{s\in U}$ are equal to $w_i^t(\mathbb{p}) = (p^t)^*(w_i)$ and $w_i^s(\mathbb{q}) = (q^s)^*(w_i)$, where w_i is the i^{th} Stiefel-Whitney class of $\mathcal{G}_d(\mathbb{R}^\infty)$. The previous commutative diagrams then translates as $\phi^t(w_i^t(\mathbb{p})) = w_i^{t-\epsilon}(\mathbb{q})$ and $\psi^s(w_i^s(\mathbb{p})) = w_i^{s-\epsilon}(\mathbb{q})$, as wanted.

Consider two vector bundle filtrations $(\mathbb{X},\mathbb{p})$, $(\mathbb{Y},\mathbb{q})$ such that there exists an ϵ-interleaving (ϕ,ψ) between their persistent cohomology modules $\mathbb{V}(\mathbb{X})$, $\mathbb{V}(\mathbb{Y})$ which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes. Let $i \in [1,d]$. Then the lifebars of their i^{th} persistent Stiefel-Whitney classes $w_i(\mathbb{p})$ and $w_i(\mathbb{q})$ are ϵ-close in the following sense: if we denote $t^\dagger(\mathbb{p}) = \inf\{t \in T, w_i^t(\mathbb{p}) \neq 0\}$ and $t^\dagger(\mathbb{q}) = \inf\{t \in T, w_i^t(\mathbb{q}) \neq 0\}$, then $|t^\dagger(\mathbb{p}) - t^\dagger(\mathbb{q})| \leq \epsilon$.

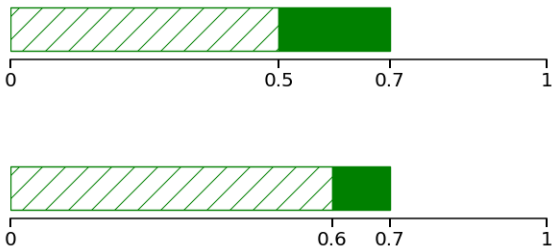


Figure V.7: Two ϵ-close lifebars, with $\epsilon = 0.1$.

Let us apply this result to the Čech bundle filtrations. Let X and Y be two subsets of $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Suppose that the Hausdorff distance $\mathrm{d_H}(X,Y)$, with respect to the norm $\|\cdot\|_\gamma$, is not greater than ϵ, meaning that the ϵ-thickenings X^ϵ and Y^ϵ satisfy $Y \subseteq X^\epsilon$ and $X \subseteq Y^\epsilon$. It is then clear that the vector bundle

filtrations are ϵ-interleaved, and we can apply Theorem V.7 to obtain the following result.

Corollary V.8. *If two subsets $X, Y \subset E$ satisfy $\mathrm{d_H}(X, Y) \leq \epsilon$, then there exists an ϵ-interleaving between the persistent cohomology modules of their corresponding Čech bundle filtrations which sends persistent Stiefel-Whitney classes on persistent Stiefel-Whitney classes.*

Example V.9. In order to illustrate Corollary V.8, consider the sets X' and Y' represented in Figure V.8. They are noisy samples of the sets X and Y defined in Example V.5. They contain 50 points each.

Figure V.8: Representation of the sets $X', Y' \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$.

Figure V.9 represents the barcodes of the Čech filtrations of the sets X' and Y', together with the lifebar of the first persistent Stiefel-Whitney class of their corresponding Čech bundle filtrations. Observe that they are close to the original descriptors of X and Y (Figure V.6).

Experimentally, we computed that the Hausdorff distances between X, X' and Y, Y' are approximately $\mathrm{d_H}(X, X') \approx 0.5$ and $\mathrm{d_H}(Y, Y') \approx 0.4$. Observe that this is coherent with the lifebar of $w_1(Y')$, which is ϵ-close to the lifebar of $w_1(Y)$ with $\epsilon \approx 0.3 \leq 0.4$.

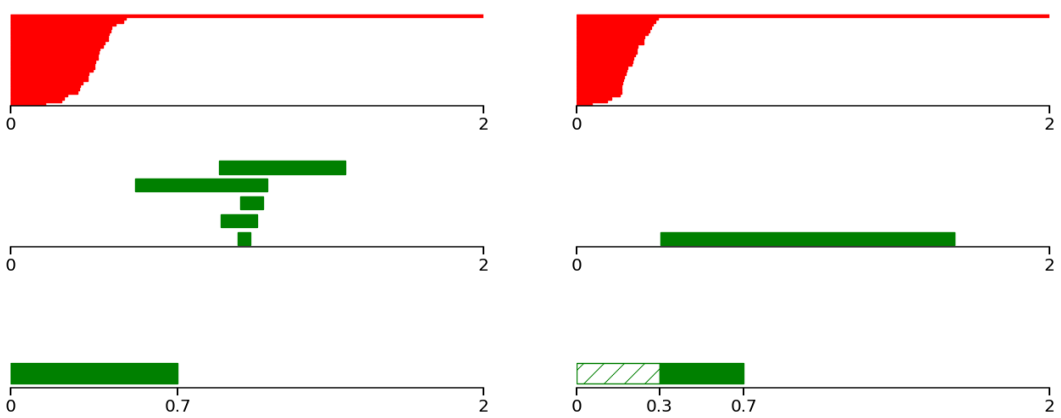


Figure V.9: Left: H^0 and H^1 barcodes of X' and lifebar of $w_1(X')$. **Right:** same for Y'.

V.1.4 Consistency

In this subsection we describe a setting where the persistent Stiefel-Whitney classes $w_i(X)$ of the Čech bundle filtration of a set X can be seen as consistent estimators of the Stiefel-Whitney classes of some underlying vector bundle.

Let $\mathcal{M}_0$ be a compact $\mathcal{C}^3$-manifold, and $u\colon \mathcal{M}_0 \to \mathcal{M} \subset \mathbb{R}^n$ an immersion. Suppose that $\mathcal{M}_0$ is given a d-dimensional vector bundle structure $p\colon \mathcal{M}_0 \to \mathcal{G}_d(\mathbb{R}^m)$. Let $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$, and consider the set

$$\check{\mathcal{M}} = \left\{ \left(u_0(x_0), P_{p(x_0)}\right), x_0 \in \mathcal{M}_0 \right\} \subset E, \tag{V.6}$$

where $P_{p(x_0)}$ denotes the orthogonal projection matrix onto the subspace $p(x_0) \subset \mathbb{R}^m$. The set $\check{\mathcal{M}}$ is called the *lift* of $\mathcal{M}_0$. Consider the *lifting map* defined as

$$\begin{aligned} \check{u}\colon \mathcal{M}_0 &\longrightarrow \check{\mathcal{M}} \subset E \\ x_0 &\longmapsto \left(u_0(x_0), P_{p(x_0)}\right). \end{aligned} \tag{V.7}$$

We make the following assumption: $\check{u}$ is an embedding. As a consequence, $\check{\mathcal{M}}$ is a submanifold of E, and $\mathcal{M}_0$ and $\check{\mathcal{M}}$ are diffeomorphic. It is worth noting that this point of view is strongly connected to Chapter IV, where we estimated the tangent bundle of an immersed manifold in order to get back to the abstract manifold $\mathcal{M}_0$.

The persistent cohomology of $\check{\mathcal{M}}$ can be used to recover the cohomology of $\mathcal{M}_0$. To see this, select $\gamma > 0$, and denote by $\mathrm{reach}\left(\check{\mathcal{M}}\right)$ the reach of $\mathcal{M}$, where E is endowed with the norm $\|\cdot\|_\gamma$. Since $\check{\mathcal{M}}$ is a $\mathcal{C}^2$-submanifold, $\mathrm{reach}\left(\check{\mathcal{M}}\right)$ is positive. Note that it depends on γ. Let $V[\check{\mathcal{M}}] = (\check{\mathcal{M}}^t)_{t\geq 0}$ be the Čech set filtration of $\check{\mathcal{M}}$ in the ambient space $(E, \|\cdot\|_\gamma)$, and let $\mathbb{V}(\check{\mathcal{M}})$ be the corresponding persistent cohomology module. For every $s, t \in [0, \mathrm{reach}\left(\check{\mathcal{M}}\right))$ such that $s \leq t$, we know that the inclusion maps $i_s^t\colon \check{\mathcal{M}}^s \hookrightarrow \check{\mathcal{M}}^t$ are homotopy equivalences (see Subsection II.3.3). Hence the persistence module $\mathbb{V}(\check{\mathcal{M}})$ is constant on the interval $[0, \mathrm{reach}\left(\check{\mathcal{M}}\right))$, and is equal to the cohomology $H^*(\check{\mathcal{M}}) = H^*(\mathcal{M}_0)$.

Consider the Čech bundle filtration $(V[\check{\mathcal{M}}], \mathbb{p})$ of $\check{\mathcal{M}}$. The following theorem shows that the persistent Stiefel-Whitney classes $w_i^t(\check{\mathcal{M}})$ are also equal to the usual Stiefel-Whitney classes of the vector bundle $(\mathcal{M}_0, p)$.

Theorem V.10. *Let $\mathcal{M}_0$ be a compact $\mathcal{C}^3$-manifold, $u\colon \mathcal{M}_0 \to \mathbb{R}^n$ an immersion and $p\colon \mathcal{M}_0 \to \mathcal{G}_d(\mathbb{R}^m)$ a continuous map. Let $\check{\mathcal{M}}$ be the lift of $\mathcal{M}_0$ (Equation* (V.6)*) and $\check{u}$ the lifting map (Equation* (V.7)*). Suppose that u is an embedding.*

Let $\gamma > 0$ and consider the Čech bundle filtration $(V[\check{\mathcal{M}}], \mathbb{p})$ of $\check{\mathcal{M}}$. Its maximal filtration value is $t_\gamma^{\max}\left(\check{\mathcal{M}}\right) = \frac{\sqrt{2}}{2}\gamma$. Denote by $w_i(\mathbb{p}) = (w_i^t(\mathbb{p}))_{t\in T}$ its persistent Stiefel-Whitney classes, $i \in [1, d]$. Denote also by i_0^t the inclusion $\check{\mathcal{M}} \to \check{\mathcal{M}}^t$, for $t \in [0, \mathrm{reach}\left(\check{\mathcal{M}}\right))$.

Let $t \geq 0$ be such that $t < \min\left(\mathrm{reach}\left(\check{\mathcal{M}}\right), t_\gamma^{\max}\left(\check{\mathcal{M}}\right)\right)$. Then the map $i_0^t \circ \check{u}\colon \mathcal{M}_0 \to \check{\mathcal{M}}^t$ induces an isomorphism $H^*(\mathcal{M}_0) \leftarrow H^*(\check{\mathcal{M}}^t)$ which maps the i^{th} persistent Stiefel-Whitney class $w_i^t(\mathbb{p})$ of $(V[\check{\mathcal{M}}], \mathbb{p})$ to the i^{th} Stiefel-Whitney class of $(\mathcal{M}_0, p)$.

Proof. Consider the following commutative diagram, defined for every $t < t_\gamma^{\max}\left(\check{\mathcal{M}}\right)$:

$$\begin{array}{ccccc} \mathcal{M}_0 & \xrightarrow{\check{u}} & \check{\mathcal{M}} & \xhookrightarrow{i_0^t} & \check{\mathcal{M}}^t \\ & \searrow^{p} & & \swarrow_{p^t} & \\ & & \mathcal{G}_d(\mathbb{R}^m) & & \end{array}$$

We obtain a commutative diagram in cohomology:

$$\begin{array}{ccccc} H^*(\mathcal{M}_0) & \xleftarrow{\check{u}^*} & H^*(\check{\mathcal{M}}) & \xleftarrow{(i_0^t)^*} & H^*(\check{\mathcal{M}}^t) \\ & \nwarrow_{p^*} & & \nearrow_{(p^t)^*} & \\ & & H^*(\mathcal{G}_d(\mathbb{R}^m)) & & \end{array}$$

Since $t < \text{reach}\left(\check{\mathcal{M}}\right)$, the map $(i_0^t)^*$ is an isomorphism (see Subsection II.3.3). So is $\check{u}^*$ since $\check{u}$ is an embedding. As a consequence, the map $i_0^t \circ \check{u}$ induces an isomorphism $H^*(\mathcal{M}_0) \simeq H^*(\check{\mathcal{M}}^t)$.

Let w_i denotes the i^{th} Stiefel-Whitney class of $\mathcal{G}_d(\mathbb{R}^m)$. By definition, the i^{th} Stiefel-Whitney class of $(\mathcal{M}_0, p)$ is $p^*(w_i)$, and the i^{th} persistent Stiefel-Whitney class of $(V[\check{\mathcal{M}}], \mathbb{p})$ is $w_i^t(\mathbb{p}) = (p^t)^*(w_i)$. By commutativity of the diagram, we obtain $p^*(w_i) = (p^t)^*(w_i)$, under the identification $H^*(\mathcal{M}_0) \simeq H^*(\check{\mathcal{M}}^t)$.

Applying Theorems V.7, V.10 and the considerations of Subsection II.3.3 yield an estimation result.

Corollary V.11. *Let $X \subset E$ be any subset such that $\mathrm{d_H}\left(X, \check{\mathcal{M}}\right) \leq \frac{1}{17}\text{reach}\left(\check{\mathcal{M}}\right)$. Define $\epsilon = \mathrm{d_H}\left(X, \check{\mathcal{M}}\right)$. Then for every $t \in \left[4\epsilon, \text{reach}\left(\check{\mathcal{M}}\right) - 3\epsilon\right)$, the composition of inclusions $\mathcal{M}_0 \hookrightarrow \check{\mathcal{M}} \hookrightarrow X^t$ induces an isomorphism $H^*(\mathcal{M}_0) \leftarrow H^*(X^t)$ which sends the i^{th} persistent Stiefel-Whitney class $w_i^t(X)$ of the Čech bundle filtration of X to the i^{th} Stiefel-Whitney class of $(\mathcal{M}_0, p)$.*

As a consequence of this corollary, on the set $[4\epsilon, \text{reach}\left(\check{\mathcal{M}}\right) - 3\epsilon)$, the i^{th} persistent Stiefel-Whitney class of the Čech bundle filtration of X is zero if and only if the i^{th} Stiefel-Whitney class of $(\mathcal{M}_0, p)$ is.

Example V.12. In order to illustrate Corollary V.11, consider the torus and the Klein bottle, immersed in $\mathbb{R}^3$ as in Figure V.10.

Figure V.10: Immersion of the torus and the Klein bottle in $\mathbb{R}^3$.

Let them be endowed with their normal bundles. They can be seen as submanifolds $\check{\mathcal{M}}, \check{\mathcal{M}}'$ of $\mathbb{R}^3 \times \mathrm{M}(\mathbb{R}^3)$. We consider two samples X, X' of $\check{\mathcal{M}}, \check{\mathcal{M}}'$, represented

in Figure V.11. They contain respectively 346 and 1489 points. We computed experimentally the Hausdorff distances $\mathrm{d_H}\left(X, \check{\mathcal{M}}\right) \approx 0.6$ and $\mathrm{d_H}\left(X', \check{\mathcal{M}}'\right) \approx 0.45$, with respect to the norm $\|\cdot\|_\gamma$ where $\gamma = 1$.

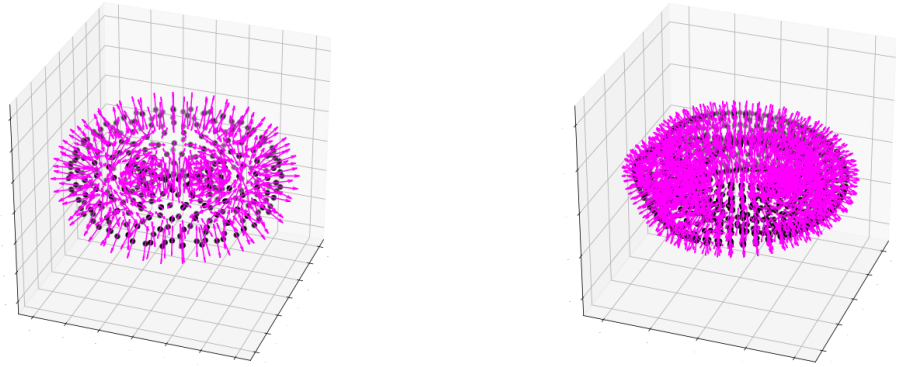


Figure V.11: Samples X and X' of $\check{\mathcal{M}}$ and $\check{\mathcal{M}}'$. The black points corresponds to the $\mathbb{R}^3$-coordinate, and the pink arrows over them correspond to the orientation of the $\mathrm{M}(\mathbb{R}^3)$-coordinate.

Figure V.12 represents the barcodes of the persistent cohomology of X and X', and the lifebars of their first persistent Stiefel-Whitney classes $w_1(X)$ and $w_1(X')$. Observe that $w_1(X)$ is always zero, while $w_1(X')$ is nonzero for $t \geq 0.3$. This is an indication that $\check{\mathcal{M}}$, the underlying manifold of X, is orientable, while $\check{\mathcal{M}}'$ is not. To see this, recall Proposition II.5: the first Stiefel-Whitney class of the tangent bundle of a manifold is zero if and only if the manifold is orientable. One can deduce the following fact: the first Stiefel-Whitney class of the normal bundle of an immersed manifold is zero if and only if the manifold is orientable (see the following lemma). Therefore, one interprets these lifebars as follows: X is sampled on an orientable manifold, while X' is sampled on a non-orientable one.

