Statistical descriptors for Topological Data Analysis using Optimal Transport

Théo Lacombe

Remerciements

À venir dans la version finale du manuscrit.
Pending for the final version of the manuscript.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction en français

1.1.1 Analyse des données, statistiques et géométrie

L'analyse des données. L'analyse des données est devenue, au cours de la dernière décennie, un des domaines de recherche en mathématiques appliquées et en informatique les plus actifs et prolifiques. L'évolution des moyens d'acquisition et des capacités de stockage en tout genre a permis de constistuer des collections de données colossales en tout genre : images [KH+09], formes 3D [CFG+15], sons [BMEWL11], données médicales et biologiques [CHF12], réseaux sociaux [YV15, KKM+16, PWZ+17]... Le but de l'analyse des données est de comprendre et de valoriser ces nouvelles informations disponibles.

Pour cela, le paradigme le plus courant de nos jours est de recourir à "l'apprentissage automatique", dont l'idée générale est la suivante : proposer un algorithme capable d'extraire une information utile d'un jeu de données, permettant par exemple de réaliser de regrouper les données par similarité (*clustering*), d'attribuer des labels à de nouvelles observations (*classification*), etc.

Statistiques et géométrie. Évidemment, il n'existe pas d'algorithme tout-puissant, qui serait capable de prédire sans erreur dans n'importe quel contexte. Si les observations et les les labels ne sont pas ou peu liés, ou si le nombre d'observations est trop faible, même le meilleur algorithme possible ne saura pas se montrer utile lorsqu'il s'agira d'extrapoler son apprentissage à de nouvelles observations.

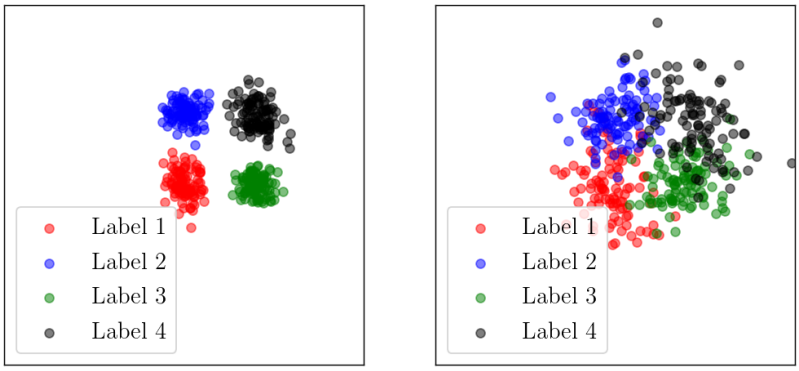


Figure 1.1: Nuages de 400 points tirés selon un mélange de quatre gaussiennes, représentant quatre classes. Sur la figure de gauche, les gaussiennes ont une faible variance, et il est facile de séparer les différentes classes. Sur la figure de droite, la séparation des classes est plus difficile.

La difficulté à résoudre un problème d'apprentissage, par exemple sa dépendance au nombre d'observations, peut se mesurer au moyen d'outils statistiques. Dans un problème de classification ou de clustering par exemple, plus les observations d'une même classe (c-à-d. partageant un même label) sont concentrées autour de leur moyenne (faible *variance*) et plus les classes sont distinctes, plus le problème sera résolu efficacement (voir Figure 1.1) De manière générale, il est important de comprendre comment nos observations sont *distribuées*. L'approche généralement adoptée en statistique est la suivante : nos observations sont *indépendantes* et proviennent d'une distribution de probabilité sous-jacente (inconnue). Connaître parfaitement cette distribution permettrait de comprendre le comportement des algorithmes d'apprentissage et d'optimiser leurs performances sur ce jeu de données (sans pour autant arriver nécessairement à un score parfait). En pratique, l'idée est donc d'estimer certaines propriétés de cette distribution sous-jacente à partir de notre *échantillon* d'observations.

L'exemple le plus élémentaire, et qui va nous amener à des considérations géométriques, est l'estimation de l'*espérance* de la loi sous-jacente, c'est-à-dire la valeur moyenne d'une observation tirée aléatoirement. Ce problème est en général très simple quand on observe des nombres réels : si on observe $(x_1 \dots x_n)$ n nombres tirés indépendamment selon une loi μ, un estimateur de l'espérance de μ est simplement la moyenne arithmétique des $(x_i)_i$, c'est-à-dire $\frac{1}{n}\sum_{i=1}^{n} x_i$. Cependant, dans les problèmes d'apprentissage modernes,

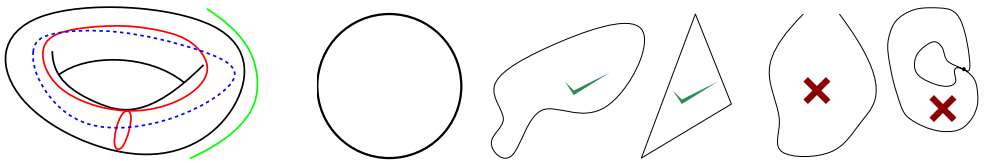


Figure 1.2: À gauche, un tore dont on identifie les propriétés topologiques. Ensuite, un cercle, deux transformations qui en préservent la topologie, et deux transformations qui ne les préservent pas (déchirement et recollement respectivement).

la géométrie sous-jacente à nos observations est généralement complexe et inconnue : graphes, images, molécules... autant de cas où la notion de moyenne, et *a fortiori* des descripteurs statistiques plus sophistiqués, est difficile à proprement définir et estimer.