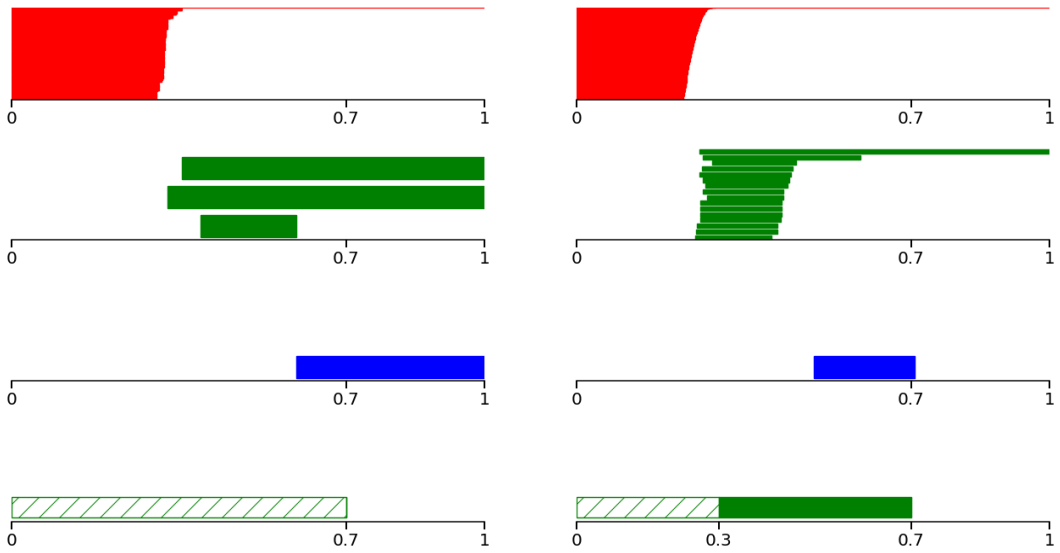


Figure V.12: Left: H^0, H^1 and H^2 barcodes of X and lifebar of $w_1(X)$. **Right:** same for X'.

Lemma V.13. *Let $\mathcal{M}_0 \to \mathcal{M}$ be an immersion of a manifold $\mathcal{M}_0$ in a Euclidean space. Then $\mathcal{M}_0$ is orientable if and only if the first Stiefel-Whitney class of its normal bundle is zero.*

Proof. Let τ and ν denote the tangent and normal bundles of $\mathcal{M}_0$. The Whitney sum $\tau \oplus \nu$ is a trivial bundle, hence its first Stiefel-Whitney class is $w_1(\tau \oplus \nu) = 0$. Using Axioms 1 and 3 of the Stiefel-Whitney classes, we obtain

$$\begin{aligned} w_1(\tau \oplus \nu) &= w_1(\tau) \smile w_0(\nu) + w_0(\tau) \smile w_1(\nu) \\ &= w_1(\tau) \smile 1 + 1 \smile w_1(\nu) \\ &= w_1(\tau) + w_1(\nu). \end{aligned}$$

Therefore, $w_1(\tau) = w_1(\nu)$. To conclude, $w_1(\tau)$ is zero if and only if $w_1(\nu)$ is zero, and Proposition II.5 yields the result.

V.2 Computation of persistent Stiefel-Whitney classes

In order to build an effective algorithm to compute the persistent Stiefel-Whitney classes, we have to find an equivalent formulation in terms of simplicial cohomology. We will make use of the well-known technique of simplicial approximation (see Subsection II.2.2 for definitions).

V.2.1 Simplicial approximation to Čech bundle filtrations

In this subsection, we apply the principle of simplicial approximation to the particular case of persistent Stiefel-Whitney classes of Čech bundle filtrations.

Let X be a subset of $E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Let us recall Definition V.4: the Čech bundle filtration associated to X is the vector bundle filtration $(\mathbb{X}, \mathbb{p})$ whose underlying filtration is the Čech filtration $\mathbb{X} = (X^t)_{t \in T}$, with $T = [0, t_\gamma^{\max}(X))$, and whose maps $\mathbb{p} = (p^t)_{t \in T}$ are given by the following composition, as in Equation (V.2):

$$X^t \xrightarrow{\mathrm{proj}_2} \mathrm{M}(\mathbb{R}^m) \setminus \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right) \xrightarrow{\mathrm{proj}(\cdot, \mathcal{G}_d(\mathbb{R}^m))} \mathcal{G}_d(\mathbb{R}^m).$$
$$p^t$$

Let $t \in T$. The aim of this subsection is to describe a simplicial approximation to $p^t \colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$. To do so, let us fix a triangulation L of $\mathcal{G}_d(\mathbb{R}^m)$. It comes with a homeomorphism $h \colon \mathcal{G}_d(\mathbb{R}^m) \to |L|$. We shall now triangulate the Čech set filtration X^t. The thickening X^t is a subset of the metric space $(E, \|\cdot\|_\gamma)$ which consists in a union of closed balls centered around points of X:

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}_\gamma(x, t),$$

where $\overline{\mathcal{B}}_\gamma(x,t)$ denotes the closed ball of center x and radius t for the norm $\|\cdot\|_\gamma$. Let $\mathcal{U}^t$ denote the cover $\left\{\overline{\mathcal{B}}_\gamma(x,t), x \in X\right\}$ of X^t, and let $\mathcal{N}(\mathcal{U}^t)$ be its nerve. By the nerve theorem for convex closed covers [BCY18, Theorem 2.9], the simplicial complex $\mathcal{N}(\mathcal{U}^t)$ is homotopy equivalent to its underlying set X^t. That is to say, there exists a continuous map $g^t\colon |\mathcal{N}(\mathcal{U}^t)| \to X^t$ which is a homotopy equivalence.

As a consequence, in cohomological terms, the map $p^t\colon X^t \to \mathcal{G}_d(E)$ is equivalent to the map q^t defined as $q^t = h \circ p^t \circ g^t$.

$$\begin{array}{ccc} X^t & \xrightarrow{\quad p^t \quad} & \mathcal{G}_d(\mathbb{R}^m) \\ {\scriptstyle g^t}\uparrow & & \downarrow{\scriptstyle h} \\ |\mathcal{N}(\mathcal{U}^t)| & \xrightarrow{\quad q^t \quad} & |L| \end{array} \tag{V.8}$$

This gives a way to compute the induced map $(p^t)^*\colon H^*(X^t) \leftarrow H^*(\mathcal{G}_d(\mathbb{R}^m))$ algorithmically:

- Subdivise $\mathcal{N}(\mathcal{U}^t)$ until q^t satisfies the star condition (as in Theorem II.1),
- Choose a simplicial approximation f^t to q^t,
- Compute the induced map between simplicial cohomology groups $(f^t)^*\colon H^*(\mathcal{N}(\mathcal{U}^t)) \leftarrow H^*(L)$.

By correspondence between simplicial and singular cohomology, the map $(f^t)^*$ corresponds to $(p^t)^*$. Hence the problem of computing $(p^t)^*$ is solved, if it were not for the following issue: in practice, the map $g^t\colon |\mathcal{N}(\mathcal{U}^t)| \to X^t$ given by the nerve theorem is not explicit. The rest of this subsection is devoted to showing that g^t can be chosen canonically as the *shadow map*.

Shadow map. We still consider X^t, the corresponding cover $\mathcal{U}^t$ and its nerve $\mathcal{N}(\mathcal{U}^t)$. The underlying vertex set of the simplicial complex $\mathcal{N}(\mathcal{U}^t)$ is the set X itself. The shadow map $g^t\colon |\mathcal{N}(\mathcal{U}^t)| \to X^t$ is defined as follows: for every simplex $\sigma = [x_0, ..., x_p] \in \mathcal{N}(\mathcal{U}^t)$ and every point $\sum_{i=0}^p \lambda_i x_i$ of $|\sigma|$ written in barycentric coordinates, associate the point $\sum_{i=0}^p \lambda_i x_i$ of E:

$$g^t\colon \sum_{i=0}^p \lambda_i x_i \in |\sigma| \longmapsto \sum_{i=0}^p \lambda_i x_i \in E.$$

We are not aware whether the shadow map is indeed a homotopy equivalence from $|\mathcal{N}(\mathcal{U}^t)|$ to X^t. Nevertheless, the following result will be enough for our purposes: the shadow map induces an isomorphism at cohomology level.

Lemma V.14. *Suppose that X is finite and in general position. Then the shadow map $g^t\colon |\mathcal{N}(\mathcal{U}^t)| \to X^t$ induces an isomorphism $(g^t)^*\colon H^*(|\mathcal{N}(\mathcal{U}^t)|) \leftarrow H^*(X^t)$.*

Proof. Recall that $\mathcal{U}^t = \left\{\overline{\mathcal{B}}_\gamma(x,t), x \in X\right\}$. Let us consider a smaller cover. For every $x \in X$, let $\mathrm{Vor}(x)$ denote the Voronoi cell of x in the ambient metric

space $(E, \|\cdot\|_\gamma)$, and define

$$\mathcal{V}^t = \left\{\overline{\mathcal{B}}(x,t) \cap \mathrm{Vor}(x), x \in X\right\}.$$

The set $\mathcal{V}^t$ is a cover of X^t, and its nerve $\mathcal{N}(\mathcal{V}^t)$ is known as the Delaunay complex (see [BE17]). Let $h^t\colon |\mathcal{N}(\mathcal{V}^t)| \to X^t$ denote the shadow map of $\mathcal{N}(\mathcal{V}^t)$. The Delaunay complex is a subcomplex of the Čech complex, hence we can consider the following diagram:

$$|\mathcal{N}(\mathcal{V}^t)| \overset{h^t}{\hookrightarrow} |\mathcal{N}(\mathcal{U}^t)| \xrightarrow[g^t]{} X^t.$$

This yields the following commutative diagram between cohomology rings:

$$H^*(|\mathcal{N}(\mathcal{V}^t)|) \xleftarrow{(h^t)^*} H^*(|\mathcal{N}(\mathcal{U}^t)|) \xleftarrow[(g^t)^*]{} H^*(X^t).$$

Now, it is proven in [Ede93, Theorem 3.2] that the shadow map $h^t\colon |\mathcal{N}(\mathcal{V}^t)| \to X^t$ is a homotopy equivalence (it is required here that X is in general position). Therefore the map $(h^t)^*\colon H^*(|\mathcal{N}(\mathcal{V}^t)|) \leftarrow H^*(X^t)$ is an isomorphism. Moreover, we know from [BE17, Theorem 5.10] that $\mathcal{N}(\mathcal{U}^t)$ collapses to $\mathcal{N}(\mathcal{V}^t)$. Therefore the inclusion $|\mathcal{N}(\mathcal{V}^t)| \hookrightarrow |\mathcal{N}(\mathcal{U}^t)|$ also is a homotopy equivalence, hence the induced map $H^*(|\mathcal{N}(\mathcal{V}^t)|) \leftarrow H^*(|\mathcal{N}(\mathcal{U}^t)|)$ is an isomorphism. We conclude from the last diagram that $(g^t)^*$ is an isomorphism.

V.2.2 A sketch of algorithm

Suppose that we are given a finite set $X \subset E = \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Choose $d \in [1, n-1]$ and $\gamma > 0$. Consider the Čech bundle filtration of dimension d of X. Let $T = \left[0, t_\gamma^{\max}(X)\right)$, $t \in T$ and $i \in [1, d]$. From the previous discussion we can infer an algorithm to solve the following problem:

Compute the persistent Stiefel-Whitney class $w_i^t(X)$ of the Čech bundle filtration of X, using a cohomology computation software.

Denote:

- $\mathbb{X} = (X^t)_{t\geq 0}$ the Čech set filtration of X,
- $\mathbb{S}$ the Čech simplicial filtration of X, and $g^t\colon |S^t| \to X^t$ the shadow map,
- L a triangulation of $\mathcal{G}_d(\mathbb{R}^n)$ and $h\colon \mathcal{G}_d(\mathbb{R}^n) \to |L|$ a homeomorphism,
- $(\mathbb{X}, \mathbb{p})$ the Čech bundle filtration of X,
- $(\mathbb{V}, \mathbb{v})$ the persistent cohomology module of $\mathbb{X}$,
- $w_i \in H^i(\mathcal{G}_d(\mathbb{R}^n))$ the i^{th} Stiefel-Whitney class of the Grassmannian.

Let $t \in T$ and consider the map q^t, as defined in Equation (V.8):

$$|S^t| \xrightarrow[g^t]{} X^t \xrightarrow[p^t]{} \mathcal{G}_d(\mathbb{R}^m) \xrightarrow[h]{} |L|\,, \qquad |S^t| \xrightarrow{\;q^t\;} |L|\,.$$

We propose the following algorithm:

- Subdivise barycentrically S^t until q^t satisfies the star condition. Denote k the number of subdivisions needed.
- Consider a simplicial approximation $f^t\colon \mathrm{sub}^k(S^t) \to L$ to q^t.
- Compute the class $(f^t)^*(w_i)$.

The output $(f^t)^*(w_i)$ is equal to the persistent Stiefel-Whitney class $w_i^t(X)$ at time t, seen in the simplicial cohomology group $H^i(S^t) = H^i(\mathrm{sub}^k(S^t))$. In the following section, we gather some technical details needed to implement this algorithm in practice.

Computing the lifebar. This algorithm also gives a way to compute the lifebar of $w_i(X)$. We remind the reader that the lifebar is determined by the value $t^\dagger = \inf\{t \in T, w_i^t(X) \neq 0\}$. This quantity can be approximated by computing the classes $w_i^t(X)$ for several values of t, by performing a dichotomic search for instance.

We now briefly describe another algorithm, in the particular case of $i = 1$. Using the previous notations, the first persistent Stiefel-Whitney class $w_1(X) = (w_i^t(X))_t$ is defined by $w_i^t(X) = (p^t)^*(w_1)$, where w_1 is the first Stiefel-Whitney class of the Grassmannian. Since $H^1(\mathcal{G}_d(\mathbb{R}^m))$ is the free abelian group generated by w_1, the map $(p^t)^*\colon H^1(X^t) \leftarrow H^1(\mathcal{G}_d(\mathbb{R}^m))$ has rank 0 or 1. Therefore the value $t^\dagger$ is equal to

$$t^\dagger = \inf\big\{t \in T, \mathrm{rank}\big((p^t)^*\big) = 1\big\}.$$

Let $C(p^t)$ denote the mapping cone of $p^t\colon X^t \to \mathcal{G}_d(\mathbb{R}^m)$. This is a usual construction in algebraic topology, that we do not detail here. The mapping cone $C(p^t)$ is a topological space, and the family of mapping cones $(C(p^t))_t$ can be seen as a filtration of topological spaces. The mapping cones come with a long exact sequence

$$\cdots \longrightarrow H^k(X^t) \longrightarrow H^{k+1}(C(p^t)) \longrightarrow H^{k+1}(\mathcal{G}_d(\mathbb{R}^m)) \longrightarrow H^{k+1}(X^t) \longrightarrow \cdots$$

from which we deduce a formula for the rank of $(p^t)^*\colon H^1(X^t) \leftarrow H^1(\mathcal{G}_d(\mathbb{R}^m))$:

$$\mathrm{rank}((p^t)^*) = \sum_{k=1}^{+\infty} (-1)^k \Big(\dim H^k(X^t) - \dim H^{k+1}(C(p^t)) + \dim H^{k+1}(\mathcal{G}_d(\mathbb{R}^m)) \Big).$$

In practice, the sum is finite. The usual persistent homology algorithm allows to compute exactly the functions $t \mapsto \dim H^k(X^t)$ and $t \mapsto \dim H^{k+1}(C(p^t))$, known as the Betti curves. Eventually, by using the formula, one identifies $t^\dagger$ as the first value of t such that $\mathrm{rank}((p^t)^*) = 1$.

V.3 An algorithm when $d = 1$

Even though the last sections described a theoretical way to compute the persistent Stiefel-Whitney classes, some concrete issues are still to be discussed:

- verifying that the star condition is satisfied,
- the Grassmann manifold has to be triangulated,
- in practice, the Vietoris-Rips filtration is preferred to the Čech filtration,
- the parameter γ has to be tuned.

The following subsections will elucidate these points. Concerning the first one, we are not aware of a computational-explicit process to triangulate the Grassmann manifolds $\mathcal{G}_d(\mathbb{R}^m)$, except when $d = 1$, which corresponds to the projective spaces $\mathcal{G}_1(\mathbb{R}^m)$. We shall then restrict to the case $d = 1$.

V.3.1 The star condition in practice

Let us get back to the context of Subsection II.2.2: K, L are two simplicial complexes, K is finite, and $g\colon |K| \to |L|$ is a continuous map. We have seen that finding a simplicial approximation to g reduces to finding a small enough barycentric subdivision $\mathrm{sub}^n(K)$ of K such that $g\colon |\mathrm{sub}^n(K)| \to |L|$ satisfies the star condition, that is, for every vertex v of $\mathrm{sub}^n(K)$, there exists a vertex w of L such that

$$g\left(\left|\overline{\mathrm{St}}(v)\right|\right) \subseteq |\mathrm{St}(w)|\,.$$

In practice, one can compute the closed star $\overline{\mathrm{St}}(v)$ from the finite simplicial complex $\mathrm{sub}^n(K)$. However, computing $g\left(\left|\overline{\mathrm{St}}(v)\right|\right)$ requires to evaluate g on the infinite set $\left|\overline{\mathrm{St}}(v)\right|$. In order to reduce the problem to a finite number of evaluations of g, we shall consider a related property that we call the *weak star condition*.