1.1.2 Descripteurs topologiques

Dans cette thèse, nous nous intéresserons à un type d'observations en particulier, les *diagrammes de persistance*, qui proviennent de *l'analyse topologique des données*. L'étude de la géométrie sous-jacente à ces objets, tout comme la conception et le calcul de descripteurs statistiques, sont des sujets de recherche actifs.

Topologie. Brièvement, la notion de topologie renvoie à celle de "forme". Étant donné un objet, par exemple une forme 3D, combien celui-ci possède-t-il de composantes connexes ? Peut-on identifier des boucles caractéristiques à sa surface ? Possède-t-il des cavités ? Le tore représenté en Figure 1.2 possède par exemple une composante connexe, deux boucles (en rouge), et une cavité (en bleu). Le cercle, lui, possède une composante connexe et une boucle.

Dans un cadre plus général, la topologie peut se comprendre comme les propriétés d'un objet qui sont préservées par des transformations continues, sans déchirement ni recollement, voir Figure 1.2. Par exemple, lorsqu'on applique une telle transformation à un cercle, si la géométrie de l'objet peut varier, sa topologie reste inchangée. D'un point de vue formel, les propriétés topologiques d'un objet sont décrites par ses *groupes d'homologie*. Les groupes d'homologie d'un espace topologique (objet) sont décrit par une suite de groupe abélien $(H_k)_k$ ($k = 0$ correspond aux propriétés topologiques de dimension 0, c'est-à-dire les composantes connexes; $k = 1$ correspond aux boucles, $k = 2$ aux cavités, etc.) dont les générateurs indentifient les

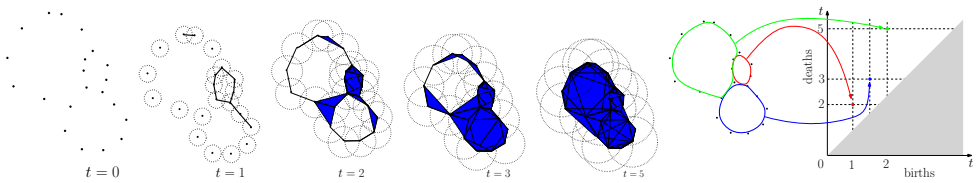


Figure 1.3: Illustration du phénomène d'évolution de la topologie à différentes échelles.

propriétés topologiques "indépendantes" de l'espace. Par exemple, pour un tore, H_1 a deux générateurs. Ceux-ci correspondent intuitivement aux deux boucles rouges sur la Figure 1.2: toutes les autres boucles que l'on peut dessiner sur le tore s'écrivent - en un sens - comme une combinaison linéaire de ces deux boucles. Le lecteur intéressé peut consulter l'Annexe A ou [Mun84] pour une présentation plus détaillée de la théorie de l'homologie.

Descripteurs multi-échelles de la topologie. En pratique néanmoins, l'usage de la topologie en apprentissage automatique n'est pas immédiat. La plupart des méthodes d'acquisition des données vont produire des objets sous forme de *nuage de points*, dont la topologie intrinsèque est excessivement simple. Les groupes d'homologie d'un nuage de points $(H_k)_k$ sont triviaux dès que $k \geq 1$, tandis que H_0 est le groupe abélien libre engendré par N éléments, où N désigne la cardinalité du nuage de points. L'homologie seule n'est pas capable de refléter la structure sous-jacente de tels objets.

L'idée fondamentale développée dans les années 2000 (quoique les fondements peuvent être tracés tout au long du XX$^{\text{e}}$ siècle) est de regarder la topologie d'un espace topologique *à différentes échelles*, et de regarder quelles sont les propriétés topologiques qui *persistent* à travers celles-ci. Considérons par exemple le nuage de points représenté à gauche sur la Figure 1.3. Initialement, il ne s'agit (topologiquement) que d'une collection de composantes connexes indépendantes. Il apparait cependant naturel d'identifier trois boucles, de tailles différentes. Regarder la topologie de cet objet à différentes échelles va nous permettre de détecter lesdites boucles. Pour introduire cette notion d'échelle, une idée est de faire grossir des boules centrées en chacun des points du nuage. Le paramètre d'échelle $t \geq 0$ correspond ici au rayon des boules. Ainsi, à partir du nuage de points X au paramètre $t = 0$, on construit une famille d'objets $(X_t)_{t \geq 0}$. On parle alors de *filtration*. Pour certaines *valeurs critiques* du paramètre t, voir Figure 1.3, la topologie[1] change : des boucles apparaissent (par exemple à

[1]Ici, par souci de simplicité, on s'intéresse uniquement aux boucles.

$t = 1$ ou $t = 2$ sur la figure) ou disparaissent (lorsqu'elles sont complètement "remplies", $t = 2$, $t = 3$, $t = 5$ sur la figure).

Ce sont ces valeurs critiques, appelées *temps de naissance* et de *mort* des propriétés topologiques, qui seront enregistrées dans un descripteur topologique : le *diagramme de persistance*. Un diagramme de persistance se présente comme une collection de points dans le plan. La présence d'un point de coordonnées (b, d) dans le diagramme va indiquer qu'une propriété topologique (composante connexe, boucle, cavité...) est apparue à l'échelle $t = b$ et a disparu à l'échelle $t = d$. Ainsi, un point proche de la diagonale, c'est-à-dire tel que $d \simeq b$, représente une propriété topologique qui est apparu et a presque immédiatement disparu : cette propriété a peu *persisté* à travers les échelles. À l'inverse, un point loin de la diagonale va représenter une composante topologique présente durant un large intervalle d'échelles, généralement considérée comme plus significative. Formellement, un diagramme de persistance est décrit comme un multi-ensemble de points[2] supporté sur le demi-plan

$$\Omega := \{(b, d),\ \in \mathbb{R}^2, d > b\}$$

ou, de façon équivalente, comme une mesure de Radon μ s'écrivant

$$\mu := \sum_{x \in X} n_x \delta_x,$$

où $X \subset \Omega$ est localement fini, n_x et un entier et δ_x désigne la masse de Dirac en $x \in X$.[3]

Bien entendu, le cadre d'application de l'analyse topologique des données ne se résume pas aux nuages de points et à des boules qui grossissent ; la théorie générale est présentée dans la Section 2.1. À ce stade, l'essentiel est de retenir qu'il est donc possible de transformer une collection d'observations complexes en une collection de diagrammes de persistance, à partir de laquelle on peut envisager de produire une analyse statistique ou de réaliser une tâche d'apprentissage automatique.

[2]Un ensemble dans lequel les points peuvent être répétés

[3]Dans leur définition la plus générale, les diagrammes de persistance peuvent avoir des points avec des coordonnées infinies ($b = -\infty$ or $d = +\infty$) qui appartiennent donc au demi-plan étendu; ou des points sur la diagonale ($b = d$). Nous ignorons ces points dans cette thèse, voir les Remarques 2.1 et 2.2.

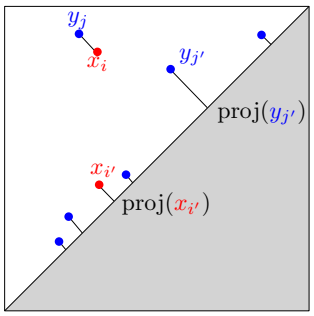


Figure 1.4: Deux diagrammes de persistance. La distance de bottleneck entre ces deux diagrammes est, par définition, la longueur de la plus longue arête dessinée.

1.1.3 Diagramme de persistance et apprentissage: limitations et enjeux

Les diagrammes de persistances portent une information riche et en général complémentaire aux autres méthodes classiques d'analyse de données. Malheureusement, leur incorporation dans les outils modernes d'apprentissage, tels que les réseaux de neurones, n'est absolument pas évidente.

Distance entre diagrammes de persistance. Pour comprendre cela, il faut dans un premier temps souligner qu'il est possible de mesurer une distance entre deux diagrammes de persistance, ce qui permet donc de comparer les deux objets initiaux d'un point de vue topologique. La distance de référence entre diagrammes est appelée *distance de bottleneck*, et se calcule de la façon suivante. Pour deux diagrammes X et Y, notons $(x_1 \dots x_n)$ et $(y_1 \dots y_m)$ leurs points respectifs (notons qu'on n'a pas nécessairement $n = m$). La première étape est de chercher à *transporter*[4] chaque point x_i du diagramme X vers un point y_j de Y ou, éventuellement, vers sa projection orthogonale sur la diagonale (voir Figure 1.4). Les points y_j de Y qui ne sont pas atteints par un point x_i de X sont eux aussi transportés sur leur projection sur la diagonale. Nous imposons de plus que ce transport soit *bijectif* : chaque point x_i est envoyé - au plus - sur un point y_j, et chaque y_j doit être atteint par - au plus - un seul x_i. En notant (x_i, y_j) quand x_i est transporté sur y_j, et $(x_i, \partial\Omega)$ (resp. $(\partial\Omega, y_j)$) lorsque x_i (resp. y_j) est transporté sur la diagonale, un *transport partiel* entre X et Y est décrit par une liste $\dots (x_i, y_j)_{ij} \dots (x_i, \partial\Omega) \dots (\partial\Omega, y_j) \dots$, où chaque point x_i de X et chaque point y_j de Y apparaît exactement une fois.

[4] On parle plus fréquemment "d'appariement", mais nous préfèrerons la terminologie de "transport" pour des raisons qui deviendront évidentes par la suite.