Definition V.15 (Weak star condition). A map $g\colon |K| \to |L|$ between topological realizations of simplicial complexes K and L satisfies the *weak star condition* if for every vertex v of $\mathrm{sub}^n(K)$, there exists a vertex w of L such that

$$\left|g\left(\overline{\mathrm{St}}(v)^{(0)}\right)\right| \subseteq |\mathrm{St}(w)|\,,$$

where $\overline{\mathrm{St}}(v)^{(0)}$ denotes the 0-skeleton of $\overline{\mathrm{St}}(v)$, i.e. its vertices.

Observe that the practical verification of the condition $\left|g\left(\overline{\mathrm{St}}(v)^{(0)}\right)\right| \subseteq |\mathrm{St}(w)|$ requires only a finite number of computations. Indeed, one just has to check whether every neighbor v' of v in the graph $K^{(1)}$, v included, satisfies $g(v') \in |\mathrm{St}(w)|$. The following lemma rephrases this condition by using the face map $\mathcal{F}_L\colon |L| \to L$. We remind the reader that the face map is defined by the relation $x \in \mathcal{F}_L(x)$ for all $x \in |L|$ (see Subsection II.2.2).

Lemma V.16. *The map g satisfies the weak star condition if and only if for every vertex v of K, there exists a vertex w of L such that for every neighbor*

v' of v in $K^{(1)}$, we have

$$w \in \mathcal{F}_L(g(v')).$$

Proof. Let us show that the assertion "$w \in \mathcal{F}_L(g(v'))$" is equivalent to "$g(v') \in |\mathrm{St}(w)|$". Remind that the open star $\mathrm{St}(w)$ consists of the simplices of L that contain w. Moreover, the topological realization $|\mathrm{St}(w)|$ is the union of the $|\sigma|$ for $\sigma \in \mathrm{St}(w)$. As a consequence, $g(v')$ belongs to $|\mathrm{St}(w)|$ if and only if it belongs to $|\sigma|$ for some simplex $\sigma \in L$ that contains w. Equivalently, the face map $\mathcal{F}_L(g(v'))$ contains w.

Suppose that g satisfies the weak star condition. Let $f\colon K^{(0)} \to L^{(0)}$ be a map between vertex sets such that for every $v \in K^{(0)}$,

$$\left|g\left(\overline{\mathrm{St}}(v)^{(0)}\right)\right| \subseteq |\mathrm{St}(f(v))|\,.$$

According to the proof of Lemma V.16, an equivalent formulation of this condition is: for all neighbor v' of v in $K^{(1)}$,

$$f(v) \in \mathcal{F}_L(g(v')). \tag{V.9}$$

Such a map is called a *weak simplicial approximation to g*. It plays a similar role as the simplicial approximations to g.

Lemma V.17. *If $f\colon K^{(0)} \to L^{(0)}$ is a weak simplicial approximation to $g\colon |K| \to |L|$, then f is a simplicial map.*

Proof. Let $\sigma = [v_0, ..., v_n]$ be a simplex of K. We have to show that $f(\sigma) = [f(v_0), ..., f(v_n)]$ is a simplex of L. Note that each closed star $\overline{\mathrm{St}}(v_i)$ contains σ. Therefore each $\left|g\left(\overline{\mathrm{St}}(v_i)^{(0)}\right)\right|$ contains $\left|g\left(\sigma^{(0)}\right)\right| = \{g(v_0), ..., g(v_n)\}$. Using the weak simplicial approximation property of f, we deduce that each $|\mathrm{St}(f(v_i))|$ contains $\{g(v_0), ..., g(v_n)\}$. Using Lemma V.18 stated below, we obtain that $[f(v_0), ..., f(v_n)]$ is a simplex of L.

Lemma V.18 ([Hat02, Lemma 2C.2]). *Let $w_0, ..., w_n$ be vertices of a simplicial complex L. Then $\bigcap_{i=0}^{n} \mathrm{St}(w_i) \neq \emptyset$ if and only if $[w_0, ..., w_n]$ is a simplex of L.*

As one can see from the definitions, the weak star condition is weaker than the star condition. Consequently, the simplicial approximation theorem admits the following corollary.

Corollary V.19. *Consider two simplicial complexes K, L with K finite, and let $g\colon |K| \to |L|$ be a continuous map. Then there exists $n \geq 0$ such that $g\colon |\mathrm{sub}^n(K)| \to |L|$ satisfies the weak star condition.*

However, some weak simplicial approximations to g may not be simplicial approximations, and may not even be homotopic to g. Figure V.13 gives such an example.

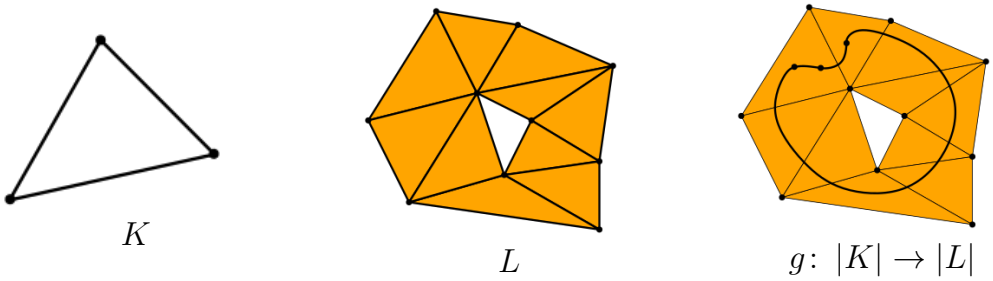


Figure V.13: The map g admits a weak simplicial approximation which is constant.

Fortunately, these two notions coincides under the star condition assumption:

Proposition V.20. *Suppose that g satisfies the star condition. Then every weak simplicial approximation to g is a simplicial approximation.*

Proof. Let f be a weak simplicial approximation to g, and f' any simplicial approximation. Let us show that f and f' are contiguous simplicial maps. Let $\sigma = [v_0, ..., v_n]$ be a simplex of K. We have to show that $[f(v_0), ..., f(v_n), f'(v_0), ..., f'(v_n)]$ is a simplex of L. As we have seen in the proof of Lemma V.17, each $\left|g\left(\overline{\mathrm{St}}(v_i)^{(0)}\right)\right|$ contains $\{g(v_0), ..., g(v_n)\}$. Therefore, by definition of weak simplicial approximations and simplicial approximations, each $|\mathrm{St}(f(v_i))|$ and $|\mathrm{St}(f'(v_i))|$ contains $\{g(v_0), ..., g(v_n)\}$. We conclude by applying Lemma V.18.

Remark that the proof of this proposition can be adapted to obtain the following fact: any two weak simplicial approximations are equivalent—as well as any two simplicial approximations.

Let us comment Proposition V.20. If K is subdivised enough, then every weak simplicial approximation to g is homotopic to g. We face the following problem in practice: the number of subdivisions needed by the star condition is not known. In order to work around this problem, we propose to subdivise the complex K until it satisfies the weak star condition, and then use a weak simplicial approximation to g. However, such a weak simplicial approximation may not be homotopic to g, and our algorithm would output a wrong result. To close this subsection, we state a lemma that gives a quantitative idea of the number of subdivisions needed. We say that a Lebesgue number for an open cover $\mathcal{U}$ of a compact metric space X is a positive number ϵ such that every subset of X with diameter less than ϵ is included in some element of the cover $\mathcal{U}$.

Lemma V.21. *Let $|K|, |L|$ be endowed with metrics. Suppose that $g\colon |K| \to |L|$ is l-Lipschitz with respect to these metrics. Let ϵ be a Lebesgue number for the open cover $\{|\mathrm{St}(w)|, w \in L\}$ of $|L|$. Let p be the dimension of K and D an upper bound on the diameter of its faces. Then for $n > \log(\frac{Dl}{\epsilon}) / \log(\frac{p+1}{p})$, the map $g\colon |\mathrm{sub}^n(K)| \to |L|$ satisfies the star condition.*

Proof. The map g satisfies the star condition if for every vertex v of K, there exists a vertex w of L such that $g(\left|\overline{\mathrm{St}}(v)\right|) \subseteq |\mathrm{St}(w)|$. Since the cover

$\{|\mathrm{St}(w)|, w \in L\}$ admits ϵ as a Lebesgue number, it is enough for v to satisfy the following inequality:

$$\mathrm{diam}\left(g(\left|\overline{\mathrm{St}}(v)\right|)\right) < \epsilon. \tag{V.10}$$

Since g is l-Lipschitz, we have $\mathrm{diam}\left(g\left(\left|\overline{\mathrm{St}}(v)\right|\right)\right) \leq l \cdot \mathrm{diam}\left(\left|\overline{\mathrm{St}}(v)\right|\right)$. Using the hypothesis $\mathrm{diam}\left(\left|\overline{\mathrm{St}}(v)\right|\right) \leq D$, Equation (V.10) leads to the condition $Dl < \epsilon$. Now, we use the fact that a barycentric subdivision reduces the diameter of each face by a factor $\frac{p}{p+1}$. After n barycentric subdivision, the last inequality rewrites $\left(\frac{p}{p+1}\right)^n Dl < \epsilon$. It admits $n > \log(\frac{Dl}{\epsilon}) / \log(\frac{p+1}{p})$ as a solution.

V.3.2 Triangulating the projective spaces

As we described in Subsection V.3.1, the algorithm we propose rests on a triangulation L of the Grassmannian $\mathcal{G}_1(\mathbb{R}^m)$, together with the map $\mathcal{F}_L \circ h \colon \mathcal{G}_1(\mathbb{R}^m) \to L$, where $h \colon \mathcal{G}_1(\mathbb{R}^m) \to |L|$ is a homeomorphism and $\mathcal{F}_L \colon \mathcal{G}_1(\mathbb{R}^m) \to L$ is the face map. In the following, we also call $\mathcal{F} := \mathcal{F}_L \circ h$ the face map.

We shall use the following folklore triangulation of the projective space $\mathcal{G}_1(\mathbb{R}^m)$. It uses the fact that the quotient of the sphere $\mathbb{S}_{m-1}$ by the antipodal relation gives $\mathcal{G}_1(\mathbb{R}^m)$. Let Δ^m denote the standard m-simplex, $v_0, ..., v_m$ its vertices, and $\partial\Delta^m$ its boundary. The simplicial complex $\partial\Delta^m$ is a triangulation of the sphere $\mathbb{S}_{m-1}$. Denote its first barycentric subdivision as $\mathrm{sub}^1(\partial\Delta^m)$. The vertices of $\mathrm{sub}^1(\partial\Delta^m)$ are in bijection with the non-empty proper subsets of $\{v_0, ..., v_m\}$ (see Subsection II.2.2). Consider the equivalence relation on these vertices which associates a vertex to its complement. The quotient simplicial complex under this relation, L, is a triangulation of $\mathcal{G}_1(\mathbb{R}^m)$.

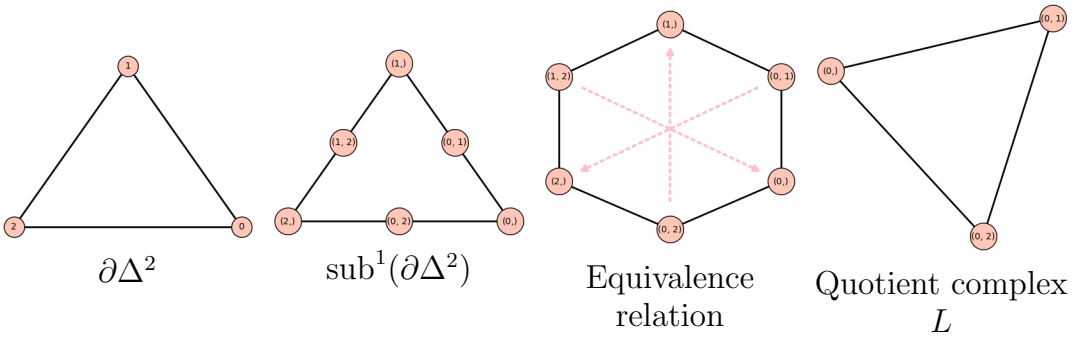
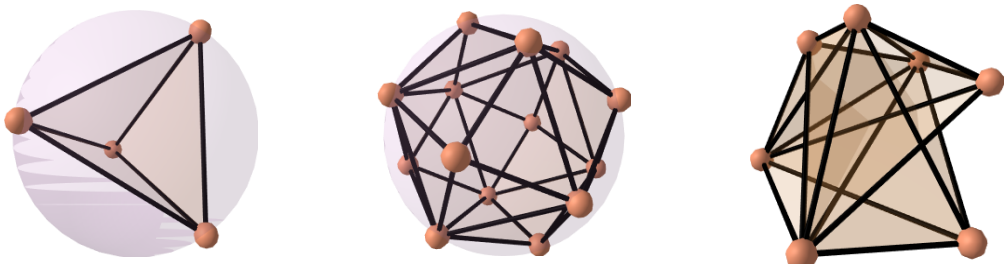


Figure V.14: Triangulating $\mathcal{G}_1(\mathbb{R}^2)$.

Let us now describe how to define the homeomorphism $h \colon \mathcal{G}_1(\mathbb{R}^m) \to |L|$. First, embed Δ^m in $\mathbb{R}^{m+1}$ via $v_i \mapsto (0, ..., 0, 1, 0, ...)$, where 1 sits at the i^{th} coordinate. Its image lies on a m-dimensional affine subspace P, with origin being the barycenter of $v_0, ..., v_m$. Seen in P, the vertices of Δ^m now belong to the sphere centered at the origin and of radius $\sqrt{\frac{m}{m+1}}$ (see Figure V.15). Let us denote this sphere as $\mathbb{S}_{m-1}$. Next, subdivise barycentrically $\partial\Delta^m$ once, and project each vertex of $\mathrm{sub}^1(\partial\Delta^m)$ on $\mathbb{S}_{m-1}$. By taking the convex hulls of its faces, we now see $\left|\mathrm{sub}^1(\partial\Delta^m)\right|$ as a subset

of P. Define an application $p\colon \mathbb{S}_{m-1} \to \left|\mathrm{sub}^1(\partial\Delta^m)\right|$ as follows: for every $x \in \mathbb{S}_{m-1}$, the image $p(x)$ is the unique intersection point between the segment $[0, x]$ and the set $\left|\mathrm{sub}^1(\partial\Delta^m)\right|$. The application p can also be seen as the inverse function of the projection on $\mathbb{S}_{m-1}$, written $\mathrm{proj}_{\mathbb{S}_{m-1}}\colon \left|\mathrm{sub}^1(\partial\Delta^m)\right| \to \mathbb{S}_{m-1}$.



$\partial\Delta^3$ is included in $\mathbb{S}_{m-1}$ $\mathrm{sub}^1(\partial\Delta^3)$ and $\mathbb{S}_{m-1}$ L

Figure V.15: Triangulating $\mathcal{G}_1(\mathbb{R}^3)$.

The next lemma shows that the antipodal relation on $\mathbb{S}_{m-1}$ can be pulled-back to $\left|\mathrm{sub}^1(\partial\Delta^m)\right|$ via p, and that it corresponds to the equivalence relation we defined on $\mathrm{sub}^1(\partial\Delta^m)$. As a consequence, we can factorize $p\colon \mathbb{S}_{m-1} \to \left|\mathrm{sub}^1(\partial\Delta^m)\right|$ as

$$h\colon \left(\mathbb{S}_{m-1}/\!\sim\right) \to \left(\left|\mathrm{sub}^1(\partial\Delta^m)\right|/\!\sim\right),$$

and we can identify these spaces with

$$h\colon \mathcal{G}_1(\mathbb{R}^m) \to |L|,$$

giving the desired homeomorphism.

Lemma V.22. *For any vertex $x \in \mathrm{sub}^1(\partial\Delta^m)$, denote by $|x|$ its embedding in P. Let $-|x|$ denote the image of $|x|$ by the antipodal relation on $\mathbb{S}_{m-1}$. Denote by y the image of x by the relation on $\mathrm{sub}^1(\partial\Delta^m)$. Then $y = -|x|$.*

More generally, pulling back the antipodal relation onto $\left|\mathrm{sub}^1(\partial\Delta^m)\right|$ via p gives the relation we defined on $\mathrm{sub}^1(\partial\Delta^m)$.

Proof. Pick a vertex x of $\mathrm{sub}^1(\partial\Delta^m)$. It can be described as a proper subset $\{v_i, i \in I\}$ of the vertex set $(\partial\Delta^m)^{(0)} = \{v_0, ..., v_m\}$, where $I \subset [0, m]$. According to the relation on $(\partial\Delta^m)$, the vertex x is in relation with the vertex y described by the proper subset $\{v_i, i \in I^c\}$.

The point x can be written in barycentric coordinates as $\frac{1}{\mathrm{card}(I)} \sum_{i\in I} |v_i|$. Seen in P, $|x|$ can be written $|x| = \mathrm{proj}_{\mathbb{S}_{m-1}}\left(\sum_{i\in I} v_i\right)$. Similarly, $|y|$ can be written $|y| = \mathrm{proj}_{\mathbb{S}_{m-1}}\left(\sum_{i\in I^c} v_i\right)$.

Now, denote by 0 the origin of the hyperplane P, and embed the vertices $v_0, ..., v_m$ in P. Observe that

$$0 = \sum_{i\leq 0} v_i = \sum_{i\in I} v_i + \sum_{i\in I^c} v_i.$$

Hence $-\sum_{i\in I} v_i = \sum_{i\in I^c} v_i$, and we deduce that

$$-|x| = \mathrm{proj}_{\mathbb{S}_{m-1}}\left(-\sum_{i\in I} v_i\right) = \mathrm{proj}_{\mathbb{S}_{m-1}}\left(\sum_{i\in I^c} v_i\right) = |y|\,.$$

Applying the same reasoning, one obtains the following result: for every simplex σ of $\mathrm{sub}^1(\partial\Delta^m)$, if ν denotes the image of σ by the relation on $\mathrm{sub}^1(\partial\Delta^m)$, then the image of $|\sigma|$ by the antipodal relation is also $|\nu|$. As a consequence, these two relations coincide.