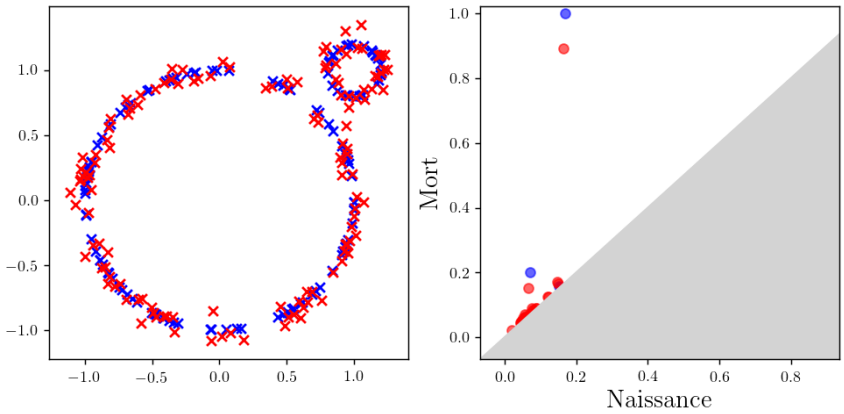


Figure 1.5: Illustration de la stabilité des diagrammes de persistance en bottleneck, et instabilité du nombre de points.

Bien entendu, il existe une multitude de façons de transporter un diagramme X sur un autre diagramme Y. Pour définir la distance entre deux diagrammes, nous allons chercher un transport partiel *optimal* au sens suivant : le *coût* d'un transport partiel est la plus grande distance parcourue entre deux points du transport. Un transport partiel optimal est un transport partiel dont le coût est minimal parmi tous les transports partiels entre X et Y possibles. Le coût minimal ainsi réalisé est, par définition, la distance de bottleneck. Formellement, si on note $\Gamma(X, Y)$ l'ensemble des transports entre deux diagrammes X et Y, on pose

$$\mathbf{d}_\infty(X, Y) := \min_{\gamma \in \Gamma(X,Y)} \max_{(x,y)\in\gamma} \|x - y\|.$$

Le choix de cette distance n'est pas arbitraire et est motivé par des considérations algébriques profondes qui sont détaillées dans la Section 2.1. Cette distance jouit de la propriété d'être *stable*, au sens où des objets initialement proches auront systématiquement des diagrammes proches en distance de bottleneck (Figure 1.5).

Limitations. Étant donnée une collection de diagrammes, grâce à la distance de bottleneck, il est possible de calculer des distances entre chaque paire de diagrammes. L'utilisation de certains algorithmes d'apprentissage basiques, comme celui du plus proche voisin, ou par exemple le *multi-dimensional scaling* (MDS), est alors possible, car ceux-ci requièrent seulement de savoir calculer des distances entre nos observations (ici les diagrammes).

Néanmoins, la plupart des algorithmes modernes d'apprentissage demandent beaucoup plus de structure, qui va faire défaut dans le cadre des diagrammes de persistance. En effet, *l'espace* des diagrammes de persistance n'est pas *linéaire*, ce qui veut essentiellement dire qu'il n'est pas possible de donner un sens[5] à la somme de deux diagrammes de persistance, ou à la multiplication d'un diagramme par un nombre réel. Il n'est donc pas possible d'incorporer naïvement les diagrammes de persistance dans les machineries d'apprentissage modernes, et leur utilisation dans la pratique en est donc compromise.

Une solution de remplacement va consister à *plonger* l'espace des diagrammes dans un espace linéaire, sur lequel il sera donc possible d'effectuer des tâches d'apprentissage aisément. Cela signifie qu'on cherche une application Φ qui transforme les diagrammes en *vecteurs*, c'est-à-dire un élément dans un espace linéaire. Une telle application est appelée une *vectorisation*. Par exemple, une vectorisation naive pourrait consister à associer à chaque diagramme X son nombre de points $\Phi(X)$, plongeant l'espace des diagrammes dans l'espace des nombres réels, bien entendu linéaire. Cette vectorisation est très décevante car deux diagrammes très différents (au sens de la distance de bottleneck) peuvent avoir le même nombre de points, mais aussi deux diagrammes très proches peuvent avoir un nombre de points très différent (comme en Figure 1.5). Autrement dit, $\Phi(X)$ n'a pas grand chose à voir avec X, ce qui est profondément regrettable.

Bien entendu, de nombreuses vectorisations bien plus satisfaisantes ont été proposées, avec d'importants succès dans les applications [Bub15, AEK+17, CCO17]. Néanmoins, l'usage d'une vectorisation soulève quelques interrogations importantes :

- Le choix de la vectorisation est arbitraire. Parmi le large catalogue de vectorisations disponibles, il n'existe pas à ce jour d'heuristique pour choisir laquelle sera adaptée à une tâche d'apprentissage donnée.

- Il semble impossible[6] d'obtenir, en toute généralité, un plongement qui soit *bi-stable*, c'est à dire pour lequel on aurait, pour toute paire de diagramme X, Y, et deux constantes $0 < A \leq B < \infty$,

$$A\|\Phi(X) - \Phi(Y)\| \leq \mathbf{d}_\infty(X, Y) \leq B\|\Phi(X) - \Phi(Y)\|,$$

  ce qui aurait le mérite d'assurer que des diagrammes proches ont des représentations proches, et inversement.

[5] Qui serait compatible avec la distance de bottleneck

[6] Ce problème est encore ouvert, mais tous les résultats actuels semblent aller dans ce sens [CB18, BW19, Wag19]

- Enfin, il faut rappeler que les diagrammes de persistance ont été construits en partie pour leur interprétabilité. Recourir à une vectorisation, et réaliser des calculs dans l'espace linéaire associé, retire cette interprétabilité. S'il est possible pour une collection de diagrammes $X_1 \dots X_n$ de calculer facilement une moyenne dans l'espace linéaire en posant $M = \frac{1}{n}\sum_{i=1}^{n} \Phi(X_i)$, il est important de comprendre que cette quantité n'a, *a priori*, rien à voir avec une notion de moyenne dans l'espace des diagrammes de persistance. Cela vaut plus généralement pour toutes les opérations réalisées dans l'espace linéaire (qui conduiraient par exemple à du clustering, etc.).

Enjeux. Si l'étude des vectorisations de l'espace des diagrammes de persistance reste un domaine de recherche pertinent, nous proposons ici de travailler *directement dans l'espace des diagrammes de persistance*. L'objectif est donc de répondre à la question

"Comment peut-on faire des statistiques dans l'espace des diagrammes de persistance ?"

Pour cela, il est important de comprendre d'abord quelles sont les propriétés géométriques de cet espace, pour lequel nous ne disposons *a priori* que d'une distance pour en étudier la structure. Cela doit permettre ensuite de proposer un cadre théorique solide dans lequel des outils statistiques sont proprement définis. Enfin, il est important que ces outils statistiques puissent être calculés et utilisés, au moins de façon approchée, en pratique.

1.1.4 Contributions et organisation du manuscrit

Afin de proposer des éléments de réponse à cette problématique, nous nous appuierons sur un autre domaine des mathématiques appliquées : la théorie du *transport optimal*, dont la présentation est faite dans la Section 2.2. Il s'agit peut-être du point essentiel de ce manuscrit : l'établissement d'un lien formel entre l'espace des diagrammes de persistance et les modèles utilisés en transport optimal. Cette connexion entre ces deux champs mathématiques est extrêmement prolifique : le transport optimal est un domaine très développé, tant dans sa théorie que dans les applications en statistiques et en apprentissage ; et nombre de ses outils s'adaptent à l'étude des diagrammes de persistance. Cela nous permettra d'obtenir de nouveaux résultats théoriques relatifs à l'utilisation des diagrammes de persistance en statistiques et en apprentissage, mais aussi de proposer divers algorithmes et outils numériques qui permettent de résoudre des problèmes difficiles

en persistance, comme l'estimation de moyennes (dites "de Fréchet"), la quantification, l'apprentissage de vectorisation, entre autres.

De façon générale, et sauf mention explicite du contraire, nous adoptons la convention suivante : les résultats (théorèmes, propositions, etc.) pour lesquels nous proposons une preuve sont le fruit de mon travail ou d'un travail joint avec mes collaborateurs : Mathieu Carrière, Frédéric Chazal, Marco Cuturi, Vincent Divol, Yuichi Ike, Steve Oudot, Martin Royer, et Yuhei Umeda.

Plan et lien avec les travaux réalisés durant la thèse.

Chapitre 2: Préliminaires. Ce chapitre présente les deux domaines impliqués dans ce travail: l'analyse topologique des données (Section 2.1) et le transport optimal (Section 2.2). L'objectif est de présenter rapidement l'état de l'art relatif aux différents problèmes abordés dans ce manuscrit. Notons néanmoins que la construction algébrique des diagrammes de persistance, developpée dans les sous-sections A.1 et A.2, est indépendante du reste du manuscrit.

Les contributions sont présentées en deux parties, en fonction de leur nature théorique (propriétés géométriques générales, etc.) ou appliquée (pour faire simple, conduisant à une implémentation).

Partie I — Théorie.

Chapitre 3: Un formalisme issu du transport optimal pour les diagrammes de persistance. Ce chapitre présente et développe les fondements théoriques de ce manuscrit. Son contenu repose essentiellement sur la section 3 de l'article [DL19], en révisions mineures au *Journal of Applied and Computational Topology*. Nous y présentons comment les métriques utilisées habituellement pour comparer les diagrammes de persistance peuvent se reformuler comme des problèmes de transport optimal *partiels*, et les nombreux résultats théoriques qui en découlent.

Chapitre 4: Moyenne de Fréchet pour les diagrammes de persistance: aspects théoriques. Ce chapitre est consacré à l'étude des moyennes de Fréchet (ou barycentres) pour les diagrammes de persistance. Il repose sur la section 4 de [DL19]. On y prouve notamment un résultat d'existence très général, et on établit un lien fort entre ces moyennes et les fameux "barycentres de Wasserstein", largement étudiés dans la littérature du transport optimal traditionnel.

Partie II — Applications.

Chapitre 5: Algorithmes efficaces pour l'estimation des moyennes de Fréchet des diagrammes de persistance. On propose ici un algorithme pour approcher les moyennes entre diagrammes; particulièrement efficace pour traiter les problèmes à grande échelle. Cette approche a été publiée dans les annales de la conférence internationale *Neural Information Processing Systems, 2018*, voir [LCO18].