At a computational level, let us describe how to compute the face map $\mathcal{F}\colon \mathcal{G}_1(\mathbb{R}^m) \to L$. Since $\mathcal{F}$ can be obtained as a quotient, it is enough to compute the face map of the sphere, $\mathcal{F}'\colon \mathbb{S}_{m-1} \to \mathrm{sub}^1(\partial\Delta^m)$, which corresponds to the homeomorphism $p\colon \mathbb{S}_{m-1} \to \left|\mathrm{sub}^1(\partial\Delta^m)\right|$. It is given by the following lemma, which can be used in practice.

Lemma V.23. *For every $x \in \mathbb{S}_{m-1}$, the image of x by $\mathcal{F}'$ is equal to the intersection of all maximal faces $\sigma = [w_0, ..., w_m]$ of $\mathrm{sub}^1(\partial\Delta^m)$ that satisfies the following conditions: denoting by x_0 any point of the affine hyperplane spanned by $\{w_0, ..., w_m\}$, and by h a vector orthogonal to the corresponding linear hyperplane,*

- *the inner product $\langle x, h\rangle$ has the same sign as $\langle x_0, h\rangle$,*
- *the point $\frac{\langle x_0,h\rangle}{\langle x,h\rangle}x$, which is included in the affine hyperplane spanned by $\{w_0, ..., w_m\}$, has nonnegative barycentric coordinates.*

Proof. Recall that for every $x \in \mathbb{S}_{m-1}$, the image $p(x)$ is defined as the unique intersection point between the segment $[0, x]$ and the set $\left|\mathrm{sub}^1(\partial\Delta^m)\right|$. Besides, the face map $\mathcal{F}'(x)$ is the unique simplex $\sigma \in \mathrm{sub}^1(\partial\Delta^m)$ such that $p(x) \in |\sigma|$. Equivalently, $\mathcal{F}'(x)$ is equal to the intersection of all maximal faces $\sigma \in \mathrm{sub}^1(\partial\Delta^m)$ such that $p(x)$ belongs to the closure $\overline{|\sigma|}$.

Consider any maximal face $\sigma = [w_0, ..., w_m]$ of $\mathrm{sub}^1(\partial\Delta^m)$. The first condition of the lemma ensures that the segment $[0, x]$ intersects the affine hyperplane spanned by $\{w_0, ..., w_m\}$. In this case, a computation shows that this intersection consists of the point $\frac{\langle x_0,h\rangle}{\langle x,h\rangle}x$. Then, the second condition of the lemma tests whether this point belongs to the convex hull of $\{w_0, ..., w_k\}$. In conclusion, if σ satisfies these two conditions, then $p(x) \in \overline{|\sigma|}$.

As a remark, let us point out that the verification of the conditions of this lemma is subject to numerical errors. In particular, the point $\frac{\langle x_0,h\rangle}{\langle x,h\rangle}x$ may have nonnegative coordinates, yet mathematical softwares may return (small) negative values. Consequently, the algorithm may recognize less maximal faces that satisfy these conditions, hence return a simplex that strictly contains the wanted simplex $\mathcal{F}'(x)$. Nonetheless, such an error will not affect the output of the algorithm. Indeed, if we denote by $\widetilde{\mathcal{F}'}$ the face map computed by the algorithm, we have that $\mathcal{F}'(x) \subseteq \widetilde{\mathcal{F}'}(x)$ for all $x \in \mathbb{S}_{m-1}$. As a consequence of Lemma V.16, $\widetilde{\mathcal{F}'}$ satisfies the weak star condition if $\mathcal{F}'$ does, and Equation (V.9) shows that every weak

simplicial approximations for $\mathcal{F}'$ are weak simplicial approximations for $\widetilde{\mathcal{F}'}$. Since every weak simplicial approximations are homotopic, we obtain that the induced maps in cohomology are equal, therefore the output of the algorithm is unchanged.

V.3.3 Vietoris-Rips version of the Čech bundle filtration

We still consider a subset $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. Denote by $\mathbb{X}$ the corresponding Čech set filtration, and by $\mathbb{S} = (S^t)_{t\geq 0}$ the simplicial Čech filtration. For every $t \geq 0$, let R^t be the flag complex of S^t, i.e. the clique complex of the 1-skeleton $(S^t)^{(1)}$ of S^t. It is known as the *Vietoris-Rips complex* of X at time t. The collection $\mathbb{R} = (R^t)_{t\geq 0}$ is called the *Vietoris-Rips filtration* of X. The simplicial filtrations $\mathbb{S}$ and $\mathbb{R}$ are multiplicatively $\sqrt{2}$-interleaved [BLM+19a, Theorem 3.1]. In other words, for every $t \geq 0$, we have

$$S^t \subseteq R^t \subseteq S^{\sqrt{2}t}.$$

Let $\gamma > 0$ and consider the Čech bundle filtration $(\mathbb{X}, \mathbb{p})$ of X. Suppose that its maximal filtration value $t_\gamma^{\max}(X)$ is positive. Let $|\mathbb{R}| = (|R^t|)_{t\geq 0}$ denote the topological realization of the Vietoris-Rips filtration. We can give $|\mathbb{R}|$ a vector bundle filtration structure with $(p')^t \colon |R^t| \to \mathcal{G}_d(\mathbb{R}^m)$ defined as

$$(p')^t = p^{\sqrt{2}t} \circ i^t,$$

where $p^{\sqrt{2}t}$ denotes the maps of the Čech bundle filtration $(\mathbb{X}, \mathbb{p})$, and i^t denotes the inclusion $|R^t| \hookrightarrow \left|S^{\sqrt{2}t}\right|$. These maps fit in the following diagram:

$$\begin{array}{ccccc} |R^t| & \xrightarrow{i^t} & \left|S^{\sqrt{2}t}\right| & = & X^{\sqrt{2}t} \\ & \searrow^{(p')^t} & & & \downarrow{\scriptstyle p^{\sqrt{2}t}} \\ & & & & \mathcal{G}_d(\mathbb{R}^m) \end{array}$$

This new vector bundle filtration is defined on the index set $T' = \left[0, \frac{1}{\sqrt{2}} t_\gamma^{\max}(X)\right)$.

It is clear from the construction that the vector bundle filtrations $(\mathbb{X}, \mathbb{p})$ and $(|\mathbb{R}|, \mathbb{p}')$ are multiplicatively $\sqrt{2}$-interleaved, with an interleaving that preserves the persistent Stiefel-Whitney classes. This property is a multiplicative equivalent of Theorem V.7.

Remember that if X is a subset of $\mathbb{R}^n \times \mathcal{G}_d(\mathbb{R}^m)$, then the maximal filtration value of the Čech bundle filtration on X is $t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2}\gamma$ (see Equation (V.5)). Consequently, the maximal filtration value of its Vietoris-Rips version is $\frac{1}{2}\gamma$.

From an application perspective, we choose to work with the Vietoris-Rips filtration since it is easier to compute. Indeed, its construction only relies on computing pairwise distances, and finding cliques in graphs.

V.3.4 Choice of the parameter γ

This subsection is devoted to discussing the influence of the parameter $\gamma > 0$. Recall that γ affects the norm $\|\cdot\|_\gamma$ we chose on $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$:

$$\|(x, A)\|_\gamma^2 = \|x\|^2 + \gamma^2 \|A\|_{\mathrm{F}}^2 .$$

Let $X \subset \mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$. If $\gamma_1 \leq \gamma_2$ are two positive real numbers, the corresponding Čech filtrations $\mathbb{X}_1$ and $\mathbb{X}_2$, as well as the Čech bundle filtrations $(\mathbb{X}_1, \mathbb{p}_1)$ and $(\mathbb{X}_2, \mathbb{p}_2)$, are $\frac{\gamma_2}{\gamma_1}$-interleaved multiplicatively. This comes from the straightforward inequality

$$\|\cdot\|_{\gamma_1} \;\leq\; \|\cdot\|_{\gamma_2} \;\leq\; \frac{\gamma_2}{\gamma_1}\|\cdot\|_{\gamma_1}.$$

Note that we also have the additive inequality

$$\|(x,A)\|_{\gamma_1} \;\leq\; \|(x,A)\|_{\gamma_2} \;\leq\; \|(x,A)\|_{\gamma_1} + \sqrt{\gamma_2^2 - \gamma_1^2}\,\|A\|_{\mathrm{F}}\,.$$

One deduces that the Čech bundle filtrations $(\mathbb{X}_1, \mathbb{p}_1)$ and $(\mathbb{X}_2, \mathbb{p}_2)$ are $\sqrt{\gamma_2^2 - \gamma_1^2} \cdot t^{\max}(X)$-interleaved additively, where $t^{\max}(X)$ is the maximal filtration value when $\gamma = 1$. As a consequence of these interleavings, when the values γ_1 and γ_2 are close, the persistence barcodes and the lifebars of the persistent Stiefel-Whitney classes are close (see Theorem V.7).

As a general principle, one would choose the parameter γ to be large, since it would lead to large filtrations. More precisely, if $t^{\max}_{\gamma_1}(X)$ and $t^{\max}_{\gamma_2}(X)$ denote respectively the maximal filtration values of $(\mathbb{X}_1, \mathbb{p}_1)$ and $(\mathbb{X}_2, \mathbb{p}_2)$, then $t^{\max}_{\gamma_1}(X) = \gamma_1 \cdot t^{\max}(X)$ and $t^{\max}_{\gamma_2}(X) = \gamma_2 \cdot t^{\max}(X)$, as in Equation (V.4). Moreover, we have the following inclusion:

$$X_1^{t^{\max}_{\gamma_1}(X)} \subseteq X_2^{t^{\max}_{\gamma_2}(X)},$$

where $X_1^{t^{\max}_{\gamma_1}(X)}$ denotes the thickening of X with respect to the norm $\|\cdot\|_{\gamma_1}$, and $X_2^{t^{\max}_{\gamma_2}(X)}$ with respect to $\|\cdot\|_{\gamma_2}$. This inclusion can be proven from the following fact, valid for every $x \in \mathbb{R}^n$ and $A \in \mathrm{M}(\mathbb{R}^m)$ such that $\|A\|_{\mathrm{F}} \leq t^{\max}(X)$:

$$\|(x,A)\|_{\gamma_1} \leq t^{\max}_{\gamma_1}(X) \implies \|(x,A)\|_{\gamma_2} \leq t^{\max}_{\gamma_2}(X)\,.$$

Hence larger parameters γ lead to larger maximal filtration values and larger filtrations.

However, as we show in the following examples, different values of γ may result in different behaviours of the persistent Stiefel-Whitney classes. In Example V.25, large values of γ highlight properties of the dataset that are not consistent with the underlying vector bundle, which is orientable. Notice that, so far, we always picked the value $\gamma = 1$, for it seemed experimentally relevant with the datasets we chose.

Example V.24. Consider the set $Y \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ representing the Mobius band, as in Example V.5 of Subsection V.1.2:

$$Y = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we show in Appendix V.B.1, Y is a circle, included in a 2-dimensional affine subspace of $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$. Its radius is $\sqrt{1 + \frac{\gamma^2}{2}}$. As a consequence, the persistence of the Čech filtration of Y consists of two bars: one H^0-feature, the bar $[0, +\infty)$, and one H^1-feature, the bar $\left[0, \sqrt{1 + \frac{\gamma^2}{2}}\right)$.

For any $\gamma > 0$, the maximal filtration value of the Čech bundle filtration of Y is $t_\gamma^{\max}(Y) = \frac{\sqrt{2}}{2}\gamma$. Moreover, the persistent Stiefel-Whitney class $w_1^t(Y)$ is nonzero all along the filtration.

In this example, we see that the parameter γ does not influence the qualitative interpretation of the persistent Stiefel-Whitney class. It is always nonzero where it is defined. The following example shows a case where γ does influence the persistent Stiefel-Whitney class.

Example V.25. Consider the set $X \subset \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ representing the normal bundle of the circle $\mathbb{S}_1$, as in Example V.5:

$$X = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we show in Appendix V.B.2, X is a subset of a 2-dimensional flat torus embedded in $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, hence can be seen as a torus knot.

Before studying the Čech bundle filtration of X, we discuss the Čech filtration $\mathbb{X}$. Its behaviour depends on γ:

- if $\gamma \leq \frac{\sqrt{2}}{2}$, then X^t retracts on a circle for $t \in [0, 1)$, X^t retracts on a 3-sphere for $t \in \left[1, \sqrt{1 + \frac{1}{2}\gamma^2}\right)$, and X^t retracts on a point for $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$.
- if $\gamma \geq \frac{\sqrt{2}}{2}$, then X^t retracts on a circle for $t \in [0, 1)$, X^t retracts on another circle for $t \in \left[1, \frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}\right)$, X^t retracts on a 3-sphere for $t \in \left[\frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}, \sqrt{1 + \frac{1}{2}\gamma^2}\right)$, and X^t has the homotopy type of a point for $t \geq \sqrt{1 + \frac{1}{2}\gamma^2}$.

Let us interpret these facts. If $\gamma \leq \frac{\sqrt{2}}{2}$, then the persistent cohomology of X looks similar to the persistent cohomology of the underlying set $\left\{ \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \theta \in [0, 2\pi) \right\} \subset \mathbb{R}^2$, but with a H^3 cohomology feature added. Besides, if $\gamma \geq \frac{\sqrt{2}}{2}$, a new topological feature appears in the H^1-barcode: the bar $\left[1, \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}\right)$. These barcodes are depicted in Figures V.16 and V.17.

Let us now discuss the corresponding Čech bundle filtrations. For any $\gamma > 0$, the maximal filtration value of the Čech bundle filtration of X is $t_\gamma^{\max}(X) = \frac{\sqrt{2}}{2}\gamma$. We observe two behaviours:

- if $\gamma \leq \frac{\sqrt{2}}{2}$, then $w_1^t(X)$ is zero all along the filtration,
- if $\gamma > \frac{\sqrt{2}}{2}$, then $w_1^t(X)$ is nonzero from $t^\dagger = 1$.

This in proven in Appendix V.B.2. To conclude, this persistent Stiefel-Whitney class is consistent with the underlying bundle—the normal bundle of the circle, which is trivial—only for $t \leq 1$.

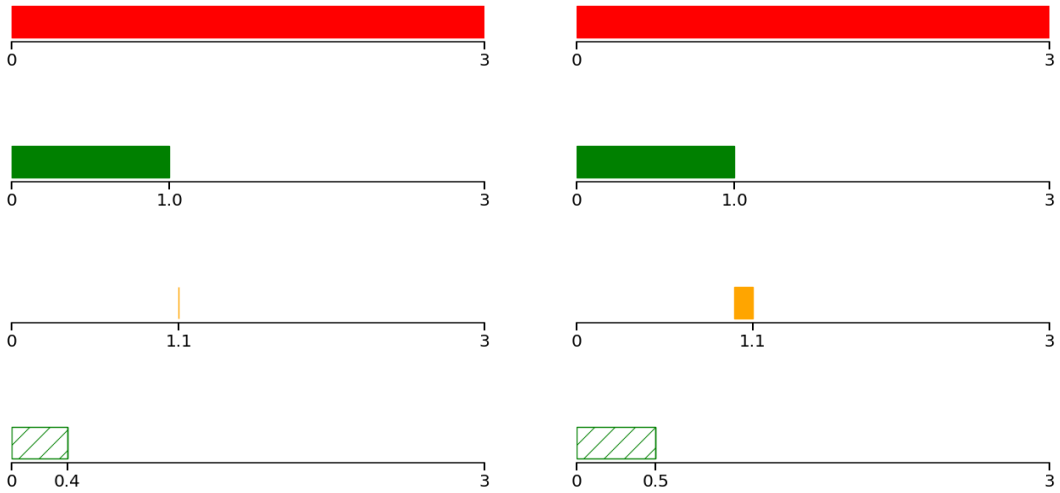


Figure V.16: H^0-, H^1-, H^3-barcodes and lifebar of the first persistent Stiefel-Whitney class of X with $\gamma = \frac{1}{2}$ **(left)** and $\gamma = \frac{\sqrt{2}}{2}$ **(right)**.

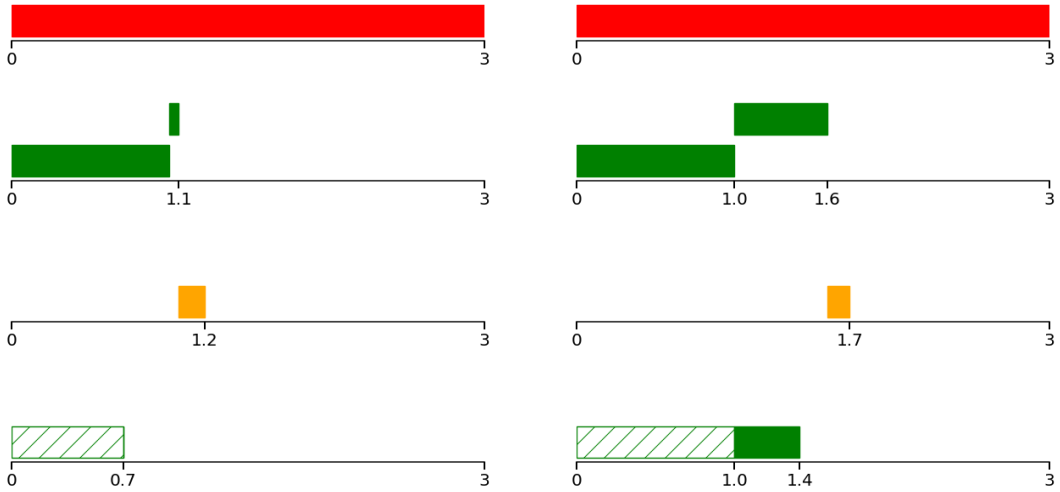


Figure V.17: H^0-, H^1-, H^3-barcodes and lifebar of the first persistent Stiefel-Whitney class of X with $\gamma = 1$ **(left)** and $\gamma = 2$ **(right)**.