Chapitre 6: Représentations linéaires des diagrammes de persistance. Ce chapitre se consacre à l'étude des représentations linéaires de diagramme de persistance, pour lesquelles on propose une caractérisation exhaustive. Une fois encore, les résultats théoriques (Section 6.1) sont issus de [DL19, §5.1]. On propose ensuite une application de ce résultat en apprentissage automatique en introduisant PersLay, une couche pour les réseaux de neurones[7] spécifiquement élaborée pour apprendre des représentations adaptées à une tâche donnée. Ce travail a été publié dans les annales de la conférence *Artificial Intelligence and Statistics, 2020*, voir [CCI+19].

Chapitre 7: Exemples complémentaires. Ce dernier chapitre regroupe diverses applications et algorithmes qui ont été développés afin de résoudre divers problèmes relatifs aux diagrammes de persistance. La Section 7.1 démontre l'intérêt du formalisme théorique développé dans le Chapitre 3 pour étudier les diagrammes de persistance dans un contexte aléatoire. On y trouve des résultats de convergence et de stabilité pour des analogues aléatoires des diagrammes de persistance. La Section 7.2 étudie la *quantisation* des diagrammes de persistance. Enfin, la Section 7.3 propose une façon simple et efficace pour estimer des distances entre diagrammes *a transformation près*. Ces résultats n'ont pas encore été publiés, mais les algorithmes correspondant ont, ou vont être, intégrés à la librairie d'analyse topologique des données `Gudhi` [GUD15]; ce manuscrit de thèse est une bonne occasion d'en présenter les rouages.

Code et contributions à la librairie `gudhi`. La plupart des méthodes présentées en Partie II ont été, ou vont être, incorporées à la librairie `Gudhi`.

[7] disponible publiquement et incorporée à la librairie `Gudhi`.

1.2 Introduction in English

1.2.1 Data Analysis, Statistics, and Geometry

Data Analysis. Data analysis has become, during the last decade, one of the most active research areas in applied mathematics and computer science. Progress in data gathering and storage has led to very large datasets of various types: images [KH+09], 3D shapes [CFG+15], music [BMEWL11], medical and biological data [CHF12], social networks [YV15, KKM+16, PWZ+17]... Data analysis aims to understand and add value to such newly available information.

To do so, the most standard paradigm nowadays is to make use of "machine learning". It aims at designing algorithms that will be able to extract useful information from a given set of observations; allowing for instance to regroup data by similarity (*clustering*), to assign labels to new observations based on the labels of the *training* data (*classification*), etc.

Statistics and geometry. Of course, there is no omnipotent algorithm that would be able to predict with no error in any context. If observations and labels are not correlated or if the number of observations is too low, even the best possible algorithm will not be of any use when it comes to generalizing its learning to new sets of observations.

The difficulty to solve a given learning problem, for instance its dependence on the number of observations, can be measured with statistical tools. In a classification or clustering problem for instance, the more the observations sharing the same label are concentrated around their mean value (low *variance*) and the more the different classes are separated from each other, the easier the problem will be to solve (see Figure 1.6). In general, it is important to understand how our observations are *distributed*. The general approach taken in statistical learning is to assume that our observations are *independent* and come from an (unknown) underlying probability distribution. Perfect knowledge of this distribution would allow us to understand the behavior of learning algorithms and to optimize their performances on a given dataset (without necessarily reaching a perfect score). In practice, it is thus useful to infer some properties of the underlying probability distribution from our *sample* of observations.

The most basic example, which will lead to geometric considerations, is the estimation of the *expectation* of the underlying law, that is the average value of a randomly sampled observation. When the observations are real numbers, this problem is pretty simple: if we observe $(x_1 \dots x_n)$ n numbers independently sampled from a law μ, an estimation of the expectation of

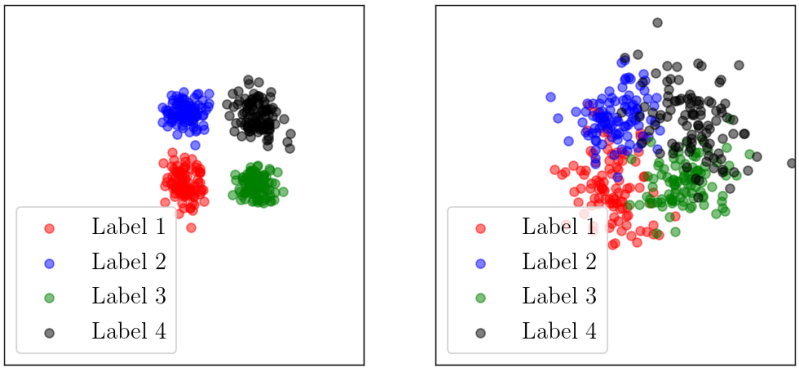


Figure 1.6: Point clouds (400 points) sampled from a mixture of four Gaussian distributions, representing four classes. On the left subfigure, the Gaussians distributions have a low variance, it is thus easy to separate the different classes. On the right subfigure, class separation is harder to perform.

μ is simply given by the arithmetic mean of the $(x_i)_i$s, that is $\frac{1}{n}\sum_{i=1}^{n} x_i$. However, in modern machine learning problems, the underlying geometry of the observations is often complex and unknown: think of graphs, images, molecular structures... These are cases where the notion of mean, and *a fortiori* more sophisticated descriptors, is hard to clearly define and to estimate.

1.2.2 Topological Descriptors

In this thesis, we will focus on a specific type of observations: the *persistence diagrams*, which come from *topological data analysis*. Understanding the underlying geometry of these objects is an active research area, as is the design and the computation of statistical descriptors built on top of persistence diagrams.

Topology. Roughly speaking, the term "topology" is related to the notion of "shape". Considering a given object, such as a 3D shape, how many connected components does it have? Can we identify characteristic loops? Does it have cavities? For instance, the torus represented in Figure 1.7 has one connected component, two loops (red), and a cavity (blue). In contrast, a circle has a single connected component and a single loop.

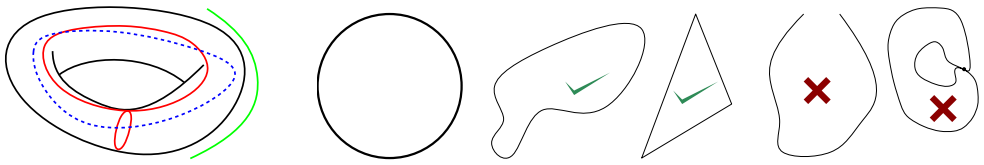


Figure 1.7: On the left, a torus and its topological properties. Then, a circle, two transformations that preserve its topology, and two transformations that do not (ripping and gluing respectively).

More generally, topology can be described as the set of properties of an object that do not change when we apply transformations without ripping nor gluing (see Figure 1.7). For instance, when such a transformation is applied to a circle, the topology (presence of one loop and one connected component) is unchanged, although the geometry is. Formally, the topological properties of an object are described by *homology groups*. Homology groups of a topological space (object) consist of a sequence of abelian groups $(H_k)_{k\geq 0}$ ($k = 0$ accounts for 0-dimensional topological properties of the space, that is connected components; $k = 1$ accounts for loops, $k = 2$ for cavities, and so on) whose generators identify "independent" topological features of the space. For instance, in the context of a torus (Figure 1.7), H_1 has two generators. These would correspond intuitively to the two red loops: any other loop on the torus is—in some sense—a linear combination of these two loops. We refer the interested reader to Appendix A or [Mun84] for a detailed presentation of homology theory.

Multi-scale topological descriptors. In practice however, the use of topology in machine learning is not straightforward. Most data acquisition methods are likely to produce objects with an extremely simple topology, such as *point clouds*. The homology groups of a point cloud $(H_k)_k$ are trivial for $k \geq 1$, while H_0 is the free abelian group with N generators, N being the cardinality of the point cloud. Homology alone is not able to reflect the underlying structure of such objects.

The core idea, whose foundations can be traced back to Morse's work [Mor40] and have been developed in [Fro92, Rob99, CFP01, ZC05], is to consider the topology of a space *at different scales*, and to look for topological properties that *persist* through scales. Let us consider as an example the point cloud represented on the left of Figure 1.8. Initially, it is (from a topological perspective) a collection of independent connected components. It is however intuitive to identify three loops of different sizes. Recording the topology of this object at different scales will allow us to detect these

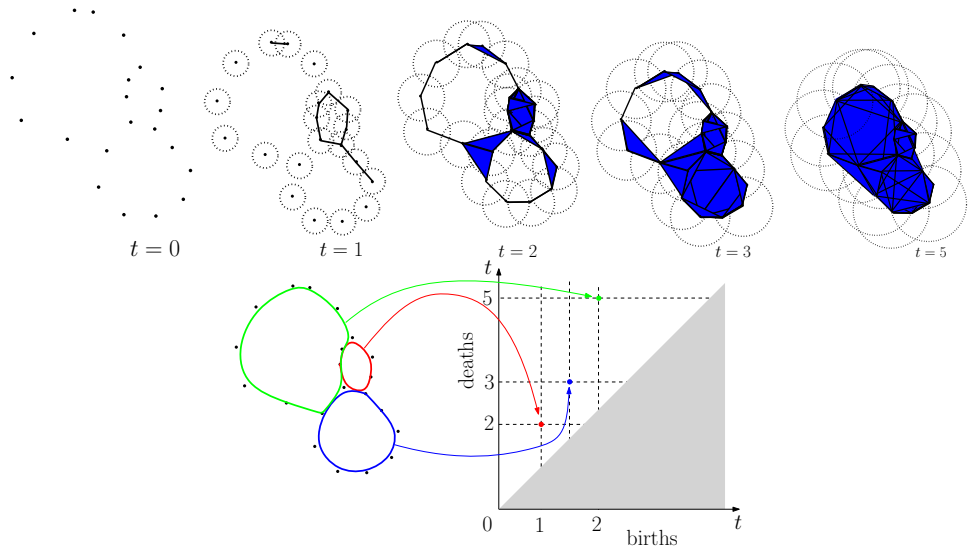


Figure 1.8: (Top) Topology of a point cloud at different scales for the Čech filtration. (Bot) The persistence diagram of the point cloud (restricted to loops).

loops. To introduce this notion of scale, an idea is to put balls centered on each point of the point cloud and to let them grow. The scale parameter $t \geq 0$ corresponds here to the radius of the balls. Therefore, starting with a point cloud X at $t = 0$, we build an increasing sequence of topological spaces $(X_t)_{t \geq 0}$, called a (Čech) *filtration*. For some *critical values* of t (see Figure 1.8) the topology[8] changes: loops appear (at $t = 1, t = 1.5, t = 2$ on the figure) or disappear (when they get "filled in", at $t = 2$, $t = 3$, $t = 5$ on the figure).