V.4 Conclusion

In this chapter we defined the persistent Stiefel-Whitney classes of vector bundle filtrations. We proved that they are stable with respect to the interleaving distance between vector bundle filtrations. We studied the particular case of Čech bundle filtrations of subsets of $\mathbb{R}^n \times \mathrm{M}(\mathbb{R}^m)$, and showed that they yield consistent estimators of the usual Stiefel-Whitney classes of some underlying vector bundle. Moreover, when the dimension of the bundle is 1 and X is finite, we proposed an algorithm to compute the persistent Stiefel-Whitney classes.

Our algorithm is limited to the bundles of dimension 1, since we only implemented triangulations of the Grassmannian $\mathcal{G}_d(\mathbb{R}^m)$ when $d = 1$. However, any other triangulation of $\mathcal{G}_d(\mathbb{R}^m)$, with a computable face map, could be included in the algorithm without any modification. We also described a way to compute the lifebar of the persistent Stiefel-Whitney classes, by evaluating the class for several values of t, or by using mapping cones.

V.A Supplementary material for Section V.1

Proof of Lemma V.3 page 200. Note that $\mathcal{G}_d(\mathbb{R}^m)$ is contained in the linear subspace $\mathcal{S}$ of symmetric matrices. Therefore, to project a matrix $A \in \mathrm{M}(\mathbb{R}^m)$ onto $\mathcal{G}_d(\mathbb{R}^m)$, we may project on $\mathcal{S}$ first. It is well known that the projection of A onto $\mathcal{S}$ is the matrix $A^s = \frac{1}{2}(A + {}^tA)$.

Suppose now that we are given a symmetric matrix B. Let it be diagonalized as $B = OD{}^tO$. A projection of B onto $\mathcal{G}_d(\mathbb{R}^m)$ is a matrix P which minimizes the following quantity:

$$\min_{P \in \mathcal{G}_d(E)} \|B - P\|_{\mathrm{F}} . \tag{V.11}$$

This problem is equivalent to

$$\min_{P \in \mathcal{G}_d(E)} \|D - P\|_{\mathrm{F}}$$

via $P \mapsto {}^tOPO$. Now, let $e_1, \cdots, e_n$ denote the canonical basis of $\mathbb{R}^m$. We have

$$\begin{aligned} \|D - P\|_{\mathrm{F}}^2 &= \|D\|_{\mathrm{F}}^2 + \|P\|_{\mathrm{F}}^2 - 2\langle D, P\rangle_{\mathrm{F}} \\ &= \|D\|_{\mathrm{F}}^2 + \|P\|_{\mathrm{F}}^2 - 2\sum \langle \lambda_i e_i, P(e_i)\rangle , \end{aligned}$$

where $\langle \cdot, \cdot \rangle_{\mathrm{F}}$ is the Frobenius inner product, and $\langle \cdot, \cdot \rangle$ the usual inner product on $\mathbb{R}^m$. Therefore, Equation (V.11) is a problem equivalent to

$$\max_{P \in \mathcal{G}_d(E)} \sum \lambda_i \langle e_i, P(e_i)\rangle .$$

Since P is an orthogonal projection, we have $\langle e_i, P(e_i)\rangle = \langle P(e_i), P(e_i)\rangle = \|P(e_i)\|^2$ for all $i \in [1, n]$. Moreover, $d = \|P\|_{\mathrm{F}}^2 = \sum \|P(e_i)\|^2$. Denoting $p_i = \|P(e_i)\|^2 \in [0, 1]$, we finally obtain the following alternative formulation of Equation (V.11):

$$\max_{\substack{p_1, \ldots p_n \in [0,1] \\ p_1 + \ldots + p_n = d}} \sum \lambda_i p_i.$$

Using that $\lambda_1 \geq \ldots \geq \lambda_n$, we see that this maximum is attained when $p_0 = \ldots = p_d = 1$ and $p_{d+1} = \ldots = p_n = 0$. Consequently, a minimizer of Equation (V.11) is $P = J_d$, where J_d is the diagonal matrix whose first d terms are 1, and the other ones are zero. Moreover, it is unique if $\lambda_d \neq \lambda_{d+1}$.

As a consequence of these considerations, we obtain the following characterization: for every $B \in \mathrm{M}(\mathbb{R}^m)$,

$$B \in \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right) \iff \lambda_d(B^s) = \lambda_{d+1}(B^s). \tag{V.12}$$

Let us now show that for every matrix $A \in \mathrm{M}(\mathbb{R}^m)$, we have

$$\mathrm{dist}\left(A, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right) = \frac{\sqrt{2}}{2}\left|\lambda_d(A^s) - \lambda_{d+1}(A^s)\right|.$$

First, remark that

$$\mathrm{dist}\left(A, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right) = \mathrm{dist}\left(A^s, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right). \tag{V.13}$$

Indeed, if B is a projection of A on $\mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)$, then B^s is still in $\mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)$ according to Equation (V.12), and

$$\mathrm{dist}\left(A, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right) = \|A - B\|_{\mathrm{F}} \geq \|A^s - B^s\|_{\mathrm{F}} \geq \mathrm{dist}\left(A^s, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right).$$

Conversely, if B is a projection of A^s on $\mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)$, then $\hat{B} = B + A - A^s$ is still in $\mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)$, and

$$\mathrm{dist}\left(A, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right) \leq \left\|A - \hat{B}\right\|_{\mathrm{F}} = \|A^s - B\|_{\mathrm{F}} = \mathrm{dist}\left(A^s, \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)\right).$$

We deduce Equation (V.13).

Now, let $A \in \mathcal{S}$ and $B \in \mathrm{med}\left(\mathcal{G}_d(\mathbb{R}^m)\right)$. Let $e_1, ..., e_n$ be a basis of $\mathbb{R}^m$ that diagonalizes A. Writing $\|A - B\|_{\mathrm{F}} = \sum \|A(e_i) - B(e_i)\|^2 = \sum \|\lambda_i(A)e_i - B(e_i)\|^2$, it is clear that the closest matrix B must satisfy $B(e_i) = \lambda_i(B)e_i$, with

- $\lambda_i(B) = \lambda_i(A)$ for $i \notin \{d, d+1\}$,
- $\lambda_d(B) = \lambda_{d+1}(B) = \frac{1}{2}(\lambda_d(A) + \lambda_{d+1}(A))$.

We finally compute

$$\begin{aligned}\|A - B\|_{\mathrm{F}}^2 &= \sum \|\lambda_i(A)e_i - \lambda_i(B)e_i\|^2 \\ &= |\lambda_d(A) - \lambda_d(B)|^2 + |\lambda_{d+1}(A) - \lambda_{d+1}(B)|^2 \\ &= \frac{1}{2}\left|\lambda_d(A) - \lambda_{d+1}(A)\right|^2.\end{aligned}$$

V.B Supplementary material for Section V.3

V.B.1 Study of Example V.24

We consider the set

$$X = \left\{ \left(\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}, \begin{pmatrix}\cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2})\\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2\end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

To study the Čech filtration of X, we shall apply the following affine transformation: let Y be the subset of $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ defined as

$$Y = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \gamma \begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

and let $\mathbb{Y} = (Y^t)_{t\geq 0}$ be the Čech filtration of Y in $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ endowed with the usual norm $\|(x, A)\|_1 = \sqrt{\|x\|^2 + \|A\|_{\mathrm{F}}^2}$. We recall that the Čech filtration of X, denoted $\mathbb{X} = (X^t)_{t\geq 0}$, is defined with respect to the norm $\|\cdot\|_\gamma$. It is clear that, for every $t \geq 0$, the thickenings X^t and Y^t are homeomorphic via the application

$$\begin{aligned} h \colon \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2) &\longrightarrow \mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2) \\ (x, A) &\longmapsto (x, \gamma A). \end{aligned}$$

As a consequence, the persistence cohomology modules associated to $\mathbb{X}$ and $\mathbb{Y}$ are isomorphic.

Next, notice that Y is a subset of the affine subspace of dimension 2 of $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ with origin O and spanned by the vectors e_1 and e_2, where

$$O = \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right), \quad e_1 = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right), \quad e_2 = \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \frac{\gamma}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

Indeed, using the equality

$$\begin{pmatrix} \cos(\frac{\theta}{2})^2 & \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\sin(\frac{\theta}{2}) & \sin(\frac{\theta}{2})^2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix},$$

we obtain

$$Y = O + \{\cos(\theta)e_1 + \sin(\theta)e_2, \quad \theta \in [0, 2\pi)\}.$$

We see that Y is a circle, of radius $\|e_1\| = \|e_2\| = \sqrt{1 + \frac{\gamma^2}{2}}$.

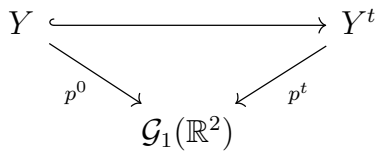
Let E denotes the affine space with origin O and spanned by the vectors e_1 and e_2. Lemma V.26, stated below, shows that the persistent cohomology of Y, seen in the ambient space $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, is the same as the persistent cohomology of Y restricted to the subspace E. As a consequence, Y has the same persistence as a circle of radius $\sqrt{1 + \frac{\gamma^2}{2}}$ in the plane. Hence its barcode can be described as follows:

- one H^0-feature: the bar $[0, +\infty)$,
- one H^1-feature: the bar $\left[0, \sqrt{1 + \frac{\gamma^2}{2}}\right)$.

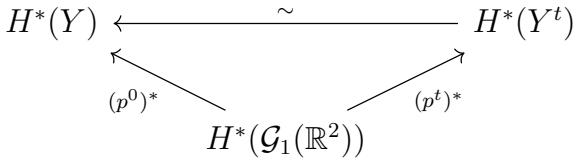
Lemma V.26. *Let $Y \subset \mathbb{R}^n$ be any subset, and define $\check{Y} = Y \times \{(0, ..., 0)\} \subset \mathbb{R}^n \times \mathbb{R}^m$. Let these spaces be endowed with the usual Euclidean norms. Then the Čech filtrations of Y and $\check{Y}$ yields isomorphic persistence modules.*

Proof. Let $\mathrm{proj}_n \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be the projection on the first n coordinates. One verifies that, for every $t \geq 0$, the map $\mathrm{proj}_n \colon \check{Y}^t \to Y^t$ is a homotopy equivalence. At cohomology level, these maps induce an isomorphism of persistence modules.

Let us now study the Čech bundle filtration of Y, denoted $(\mathbb{Y}, \mathbb{p})$. According to Equation (V.5), its filtration maximal value is $t^{\max}(Y) = t^{\max}_{\gamma}(X) = \frac{\gamma}{\sqrt{2}}$. Note that $\frac{\gamma}{\sqrt{2}}$ is lower than $\sqrt{1 + \frac{\gamma^2}{2}}$, which is the radius of the circle Y. Hence, for $t < t^{\max}(Y)$, the inclusion $Y \hookrightarrow Y^t$ is a homotopy equivalence. Consider the following commutative diagram:



It induces the following diagram in cohomology:



The horizontal arrow is an isomorphism. Hence the map $(p^t)^* \colon H^*(Y^t) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$ is equal to $(p^0)^*$. We only have to understand $(p^0)^*$.

Remark that the map $p^0 \colon Y \to \mathcal{G}_1(\mathbb{R}^2)$ can be seen as the universal bundle of the circle. Therefore $(p^0)^* \colon H^*(Y) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^m))$ is nontrivial. Alternatively, p^0 can be seen as a map between two circles. It is injective, hence its degree (modulo 2) is one. We still deduce that $(p^0)^*$ is nontrivial. As a consequence, the persistent Stiefel-Whitney class $w_1^t(X)$ is nonzero for every $t < t^{\max}(Y)$.

V.B.2 Study of Example V.25

We consider the set

$$X = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

As we explained in the previous subsection, the Čech filtration of X with respect to the norm $\|\cdot\|_\gamma$ yields the same persistence as the Čech filtration of Y with respect to the usual norm $\|\cdot\|$, where

$$Y = \left\{ \left(\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \gamma \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0, 2\pi) \right\}.$$

Notice that Y is a subset of the affine subspace of dimension 4 of $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$ with origin $O = \left(\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right), \frac{1}{2}\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right)$ and spanned by the vectors e_1, e_2, e_3 and e_4, where

$$e_1 = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0&0\\0&0\end{pmatrix}\right), \qquad e_2 = \left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0&0\\0&0\end{pmatrix}\right),$$
$$e_3 = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1&0\\0&-1\end{pmatrix}\right), \qquad e_4 = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0&1\\1&0\end{pmatrix}\right).$$

Indeed, Y can be written as

$$Y = O + \left\{\cos(\theta)e_1 + \sin(\theta)e_2 + \frac{\gamma}{\sqrt{2}}\cos(2\theta)e_3 + \frac{\gamma}{\sqrt{2}}\sin(2\theta)e_4, \quad \theta \in [0, 2\pi)\right\}.$$

This comes from the equality

$$\begin{pmatrix}\cos(\theta)^2 & \cos(\theta)\sin(\theta)\\ \cos(\theta)\sin(\theta) & \sin(\theta)^2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}.$$

Observe that Y is a torus knot, i.e. a simple closed curve included in the torus $\mathbb{T}$, defined as

$$\mathbb{T} = O + \left\{\cos(\theta)e_1 + \sin(\theta)e_2 + \frac{\gamma}{\sqrt{2}}\cos(\nu)e_3 + \frac{\gamma}{\sqrt{2}}\sin(\nu)e_4, \quad \theta, \nu \in [0, 2\pi)\right\}.$$

The curve Y winds one time around the first circle of the torus, and two times around the second one. It is known as the torus knot $(1, 2)$.

Let E denotes the affine subspace with origin O and spanned by e_1, e_2, e_3, e_4. Since Y is a subset of E, it is equivalent to study the Čech filtration of Y restricted to E (as in Lemma V.26). We shall denote the coordinates of points $x \in E$ with respect to the orthonormal basis (e_1, e_2, e_3, e_4). That is, a tuple (x_1, x_2, x_3, x_4) shall refer to the point $O + x_1e_1 + x_2e_2 + x_3e_3 + x_4e_4$ of E. Seen in E, the set Y can be written as

$$Y = \left\{\left(\cos(\theta), \sin(\theta), \frac{\gamma}{\sqrt{2}}\cos(2\theta), \frac{\gamma}{\sqrt{2}}\sin(2\theta)\right), \theta \in [0, 2\pi)\right\}.$$

For every $\theta \in [0, 2\pi)$, we shall denote $y_\theta = \left(\cos(\theta), \sin(\theta), \frac{\gamma}{\sqrt{2}}\cos(2\theta), \frac{\gamma}{\sqrt{2}}\sin(2\theta)\right)$.

Figure V.18: Representations of the set Y, lying on a torus, for a small value of γ **(left)** and a large value of γ **(right)**.

We now state two lemmas that will be useful in what follows.

Lemma V.27. *For every $\theta \in [0, 2\pi)$, the map $\theta' \mapsto \|y_\theta - y_{\theta'}\|$ admits the following critical points:*

- $\theta' - \theta = 0$ *and* $\theta' - \theta = \pi$ *if* $\gamma \leq \frac{1}{\sqrt{2}}$,
- $\theta' - \theta = 0$, π, $\arccos(-\frac{1}{2\gamma^2})$ *and* $-\arccos(-\frac{1}{2\gamma^2})$ *if* $\gamma \geq \frac{1}{\sqrt{2}}$.

They correspond to the values

- $\|y_\theta - y_{\theta'}\| = 0$ *if* $\theta' - \theta = 0$,
- $\|y_\theta - y_{\theta'}\| = 2$ *if* $\theta' - \theta = \pi$,
- $\|y_\theta - y_{\theta'}\| = \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$ *if* $\theta' - \theta = \pm\arccos(-\frac{1}{2\gamma^2})$.

Moreover, we have $\sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}} \geq 2$ when $\gamma \geq \frac{1}{\sqrt{2}}$.

Proof. Let $\theta, \theta' \in [0, 2\pi)$. One computes that

$$\|y_\theta - y_{\theta'}\|^2 = 4\sin^2\left(\frac{\theta - \theta'}{2}\right) + 2\gamma^2\sin^2(\theta - \theta').$$

Consider the map $f\colon x \in [0, 2\pi) \mapsto 4\sin^2\left(\frac{x}{2}\right) + 2\gamma^2\sin^2(x)$. Its derivative is

$$\begin{aligned} f'(x) &= 4\cos\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right) + 4\gamma^2\cos(x)\sin(x) \\ &= 2\sin(x)\left(1 + 2\gamma^2\cos(x)\right). \end{aligned}$$

It vanishes when $x = 0$, $x = \pi$, or $x = \pm\arccos(-\frac{1}{2\gamma^2})$ if $\gamma \geq \frac{1}{\sqrt{2}}$. A computation shows that $f(0) = 0$, $f(\pi) = 4$ and $f\left(\pm\arccos\left(-\frac{1}{2\gamma^2}\right)\right) = 2\left(1 + \gamma^2 + \frac{1}{4\gamma^2}\right)$.

Lemma V.28. *For every $x \in E$ such that $x \neq 0$, the map $\theta \mapsto \|x - y_\theta\|$ admits at most two local maxima and two local minima.*

Proof. Consider the map $g\colon \theta \in [0, 2\pi) \mapsto \|x - y_\theta\|^2$. It can be written as

$$\begin{aligned} g(\theta) &= \|x\|^2 + \|y_\theta\|^2 - 2\langle x, y_\theta\rangle \\ &= \|x\|^2 + 1 + \frac{\gamma^2}{2} - 2\langle x, y_\theta\rangle. \end{aligned}$$

Let us show that its derivative g' vanishes at most four times on $[0, 2\pi)$, which would show the result. Using the expression of y_θ, we see that g' can be written as

$$g'(\theta) = a\cos(\theta) + b\sin(\theta) + c\cos(2\theta) + d\sin(2\theta),$$

where $a, b, c, d \in \mathbb{R}$ are not all zero. Denoting $\omega = \cos(\theta)$ and $\xi = \sin(\theta)$, we have $\xi^2 = 1 - \omega^2$, $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 2\omega^2 - 1$ and $\sin(2\theta) = 2\cos(\theta)\sin(\theta) = 2\omega\xi$. Hence

$$g'(\theta) = a\omega + b\xi + 2c\omega^2 + 2d\omega\xi.$$

Now, if $g'(\theta) = 0$, we get

$$a\omega + 2c\omega^2 = -(b + 2d\omega)\xi \tag{V.14}$$

Squaring this equality yields $(a\omega + 2c\omega^2)^2 = (b + 2d\omega)^2(1 - \omega^2)$. This degree four equation, with variable ω, admits at most four roots. To each of these w, there exists a unique $\xi = \pm\sqrt{1 - w^2}$ that satisfies Equation (V.14). In other words, the corresponding $\theta \in [0, 2\pi)$ such that $\omega = \cos(\theta)$ is unique. We deduce the result.

Before studying the Čech filtration of Y, let us describe some geometric quantities associated to it. Using a symbolic computation software, we see that the curvature of Y is constant and equal to

$$\rho = \frac{\sqrt{1 + 8\gamma^2}}{1 + 2\gamma^2}.$$

In particular, we have $\rho \geq 1$ if $\gamma \leq 1$, and $\rho < 1$ if $\gamma > 1$. We also have an expression for the diameter of Y:

$$\frac{1}{2}\mathrm{diam}\,(Y) = \begin{cases} 1 & \text{if } \gamma \leq \frac{1}{\sqrt{2}}, \\ \frac{1}{\sqrt{2}}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}} & \text{if } \gamma \geq \frac{1}{\sqrt{2}}. \end{cases}$$

It is a consequence of Lemma V.27. We now describe the reach of Y:

$$\mathrm{reach}\,(Y) = \begin{cases} \frac{1+2\gamma^2}{\sqrt{1+8\gamma^2}} & \text{if } \gamma \leq 1, \\ 1 & \text{if } \gamma \geq 1. \end{cases} \tag{V.15}$$

Let us prove this by using the results of Subsection II.3.3. We define a bottleneck of Y as pair of distinct points $(y, y') \in Y^2$ such that the open ball $\mathcal{B}(\frac{1}{2}(y + y'), \frac{1}{2}\|y - y'\|)$ does not intersect Y. Its length is defined as $\frac{1}{2}\|y - y'\|$. According to Theorem II.13, the reach of Y is equal to

$$\mathrm{reach}\,(Y) = \min\left\{\frac{1}{\rho}, \delta\right\},$$

where $\frac{1}{\rho}$ is the inverse curvature of Y, and δ is the minimal length of bottlenecks of Y. As we computed, $\frac{1}{\rho}$ is equal to $\frac{1+2\gamma^2}{\sqrt{1+8\gamma^2}}$. Besides, according to Lemma V.27, a bottleneck $(y_\theta, y_{\theta'})$ has to satisfy $\theta' - \theta = \pi$ or $\pm\arccos(-\frac{1}{2\gamma^2})$. The smallest length is attained when $\theta' - \theta = \pi$, for which $\frac{1}{2}\|y_\theta - y_{\theta'}\| = 1$. It is straightforward to verify that the pair $(y_\theta, y_{\theta'})$ is indeed a bottleneck. Therefore we have $\delta = 1$, and we deduce the expression of $\mathrm{reach}\,(Y)$.

Last, the weak feature size of Y does not depend on γ and is equal to 1:

$$\mathrm{wfs}\,(Y) = 1. \tag{V.16}$$

We shall prove it by using the characterization of Subsection II.3.4: $\mathrm{wfs}\,(Y)$ is the infimum of distances $\mathrm{dist}\,(x, Y)$, where $x \in E$ is a critical point of the distance function d_Y. In this context, x is a critical point if it lies in the convex hull of its projections on Y. Remark that, if $x \neq 0$, then x admits at most two projections on Y. This follows from Lemma V.28. As a consequence, if x is a critical point, then there exists $y, y' \in Y$ such that x lies in the middle of the segment $[y, y']$, and the open ball $\mathcal{B}(x, \mathrm{dist}\,(x, Y))$ does not intersect Y. Therefore y' is a critical point of $y' \mapsto \|y - y'\|$, hence Lemma V.27 gives that $\|y - y'\| \geq 2$. We deduce the result.

We now describe the thickenings Y^t. They present four different behaviours:

- $0 \leq t < 1$: Y^t is homotopy equivalent to a circle,
- $1 \leq t < \frac{1}{2}\mathrm{diam}\,(Y)$: Y^t is homotopy equivalent to a circle,
- $\frac{1}{2}\mathrm{diam}\,(Y) \leq t < \sqrt{1 + \frac{\gamma^2}{2}}$: Y^t is homotopy equivalent to a 3-sphere,
- $t \geq \sqrt{1 + \frac{\gamma^2}{2}}$: Y^t is homotopy equivalent to a point.

Recall that, in the case where $\gamma \leq \frac{1}{\sqrt{2}}$, we have $\frac{1}{2}\mathrm{diam}\,(Y) = 1$. Consequently, the interval $\left[1, \frac{1}{2}\mathrm{diam}\,(Y)\right)$ is empty, and the second point does not appear in this case.

Study of the case $0 \leq t < 1$. For $t \in [0, 1)$, let us show that Y^t deform retracts on Y. According to Equation (V.16), we have $\mathrm{wfs}\,(Y) = 1$. Moreover, Equation (V.15) gives that $\mathrm{reach}\,(Y) > 0$. Using the results of Subsection II.3.4, we deduce that Y^t is isotopic to Y.

Study of the case $1 \leq t < \frac{1}{2}\mathrm{diam}\,(Y)$. Denote $z_\theta = \left(0, 0, \frac{\gamma}{\sqrt{2}}\cos(2\theta), \frac{\gamma}{\sqrt{2}}\sin(2\theta)\right)$, and define the circle $Z = \{z_\theta, \theta \in [0, \pi)\}$.

Figure V.19: Representation of the set Y (black) and the circle Z (red).

We claim that Y^t deform retracts on Z. To prove so, we shall define a continuous application $f\colon Y^t \to Z$ such that, for every $y \in Y^t$, the segment $[y, f(y)]$ is included in Y^t. This would lead to a deformation retraction of Y^t onto Z, via

$$(s, y) \in [0, 1] \times Y^t \mapsto (1 - s)y + sf(y).$$

Equivalently, we shall define an application $\Theta\colon Y^t \to [0, \pi)$ such that the segment $[y, z_{\Theta(y)}]$ is included in Y^t.

Let $y \in Y^t$. According to Lemma V.28, y admits at most two projection on Y. We start with the case where y admits only one projection, namely y_θ with $\theta \in [0, 2\pi)$. Let $\overline{\theta} \in [0, \pi)$ be the reduction of θ modulo π, and consider the point $z_{\overline{\theta}}$ of Z. A computation shows that the distance $\|y_\theta - z_{\overline{\theta}}\|$ is equal to 1. Besides, since $y \in Y^t$, the distance $\|y_\theta - y\|$ is at most t. By convexity, the segment $[y, z_{\overline{\theta}}]$ is included in the ball $\overline{\mathcal{B}}(y_\theta, t)$, which is a subset of Y^t. We then define $\Theta(y) = \overline{\theta}$.

Now suppose that y admits exactly two projection y_θ and $y_{\theta'}$. According to Lemma V.27, these angles must satisfy $\theta' - \theta = \pi$. Indeed, the case $\|y_\theta - y_{\theta'}\| = \sqrt{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$ does not occur since we chose $t < \frac{1}{2}\mathrm{diam}\,(Y) = \frac{\sqrt{2}}{2}\sqrt{1 + \gamma^2 + \frac{1}{4\gamma^2}}$. The angles θ and θ' correspond to the same reduction modulo π, denoted $\overline{\theta}$, and we also define $\Theta(y) = \overline{\theta}$.

Study of the case $t \in \left[\frac{1}{2}\mathrm{diam}\,(X), \sqrt{1 + \frac{\gamma^2}{2}}\right)$. Let $\mathbb{S}_3$ denotes the unit sphere of E. For every $v = (v_1, v_2, v_3, v_4) \in \mathbb{S}_3$, we shall denote by $\langle v \rangle$ the linear subspace spanned by v, and by $\langle v \rangle_+$ the cone $\{\lambda v, \lambda \geq 0\}$. Moreover, we define the quantity

$$\delta(v) = \min_{y \in Y} \mathrm{dist}\,(y, \langle v \rangle_+)\,.$$

and the set

$$S = \{\delta(v)v, v \in \mathbb{S}_3\}\,.$$

We claim that S is a subset of Y^t, and that Y^t deform retracts on it. This follows from the two following facts: for every $v \in \mathbb{S}_3$,

1. $\delta(v)$ is not greater than $\frac{1}{2}\mathrm{diam}\,(Y)$,
2. $\langle v \rangle_+ \cap Y^t$ consists of one connected component: an interval centered on $\delta(v)v$, that does not contain the point 0.

Suppose that these assertions are true. Then one defines a deformation retraction of Y^t on S by retracting each fiber $\langle v \rangle_+ \cap Y^t$ linearly on the singleton $\{\delta(v)v\}$. We shall now prove the two items.

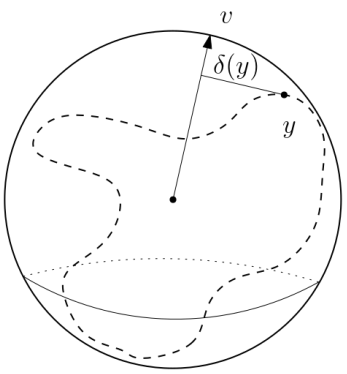


Figure V.20: Representation of the set Y (dashed), lying on a 3-sphere of radius $\sqrt{1 + \frac{\gamma^2}{2}}$.

Item 1.
Note that Item 1 can be reformulated as follows:

$$\max_{v\in\mathbb{S}_3}\min_{y\in Y}\operatorname{dist}(y,\langle v\rangle_+)\leq\frac{1}{2}\operatorname{diam}(Y)\,. \tag{V.17}$$

Let us justify that the pairs (v,y) that attain this maximum-minimum are the same as in

$$\max_{v\in\mathbb{S}_3}\min_{y\in Y}\|y-v\|\,. \tag{V.18}$$

From the definition of $Y=\{y_\theta,\theta\in[0,2\pi)\}$, we see that $\min_{y\in Y}\operatorname{dist}(y,\langle v\rangle_+)=\min_{y\in Y}\operatorname{dist}(y,\langle v\rangle)$. A vector $v\in\mathbb{S}_3$ being fixed, let us show that $y\mapsto\operatorname{dist}(y,\langle v\rangle)$ is minimized when $y\mapsto\|v-y\|$ is. Let $y\in Y$. Since v is a unit vector, the projection of y on $\langle v\rangle$ can be written as $\langle y,v\rangle\,v$. Hence $\operatorname{dist}(y,\langle v\rangle)^2=\|\langle y,v\rangle\,v-y\|^2$, and expanding this norm yields

$$\operatorname{dist}(y,\langle v\rangle)^2=\|y\|^2-\langle y,v\rangle^2\,.$$

Expanding the norm $\|y-v\|^2$ and using that $\|y\|^2=1+\frac{\gamma^2}{2}$, we get $\langle y,v\rangle=\frac{1}{2}\left(2+\frac{\gamma^2}{2}-\|y-v\|^2\right)$. We inject this relation in the preceding equation to obtain

$$\operatorname{dist}(y,\langle v\rangle)^2=-\left(\frac{\gamma}{2}\right)^4+\gamma^2+\frac{1}{4}\|y-v\|^2\left(4+\gamma^2-\|y-v\|^2\right).$$

Now we can deduce that $y\mapsto\operatorname{dist}(y,\langle v\rangle)^2$ is minimized when $y\mapsto\|y-v\|$ is minimized. Indeed, the map $\|y-v\|\mapsto\frac{1}{4}\|y-v\|^2\left(4+\gamma^2-\|y-v\|^2\right)$ is increasing on $\left[0,\frac{1}{2}(4+\gamma^2)\right]$. But $\|y-v\|\leq\|y\|+\|v\|=\frac{1}{2}(4+\gamma^2)$.

We deduce that studying the left hand term of Equation (V.17) is equivalent to studying Equation (V.18). We shall denote by $g\colon\mathbb{S}_3\to\mathbb{R}$ the map

$$g(v)=\min_{y\in Y}\|y-v\|\,. \tag{V.19}$$

Let $v\in\mathbb{S}_3$ that attains the maximum of g, and let y be a corresponding point that attains the minimum of $\|y-v\|$. The points v and y attains the quantity in Equation (V.17). In order to prove that $\operatorname{dist}(y,\langle v\rangle)\leq\frac{1}{2}\operatorname{diam}(Y)$, let $p(y)$ denotes the projection of y on $\langle v\rangle$. We shall show that there exists another point $y'\in Y$ such that $p(y)$ is equal to $\frac{1}{2}(y+y')$ Consequently, we would have $\|y-p(y)\|=\frac{1}{2}\|y'-y\|\leq\frac{1}{2}\operatorname{diam}(Y)$, i.e.

$$\operatorname{dist}(y,\langle v\rangle)\leq\frac{1}{2}\operatorname{diam}(Y)\,.$$

Remark the following fact: if $w\in\mathbb{S}_3$ is a unit vector such that $\langle p(y)-y,w\rangle>0$, then for $\epsilon>0$ small enough, we have

$$\operatorname{dist}(y,\langle v+\epsilon w\rangle)>\operatorname{dist}(y,\langle v\rangle)\,.$$

Equivalently, this statement reformulates as $0\leq\left\langle y,\frac{1}{\|v+\epsilon w\|}(v+\epsilon w)\right\rangle<\langle y,v\rangle$. Let us show that

$$\left\langle y,\frac{1}{\|v+\epsilon w\|}(v+\epsilon w)\right\rangle=\langle y,v\rangle-\epsilon\kappa+o(\epsilon), \tag{V.20}$$

where $\kappa = \langle p(y) - y, w\rangle > 0$, and where $o(\epsilon)$ is the little-o notation. Note that $\frac{1}{\|v+\epsilon w\|} = 1 - \epsilon \langle v, w\rangle + o(\epsilon)$. We also have

$$\frac{1}{\|v+\epsilon w\|}(v+\epsilon w) = v + \epsilon\,(w - \langle v, w\rangle\, v) + o(\epsilon).$$

Expanding the inner product in Equation (V.20) gives

$$\begin{aligned}\left\langle y, \frac{1}{\|v+\epsilon w\|}(v+\epsilon w)\right\rangle &= \langle y, v\rangle + \epsilon\Big(\ \langle y, w\rangle - \langle v, w\rangle\,\langle y, v\rangle\ \Big) + o(\epsilon)\\ &= \langle y, v\rangle + \epsilon\left\langle y - \langle y, v\rangle\, v, w\right\rangle + o(\epsilon)\\ &= \langle y, v\rangle + \epsilon\,\langle y - p(y), w\rangle + o(\epsilon),\end{aligned}$$

and we obtain the result.

Next, let us prove that y is not the only point of Y that attains the minimum in Equation (V.19). Suppose that it is the case by contradiction. Let $w \in \mathbb{S}_3$ be a unit vector such that $\langle p(y) - y, w\rangle > 0$. For ϵ small enough, let us prove that the vector $v' = \frac{1}{\|v+\epsilon w\|}(v+\epsilon w)$ of $\mathbb{S}_3$ contradicts the maximality of v. That is, let us prove that $g(v') > g(v)$. Let $y' \in Y$ be a minimizer $\|y' - v'\|$. We have to show that $\|y' - v'\| > \|y - v\|$. This would lead to $g(v') > g(v)$, hence the contradiction.

Expanding the norm yields

$$\|v' - y'\|^2 = \|v' - v + v - y'\|^2 \geq \|v' - v\|^2 + \|v - y'\|^2 - 2\,\langle v' - v, v - y'\rangle\,.$$

Using $\|v' - v\|^2 \geq 0$ and $\|v - y'\|^2 \geq \|v - y\|^2$ by definition of y, we obtain

$$\|v' - y'\|^2 \geq \|v - y\|^2 - 2\,\langle v' - v, v - y'\rangle\,.$$

We have to show that $\langle v' - v, y - y'\rangle$ is positive for ϵ small enough. By writing $v - y' = v - y + (y - y')$ we get

$$\langle v' - v, v - y'\rangle = \langle v' - v, v\rangle - \langle v' - v, y\rangle + \langle v' - v, y - y'\rangle$$

According to Equation (V.20), $-\langle v' - v, y\rangle = \epsilon\kappa + o(\epsilon)$. Besides, using $v' - v = \epsilon(w - \langle v, w\rangle\, v) + o(\epsilon)$, we get $\langle v' - v, v\rangle = o(\epsilon)$. Last, Cauchy-Schwarz inequality gives $|\,\langle v' - v, y - y'\rangle\,| \leq \|v' - v\|\,\|y - y'\|$. Therefore, $\langle v' - v, y - y'\rangle = O(\epsilon)\,\|y - y'\|$, where $O(\epsilon)$ is the big-o notation. Gathering these three equalities, we obtain

$$\langle v' - v, v - y'\rangle = o(\epsilon) + \epsilon\kappa + O(\epsilon)\,\|y - y'\|\,.$$

As we can read from this equation, if $\|y - y'\|$ goes to zero as ϵ does, then $\langle v' - v, v - y'\rangle$ is positive for ϵ small enough. Observe that v' goes to v when ϵ goes to 0. By assumption y is the only minimizer in Equation (V.19). By continuity of g, we deduce that y' goes to y.

By contradiction, we deduce that there exists another point y' which attains the minimum in $g(v)$. Note that it is the only other one, according to Lemma V.28. Let us show that $p(y)$ lies in the middle of the segment $[y, y']$. Suppose that it is not the case. Then $p(y) - y$ is not equal to $-(p(y') - y')$, where $p(y')$ denotes the

projection of y' on $\langle v \rangle$. Consequently, the half-spaces $\{w \in E, \langle p(y) - y, w \rangle > 0\}$ and $\{w \in E, \langle p(y') - y', w \rangle > 0\}$ intersects. Let w be any vector in the intersection. For $\epsilon > 0$, denote $v' = \frac{1}{\|v+\epsilon w\|}(1 + \epsilon w)$. If ϵ is small enough, the same reasoning as before shows that v' contradicts the maximality of v.

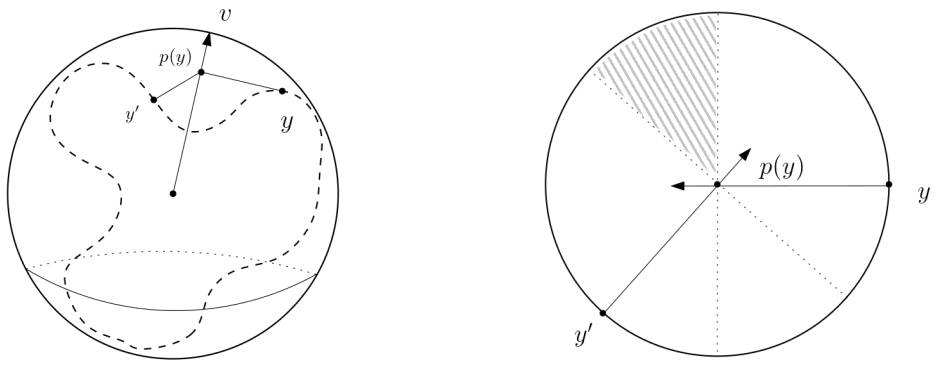


Figure V.21: Left: Representation of the situation where y and y' are minimizers of Equation (V.19). **Right:** Representation in the plane passing through the points y, y' and $p(y)$. The dashed area corresponds to the intersection of the half-spaces $\{w \in E, \langle p(y) - y, w \rangle > 0\}$ and $\{w \in E, \langle p(y') - y', w \rangle > 0\}$.

Item 2.
Let $v \in \mathbb{S}_3$. The set $\langle v \rangle_+ \cap Y^t$ can be described as

$$\langle v \rangle_+ \cap \bigcup_{y \in Y} \overline{\mathcal{B}}(y, t).$$

Let $y \in Y$ such that $\langle v \rangle_+ \cap \overline{\mathcal{B}}(y, t) \neq \emptyset$. Denote by $p(y)$ the projection of y on $\langle v \rangle_+$. It is equal to $\langle y, v \rangle v$. Using Pythagoras' theorem, we obtain that the set $\langle v \rangle_+ \cap \overline{\mathcal{B}}(y, t)$ is equal to the interval

$$\left[p(y) \pm \sqrt{t^2 - \operatorname{dist}(y, \langle v \rangle)^2} v \right].$$

Using the identity $\operatorname{dist}(y, \langle v \rangle)^2 = \|y\| - \langle y, v \rangle^2 = 1 + \frac{\gamma^2}{2} - \langle y, v \rangle^2$, we can write this interval as

$$\left[I_1(y) \cdot v,\ I_2(y) \cdot v \right],$$

where $I_1(y) = \langle y, v \rangle - \sqrt{\langle y, v \rangle^2 - (1 + \frac{\gamma^2}{2} - t^2)}$ and $I_2(y) = \langle y, v \rangle + \sqrt{\langle y, v \rangle^2 - (1 + \frac{\gamma^2}{2} - t^2)}$. Seen as functions of $\langle y, v \rangle$, the map I_1 is decreasing, and the map I_2 is increasing (see Figure V.22). Let $y^* \in Y$ that minimizes $\operatorname{dist}(y, \langle v \rangle)$. Equivalently, y^* maximizes $\langle y, v \rangle$. It follows that the corresponding interval $\left[I_1(y^*) \cdot v,\ I_2(y^*) \cdot v \right]$ contains all the others. We deduce that the set $\langle v \rangle_+ \cap Y^t$ is equal to this interval.

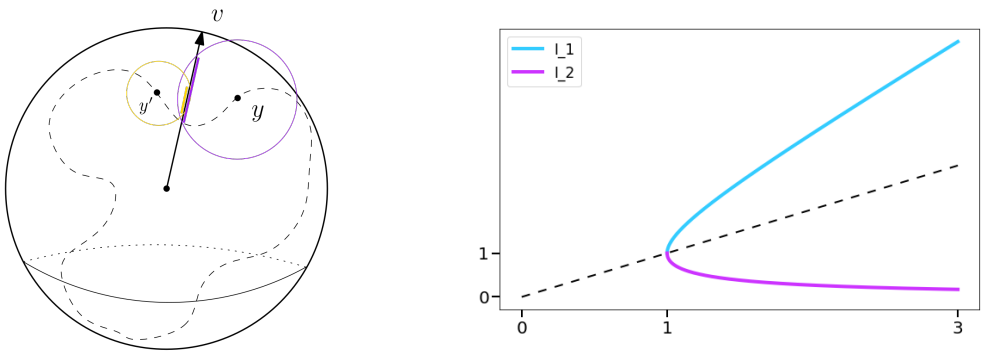


Figure V.22: Left: Representation of two intervals $\langle v\rangle_+ \cup \overline{\mathcal{B}}(y,t)$ and $\langle v\rangle_+ \cup \overline{\mathcal{B}}(y',t)$. **Right:** Representation of the maps $x \mapsto x \pm \sqrt{x^2-1}$.

Study of the case $t \geq \sqrt{1+\frac{1}{2}\gamma^2}$. For every $y \in Y$, we have $\|y\| = \sqrt{1+\frac{1}{2}\gamma^2}$. Therefore, if $t \geq \sqrt{1+\frac{1}{2}\gamma^2}$, then Y^t is star shaped around the point 0, hence it deform retracts on it.

Čech bundle filtration of Y. To close this subsection, let us study the Čech bundle filtration $(\mathbb{Y}, \mathbb{p})$ of Y. According to Equation (V.5), its filtration maximal value is $t^{\max}(Y) = t^{\max}_{\gamma}(X) = \frac{\gamma}{\sqrt{2}}$. Note that $\frac{\gamma}{\sqrt{2}}$ is lower than $\frac{1}{2}\mathrm{diam}(Y)$. Consequently, only two cases are to be studied: $t \in [0,1)$, and $t \in \left[1, \frac{1}{2}\mathrm{diam}(Y)\right)$.

The same argument as in Subsection V.B.2 yields that for every $t \in [0,1)$, the persistent Stiefel-Whitney class $w_1^t(Y)$ is equal to $w_1^0(Y)$. Accordingly, for every $t \in \left[1, \frac{1}{2}\mathrm{diam}(Y)\right)$, the class $w_1^t(Y)$ is equal to $w_1^1(Y)$. Let us show that $w_1^0(Y)$ is zero, and that $w_1^1(Y)$ is not.

First, remark that the map $p^0 \colon Y \to \mathcal{G}_1(\mathbb{R}^2)$ can be seen as the normal bundle of the circle. Hence $(p^0)^* \colon H^*(Y) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$ is nontrivial, and we deduce that $w_1^0(Y) = 0$. As a consequence, the persistent Stiefel-Whitney class $w_1^t(X)$ is nonzero for every $t < 1$.

Next, consider $p^1 \colon Y^1 \to \mathcal{G}_1(\mathbb{R}^2)$. Recall that Y^1 deform retracts on the circle

$$Z = \left\{ \left(0, 0, \frac{\gamma}{\sqrt{2}}\cos(2\theta), \frac{\gamma}{\sqrt{2}}\sin(2\theta)\right), \theta \in [0,\pi) \right\}.$$

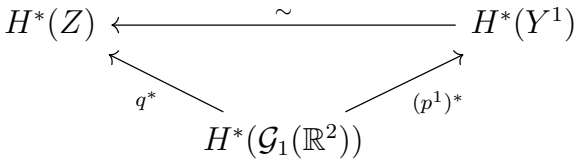
Seen in $\mathbb{R}^2 \times \mathrm{M}(\mathbb{R}^2)$, we have

$$Z = \left\{ \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \gamma \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \right), \theta \in [0,\pi) \right\}.$$

Notice that the map $q \colon Z \to \mathcal{G}_1(\mathbb{R}^2)$, the projection on $\mathcal{G}_1(\mathbb{R}^2)$, is injective. Seen as a map between two circles, it has degree (modulo 2) equal to 1. We deduce that $q^* \colon H^*(Z) \leftarrow H^*(\mathcal{G}_1(\mathbb{R}^2))$ is nontrivial. Now, remark that the map q factorizes through p^1:

$$\begin{array}{ccc} Z & \hookrightarrow & Y^1 \\ & {\scriptstyle q}\searrow \quad \swarrow{\scriptstyle p^1} & \\ & \mathcal{G}_1(\mathbb{R}^2) & \end{array}$$

It induces the following diagram in cohomology:



Since q^* is nontrivial, this commutative diagram yields that the persistent Stiefel-Whitney class $w_1^1(Y)$ is nonzero. As a consequence, the persistent Stiefel-Whitney class $w_1^t(Y)$ is nonzero for every $t \geq 1$.

Bibliography

[Aam18] Eddie Aamari. *Vitesses de convergence en inférence géométrique*. PhD thesis, Université Paris-Saclay, 2018. https://tel.archives-ouvertes.fr/tel-01607782.

[AB06] Stephanie B Alexander and Richard L Bishop. Gauss equation and injectivity radii for subspaces in spaces of curvature bounded above. *Geometriae Dedicata*, 117(1):65–84, 2006. https://arxiv.org/abs/math/0511570.

[AB18] Catherine Aaron and Olivier Bodart. Convergence rates for estimators of geodesic distances and Fréchet expectations. *Journal of Applied Probability*, 55(4):1001–1013, 2018. https://www.cambridge.org/core/journals/journal-of-applied-probability/article/convergence-rates-for-estimators-of-geodesic-distances-and-frechet-expectations/B52D9E022580EDF9041C7235C41315A4.

[ACG+18] Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. DTM-based filtrations. In *Symposium on Computational Geometry*, 2018. https://drops.dagstuhl.de/opus/volltexte/2019/10462/.

[ACG+20] Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. DTM-based filtrations. In *Topological Data Analysis*, pages 33–66. Springer, 2020. https://link.springer.com/chapter/10.1007/978-3-030-43408-3_2.

[ACLZ17] Ery Arias-Castro, Gilad Lerman, and Teng Zhang. Spectral clustering based on local PCA. *The Journal of Machine Learning Research*, 18(1):253–309, 2017. http://www.jmlr.org/papers/volume18/14-318/14-318.pdf.

[AKC+19] Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the Reach of a Manifold. *Electronic journal of statistics*, 2019. https://arxiv.org/abs/1705.04565.

[AL19] Eddie Aamari and Clément Levrard. Nonasymptotic rates for manifold, tangent space and curvature estimation. *The Annals of Statistics*, 47(1):177–204, 2019. https://arxiv.org/abs/1705.00989.

[ARC14] Aaron Adcock, Daniel Rubin, and Gunnar Carlsson. Classification of hepatic lesions using the matching metric. *Computer vision and image understanding*, 121:36–42, 2014. https://arxiv.org/abs/1210.0866.

[Aub11] HB Aubrey. *Persistent cohomology operations*. PhD thesis, Duke University, 2011. https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/3867/HB_duke_0066D_10919.pdf.

[Azu50] Gorô Azumaya. Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem. *Nagoya Mathematical Journal*, 1:117–124, 1950.

[BCB20] Magnus Botnan and William Crawley-Boevey. Decomposition of persistence modules. *Proceedings of the American Mathematical Society*, 148(11):4581–4596, 2020. https://arxiv.org/abs/1811.08946.

[BCOS16] Mickaël Buchet, Frédéric Chazal, Steve Y Oudot, and Donald R Sheehy. Efficient and robust persistent homology for measures. *Computational Geometry*, 58:70–96, 2016. https://arxiv.org/abs/1306.0039.

[BCY18] Jean-Daniel Boissonnat, Frédéric Chazal, and Mariette Yvinec. *Geometric and topological inference*, volume 57. Cambridge University Press, 2018. https://hal.inria.fr/hal-01615863/.

[BDG+] Jean-Daniel Boissonnat, Ramsay Dyer, Arijit Ghosh, André Lieutier, and Mathijs Wintraecken. Local conditions for triangulating submanifolds of Euclidean space. Preprint. https://hal.inria.fr/hal-02267620.

[BDG18] Jean-Daniel Boissonnat, Ramsay Dyer, and Arijit Ghosh. Delaunay triangulation of manifolds. *Foundations of Computational Mathematics*, 18(2):399–431, 2018. https://hal.inria.fr/hal-01509888.

[BE17] Ulrich Bauer and Herbert Edelsbrunner. The Morse theory of Čech and Delaunay complexes. *Transactions of the American Mathematical Society*, 369(5):3741–3762, 2017. https://arxiv.org/abs/1312.1231.

[BF04] Jean-Daniel Boissonnat and Julia Flötotto. A coordinate system associated with points scattered on a surface. *Computer-Aided Design*, 36(2):161–174, 2004. https://www.sciencedirect.com/science/article/abs/pii/S0010448503000599.

[BG14] Jean-Daniel Boissonnat and Arijit Ghosh. Manifold reconstruction using tangential Delaunay complexes. *Discrete & Computational Geometry*, 51(1):221–267, 2014. https://hal.inria.fr/hal-00932209/PDF/Journal-version.pdf.

[BGO09] Jean-Daniel Boissonnat, Leonidas J Guibas, and Steve Y Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. *Discrete & Computational Geometry*, 42(1):37–70, 2009. https://link.springer.com/article/10.1007/s00454-009-9175-1.

[BGO19] Nicolas Berkouk, Grégory Ginot, and Steve Oudot. Level-sets persistence and sheaf theory. Preprint. https://arxiv.org/abs/1907.09759, 2019.

[BL13] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. Preprint. https://arxiv.org/abs/1311.3681, 2013.

[BL19] Claire Brécheteau and Clément Levrard. A k-points-based distance for robust geometric inference. Preprint. https://hal.archives-ouvertes.fr/hal-02266408, 2019.

[BLM17] Blanche Buet, Gian Paolo Leonardi, and Simon Masnou. A varifold approach to surface approximation. *Archive for Rational Mechanics and Analysis*, 226(2):639–694, 2017. https://arxiv.org/abs/1609.03625.

[BLM+19a] Gregory Bell, Austin Lawson, Joshua Martin, James Rudzinski, and Clifford Smyth. Weighted persistent homology. *Involve, a Journal of Mathematics*, 12(5):823–837, 2019. https://arxiv.org/abs/1709.00097.

[BLM19b] Blanche Buet, Gian Paolo Leonardi, and Simon Masnou. Weak and approximate curvatures of a measure: a varifold perspective. Preprint. https://arxiv.org/abs/1904.05930, 2019.

[BLW19] Jean-Daniel Boissonnat, André Lieutier, and Mathijs Wintraecken. The reach, metric distortion, geodesic convexity and the variation of tangent spaces. *Journal of Applied and Computational Topology*, 3(1-2):29–58, 2019. https://link.springer.com/article/10.1007/s41468-019-00029-8.

[Buc14] Mickaël Buchet. *Topological inference from measures*. PhD thesis, Paris 11, 2014. https://hal.inria.fr/tel-01108521/document.

[BW] Jean-Daniel Boissonnat and Mathijs Wintraecken. The topological correctness of PL-approximations of isomanifolds. Preprint. https://hal.inria.fr/hal-02386193.

[Cam03] Francesco Camastra. Data dimensionality estimation methods: a survey. *Pattern recognition*, 36(12):2945–2954, 2003. https://www.sciencedirect.com/science/article/abs/pii/S0031320303001766.

[CB15] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 14(05):1550066, 2015. https://arxiv.org/abs/1210.0819.

[CCBDS14] Frédéric Chazal, William Crawley-Boevey, and Vin De Silva. The observable structure of persistence modules. Preprint. https://arxiv.org/abs/1405.5644, 2014.

[CCSL09] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in Euclidean space. *Discrete & Computational Geometry*, 41(3):461–479, 2009. https://link.springer.com/article/10.1007/s00454-009-9144-8.

[CCSL+17] Frédéric Chazal, David Cohen-Steiner, André Lieutier, Quentin Mérigot, and Boris Thibert. Inference of curvature using tubular neighborhoods. In *Modern Approaches to Discrete Curvature*, pages 133–158. Springer, 2017. https://hal.archives-ouvertes.fr/hal-01425558.

[CCSM11] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Journal on Found. of Comp. Mathematics*, 11(6):733–751, 2011. https://geometrica.saclay.inria.fr/team/Fred.Chazal/papers/ccsm-gipm-11/ccsm-gipm-11.pdf.

[CDS10] Gunnar Carlsson and Vin De Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010. https://link.springer.com/article/10.1007/s10208-010-9066-0.

[CdSGO16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics, 2016. https://arxiv.org/abs/1207.3674.

[CDSO14] Frédéric Chazal, Vin De Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, 2014. https://arxiv.org/abs/1207.3885.

[CG20] Gunnar Carlsson and Rickard Brüel Gabrielsson. Topological approaches to deep learning. In *Topological Data Analysis*, pages 119–146. Springer, 2020. https://arxiv.org/abs/1811.01122.

[CIDSZ08] Gunnar Carlsson, Tigran Ishkhanov, Vin De Silva, and Afra Zomorodian. On the local behavior of spaces of natural images. *International journal of computer vision*, 76(1):1–12, 2008. https://www.researchgate.net/publication/220659347_On_the_Local_Behavior_of_Spaces_of_Natural_Images.

[CO08] Frédéric Chazal and Steve Yann Oudot. Towards persistence-based reconstruction in Euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, SCG '08,

pages 232–241, New York, NY, USA, 2008. ACM. https://geometrica.saclay.inria.fr/team/Steve.Oudot/papers/co-tpbr-08/co-tpbr-08.pdf.

[CS18] Nicholas J Cavanna and Donald R Sheehy. The generalized persistent nerve theorem. Preprint. https://arxiv.org/abs/1807.07920, 2018.

[CSEH07] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & computational geometry*, 37(1):103–120, 2007. https://link.springer.com/article/10.1007/s00454-006-1276-5.

[CZ09] Gunnar Carlsson and Afra Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, 2009. https://link.springer.com/article/10.1007/s00454-009-9176-0.

[DAE+08] Mary-Lee Dequeant, Sebastian Ahnert, Herbert Edelsbrunner, Thomas MA Fink, Earl F Glynn, Gaye Hattem, Andrzej Kudlicki, Yuriy Mileyko, Jason Morton, Arcady R Mushegian, et al. Comparison of pattern detection methods in microarray time series of the segmentation clock. *PLoS One*, 3(8):e2856, 2008. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002856.

[dC92] M.P. do Carmo. *Riemannian Geometry*. Mathematics (Boston, Mass.). Birkhäuser, 1992.

[DVW15] Ramsay Dyer, Gert Vegter, and Mathijs Wintraecken. Riemannian simplices and triangulations. *Geometriae Dedicata*, 179(1):91–138, 2015. https://drops.dagstuhl.de/opus/volltexte/2015/5136/.

[Ede93] Herbert Edelsbrunner. The union of balls and its dual shape. In *Proceedings of the ninth annual symposium on Computational geometry*, pages 218–231, 1993. https://link.springer.com/article/10.1007/BF02574053.

[ER14] Kevin J Emmett and Raul Rabadan. Characterizing scales of genetic recombination and antibiotic resistance in pathogenic bacteria using topological data analysis. In *International Conference on Brain Informatics and Health*, pages 540–551. Springer, 2014. https://arxiv.org/abs/1406.1219.

[Fed59] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959. https://www.ams.org/journals/tran/1959-093-03/S0002-9947-1959-0110078-1/S0002-9947-1959-0110078-1.pdf.

[Fed14] Herbert Federer. *Geometric measure theory*. Springer, 2014.

[FG15] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015. https://arxiv.org/abs/1312.2128.

[Fre02] Daniel Freedman. Efficient simplicial reconstructions of manifolds from their samples. *IEEE transactions on pattern analysis and machine intelligence*, 24(10):1349–1357, 2002. http://www.cs.jhu.edu/~misha/Fall05/Papers/freedman02.pdf.

[Gai05] Alexander A Gaifullin. Computation of characteristic classes of a manifold from a triangulation of it. *Russian Mathematical Surveys*, 60(4):615, 2005. https://www.maths.ed.ac.uk/~v1ranick/papers/gaifullin.pdf.

[GDR03] Rocio Gonzalez-Diaz and Pedro Real. Computation of cohomology operations of finite simplicial complexes. *Homology, Homotopy and Applications*, 5(2):83–93, 2003. https://arxiv.org/abs/math/0110332.

[GHI+15] Marcio Gameiro, Yasuaki Hiraoka, Shunsuke Izumi, Miroslav Kramar, Konstantin Mischaikow, and Vidit Nanda. A topological measurement of protein compressibility. *Japan Journal of Industrial and Applied Mathematics*, 32(1):1–17, 2015. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.363.1768&rep=rep1&type=pdf.

[GMM13] Leonidas Guibas, Dmitriy Morozov, and Quentin Mérigot. Witnessed k-distance. *Discrete & Computational Geometry*, 49(1):22–45, 2013. https://arxiv.org/abs/1102.4972.

[Hat02] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[HGK12] Giseon Heo, Jennifer Gamble, and Peter T Kim. Topological analysis of variance and the maxillary complex. *Journal of the American Statistical association*, 107(498):477–492, 2012. https://www.tandfonline.com/doi/abs/10.1080/01621459.2011.641430.

[Hir12] Morris W Hirsch. *Differential topology*, volume 33. Springer Science & Business Media, 2012. http://people.dm.unipi.it/benedett/HIRSCH.pdf.

[Jos04] Michael Joswig. Computing invariants of simplicial manifolds. Preprint. https://arxiv.org/abs/math/0401176, 2004.

[KRW16] Jisu Kim, Alessandro Rinaldo, and Larry Wasserman. Minimax rates for estimating the dimension of a manifold. Preprint. https://arxiv.org/abs/1605.01011, 2016.

[KSC+20] Jisu Kim, Jaehyeok Shin, Frédéric Chazal, Alessandro Rinaldo, and Larry Wasserman. Homotopy reconstruction via the Čech complex and the Vietoris-Rips complex. In *The 36th International Symposium on Computational Geometry (SoCG 2020)*, 2020. https://drops.dagstuhl.de/opus/volltexte/2020/12212/.

[Lab18] Fujitsu Laboratories. Estimating the degradation state of old bridges—Fujitsu supports ever-increasing bridge inspection tasks with AI technology. *Fujitsu Journal*, March 2018. https://journal.jp.fujitsu.com/en/2018/03/01/01/.

[LD19] Didong Li and David B Dunson. Geodesic distance estimation with spherelets. Preprint. https://arxiv.org/abs/1907.00296, 2019.

[Lee13] John M Lee. Smooth manifolds. In *Introduction to Smooth Manifolds*, pages 1–31. Springer, 2013.

[LMR17] Anna V Little, Mauro Maggioni, and Lorenzo Rosasco. Multiscale geometric methods for data sets I: Multiscale SVD, noise and curvature. *Applied and Computational Harmonic Analysis*, 43(3):504–567, 2017. http://cbcl.mit.edu/publications/ai-publications/2012/MIT-CSAIL-TR-2012-029.pdf.

[Lyt05] Alexander Lytchak. Almost convex subsets. *Geometriae Dedicata*, 115(1):201–218, 2005. http://www.mi.uni-koeln.de/~alytchak/preprints/almost_convex.pdf.

[MLX+08] Deyu Meng, Yee Leung, Zongben Xu, Tung Fung, and Qingfu Zhang. Improving geodesic distance estimation based on locally linear assumption. *Pattern Recognition Letters*, 29(7):862–870, 2008. https://www.sciencedirect.com/science/article/abs/pii/S0167865508000093.

[MMM18] Diego H Díaz Martínez, Facundo Mémoli, and Washington Mio. The shape of data and probability measures. *Applied and Computational Harmonic Analysis*, 2018. https://arxiv.org/abs/1509.04632.

[MS16] John Milnor and James D Stasheff. *Characteristic Classes.(AM-76)*, volume 76. Princeton university press, 2016. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.448.869&rep=rep1&type=pdf.

[MSW19] Facundo Memoli, Zane Smith, and Zhengchao Wan. The Wasserstein transform. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4496–4504, Long Beach, California, USA, 09–15 Jun 2019. PMLR. http://proceedings.mlr.press/v97/memoli19a/memoli19a.pdf.

[MSWW19] John Malik, Chao Shen, Hau-Tieng Wu, and Nan Wu. Connecting dots: from local covariance to empirical intrinsic geometry and locally linear embedding. *Pure and Applied Analysis*, 1(4):515–542, 2019. https://arxiv.org/abs/1804.02811.

[MTCW10] Shawn Martin, Aidan Thompson, Evangelos A Coutsias, and Jean-Paul Watson. Topology of cyclo-octane energy landscape. *The journal of chemical physics*, 132(23):234115, 2010. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3188624/.

[Mun16] James R Munkres. *Elementary Differential Topology.(AM-54)*, volume 54. Princeton University Press, 2016.

[NLC11] Monica Nicolau, Arnold J Levine, and Gunnar Carlsson. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences*, 108(17):7265–7270, 2011. https://www.pnas.org/content/108/17/7265.

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008. http://people.cs.uchicago.edu/~niyogi/papersps/NiySmaWeiHom.pdf.

[Oud15] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society Providence, 2015. https://geometrica.saclay.inria.fr/team/Steve.Oudot/books/o-pt-fqrtda-15/surv-209.pdf.

[PC14] Jose A Perea and Gunnar Carlsson. A Klein-bottle-based dictionary for texture representation. *International journal of computer vision*, 107(1):75–97, 2014. https://fds.duke.edu/db/attachment/2638.

[PD07] Sylvain Paris and Frédo Durand. A topological approach to hierarchical segmentation using mean shift. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007. https://ieeexplore.ieee.org/document/4270253.

[PEVdW+17] Pratyush Pranav, Herbert Edelsbrunner, Rien Van de Weygaert, Gert Vegter, Michael Kerber, Bernard JT Jones, and Mathijs Wintraecken. The topology of the cosmic web in terms of persistent Betti numbers. *Monthly Notices of the Royal Astronomical Society*, 465(4):4281–4310, 2017. https://arxiv.org/abs/1608.04519.

[PWZ15] J. Phillips, B. Wang, and Y Zheng. Geometric inference on kernel density estimates. In *Proc. 31st Annu. Sympos. Comput. Geom (SoCG 2015)*, pages 857–871, 2015.

[PZ16] Gerlind Plonka and Yi Zheng. Relation between total variation and persistence distance and its application in signal processing. *Advances*

in Computational Mathematics, 42(3):651–674, 2016. http://na.math.uni-goettingen.de/pdf/plonka_zheng.pdf.

[SDB16] Lee M Seversky, Shelby Davis, and Matthew Berger. On time-series topological data analysis: New data and opportunities. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 59–67, 2016. https://www.cv-foundation.org/openaccess/content_cvpr_2016_workshops/w23/papers/Seversky_On_Time-Series_Topological_CVPR_2016_paper.pdf.

[SHCP18] Leo Speidel, Heather A Harrington, S Jonathan Chapman, and Mason A Porter. Topological data analysis of continuum percolation with disks. *Physical Review E*, 98(1):012318, 2018. https://arxiv.org/abs/1804.07733.

[She13] Donald R. Sheehy. Linear-size approximations to the Vietoris-Rips filtration. *Discrete & Computational Geometry*, 49(4):778–796, 2013. https://arxiv.org/abs/1203.6786.

[SMC07] Gurjeet Singh, Facundo Mémoli, and Gunnar E Carlsson. Topological methods for the analysis of high dimensional data sets and 3d object recognition. *SPBG*, 91:100, 2007. https://research.math.osu.edu/tgda/mapperPBG.pdf.

[Sou11] Thierry Sousbie. The persistent cosmic web and its filamentary structure–I. Theory and implementation. *Monthly Notices of the Royal Astronomical Society*, 414(1):350–383, 2011. https://academic.oup.com/mnras/article/414/1/384/1091264.

[Spi70] Michael D Spivak. *A comprehensive introduction to differential geometry*. Publish or perish, 1970.

[SPK11] Thierry Sousbie, Christophe Pichon, and Hajime Kawahara. The persistent cosmic web and its filamentary structure–II. Illustrations. *Monthly Notices of the Royal Astronomical Society*, 414(1):384–403, 2011. https://academic.oup.com/mnras/article/414/1/384/1091264.

[SW11] Amit Singer and Hau-tieng Wu. Orientability and diffusion maps. *Applied and computational harmonic analysis*, 31(1):44–58, 2011. https://www.sciencedirect.com/science/article/pii/S1063520310001144.

[Tin19] Raphaël Tinarrage. Recovering the homology of immersed manifolds. Preprint. https://arxiv.org/abs/1912.03033, 2019.

[Tin20] Raphaël Tinarrage. Computing persistent Stiefel-Whitney classes of line bundles. Preprint. https://arxiv.org/abs/2005.12543, 2020.

[TVF13] Hemant Tyagi, Elıf Vural, and Pascal Frossard. Tangent space estimation for smooth embeddings of Riemannian manifolds. *Information and Inference: A Journal of the IMA*, 2(1):69–114, 2013. https://arxiv.org/abs/1208.1065.

[Ume17] Yuhei Umeda. Time series classification via topological data analysis. *Transactions of the Japanese Society for Artificial Intelligence*, 32(3):D–G72_1, 2017. https://pdfs.semanticscholar.org/18e9/f6dfbe544d378b3c90f19892328ca69062d5.pdf.

[Vil08] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008. https://ljk.imag.fr/membres/Emmanuel.Maitre/lib/exe/fetch.php?media=b07.stflour.pdf.

[WOC14] Yuan Wang, Hernando Ombao, and Moo K Chung. Persistence landscape of functional signal and its application to epileptic electroencephalogram data. *ENAR Distinguished Student Paper Award*, 2014. https://pdfs.semanticscholar.org/69fd/fa254fd019ce4b7d7879de1ceddd61040d2d.pdf.

[Yar10] Andrew Yarmola. *Persistence and computation of the cup product*. PhD thesis, Stanford University, 2010. https://web.math.princeton.edu/~yarmola/assets/pdf/cup_prod.pdf.

université PARIS-SACLAY | ÉCOLE DOCTORALE de mathématiques Hadamard (EDMH)

Titre: Inférence topologique à partir de mesures et de fibrés vectoriels

Mots clés: Homologie persistante, distance de Wasserstein, classes caractéristiques, Analyse Topologique des Données

Résumé: Nous contribuons à l'inférence topologique, basée sur la théorie de l'homologie persistante, en proposant trois familles de filtrations. Nous établissons pour chacune d'elles des résultats de consistance – c'est-à-dire de qualité d'approximation d'un objet géométrique sous-jacent –, et de stabilité – c'est-à-dire de robustesse face à des erreurs de mesures initiales. Nous proposons des algorithmes concrets afin de pouvoir utiliser ces méthodes en pratique.

La première famille, les filtrations-DTM, est une alternative robuste à la filtration de Čech habituelle lorsque le nuage de points est bruité ou contient des points aberrants. Elle repose sur la notion de distance à la mesure, qui permet d'obtenir une stabilité au sens de la distance de Wasserstein.

Deuxièmement, nous proposons les filtrations relevées, qui permettent d'estimer l'homologie des variétés immergées, même quand leur portée est nulle. Nous introduisons la notion de portée normale, et montrons qu'elle permet de contrôler des quantités géométriques associées à la variété. Nous étudions l'estimation des espaces tangents par les matrices de covariance locale.

Troisièmement, nous développons un cadre théorique pour les filtrations de fibrés vectoriels, et définissons les classes de Stiefel-Whitney persistantes. Nous montrons que les classes de Stiefel-Whitney persistantes associées aux filtrations de fibrés de Čech sont consistantes et stables en distance de Hausdorff. Pour permettre leur mise en œuvre algorithmique, nous introduisons la notion de condition étoile faible.

Title: Topological inference from measures and vector bundles

Keywords: Persistent homology, Wasserstein distance, characteristic classes, Topological Data Analysis

Abstract: We contribute to the theory of topological inference, based on the theory of persistent homology, by proposing three families of filtrations. For each of them, we prove consistency results—that is, quality of approximation of an underlying geometric object—, and stability results—that is, robustness against initial measurement errors. We propose concrete algorithms in order to use these methods in practice.

The first family, the DTM-filtrations, is a robust alternative to the usual Čech filtration when the point cloud is noisy or contains anomalous points. It is based on the notion of distance to measure, which allows to obtain stability in the sense of the Wasserstein distance.

Secondly, we propose the lifted filtrations, which make it possible to estimate the homology of immersed manifolds, even when their reach is zero. We introduce the notion of normal reach, and show that it allows to control geometric quantities associated to the manifold. We study the estimation of tangent spaces by local covariance matrices.

Thirdly, we develop a theoretical framework for vector bundle filtrations, and define the persistent Stiefel-Whitney classes. We show that the persistent Stiefel-Whitney classes associated to the Čech bundle filtrations are Hausdorff-stable and consistent. To allow their algorithmic implementation, we introduce the notion of weak star condition.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France