These critical values, called *birth times* and *death times*, are recorded in a topological descriptor: the *persistence diagram*, which is a collection of points in the plane. Each point of coordinate (b, d) in the diagram will account for a topological property (connected component, loop, caveat...) that appeared at scale $t = b$ and disappeared at scale $t = d$. Therefore, a point close to the diagonal (that is, such that $b \simeq d$) represents a topological property that appeared then almost instantly disappeared: this feature did not *persist* much. In contrast, a point away from the diagonal represents a topological feature that was recorded through a large interval of scales, generally considered as being more significant. Formally, a persistence diagram

[8]Here, for the sake of simplicity, we only consider loops.

can be described as a multi-set of points[9] supported on a half-plane

$$\Omega := \{(b, d) \in \mathbb{R}^2,\ d > b\}$$

or, equivalently, as a *point measure*, that is a Radon measure μ of the form

$$\mu = \sum_{x \in X} n_x \delta_x,$$

where X is a locally finite subset of Ω, n_x is an integer and δ_x denotes the Dirac mass located at $x \in X$.[10]

Of course, the situations in which topological data analysis can be applied are not reduced to point clouds and growing balls: the general theory is detailed in Appendix A. The important point is that one can transform a set of complex observations into a collection of persistence diagrams, from which we might consider producing a statistical analysis or performing a machine learning task.

1.2.3 Persistence diagrams and learning: challenges and limitations

In practical applications, incorporating persistence diagrams in modern machine learning toolboxes, such as neural networks, is unfortunately not straightforward.

Distance between persistence diagrams. In order to understand this, we must first mention that one can compute a distance between two persistence diagrams, which thus allows us to compare two objects adopting a topological viewpoint. The standard distance used to compare diagrams is called the *bottleneck distance*, and can be computed in the following way. For two diagrams X and Y, let $(x_1 \dots x_n)$ and $(y_1 \dots y_m)$ denote their respective points (note that we do not necessarily have $n = m$). The first step is to *transport*[11] each point x_i in the diagram X to a point y_j in Y or, possibly, to its orthogonal projection onto the diagonal (see Figure 1.9). The points y_j in Y which are not reached by a point x_i in X are also transported

[9] A set where points can be repeated.

[10] In the greatest generality, persistence diagrams can have points with infinite coordinates ($b = -\infty$ or $d = +\infty$) thus belonging to the extended half-plane; or points supported exactly on the diagonal ($b = d$). We will not consider such points in this thesis; this is discussed in Remark 2.1 and Remark 2.2.

[11] It is more standard to use the wording "matching", but we will prefer "transport" for reasons that will become obvious later on.

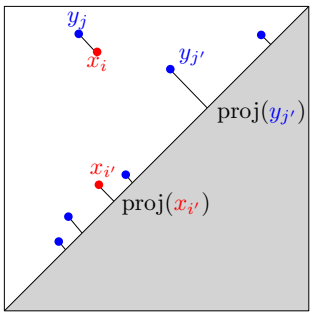


Figure 1.9: Two persistence diagrams. The bottleneck distance between these two diagrams is given, by definition, by the length of the longest edge drawn here.

on their respective projections onto the diagonal. We also enforce that each point x_i is transported to—at most—a single point y_j, and each y_j must be reached by—at most—a single x_i. Denote by (x_i, y_j) the fact that x_i is transported on y_j, and $(x_i, \partial\Omega)$ (resp. $(\partial\Omega, y_j)$) when x_i (resp. y_j) is transported to the diagonal. A *partial transport* between X and Y is given by a list $\dots (x_i, y_j)_{ij} \dots (x_i, \partial\Omega) \dots (\partial\Omega, y_j) \dots$, where each x_i, y_j must appear exactly once.

Of course, there exist many ways to transport a diagram X onto another diagram Y. In order to define a distance, one seeks for an *optimal* partial transport in the following sense: the *cost* of a partial transport is given by the longest distance traveled when performing the transport. A partial transport is optimal if it has minimum cost among all the possible partial transport between x and Y. The bottleneck distance is then defined as the minimal cost achieved by this optimal transport. Formally, let $\Gamma(X, Y)$ denote the set of all possible partial transports between X and Y, and define the bottleneck distance as

$$\mathbf{d}_\infty(X, Y) := \min_{\gamma \in \Gamma(X,Y)} \max_{(x,y)\in\gamma} \|x - y\|.$$

The choice of such a distance is motivated by algebraic considerations detailed in Appendix A and Section 2.1. This distance has the nice property of being *stable*, in the sense that initially close objects will always have diagrams that are close in the bottleneck distance (Figure 1.10). For instance, in the context of two point clouds P, P' and corresponding diagrams μ, μ' built using the Čech filtration (Figure 1.8), it yields

$$\mathbf{d}_\infty(\mu, \mu') \leq d_H(P, P'),$$

where d_H stands for the Hausdorff distance between the point clouds (see Example A.5 for details).

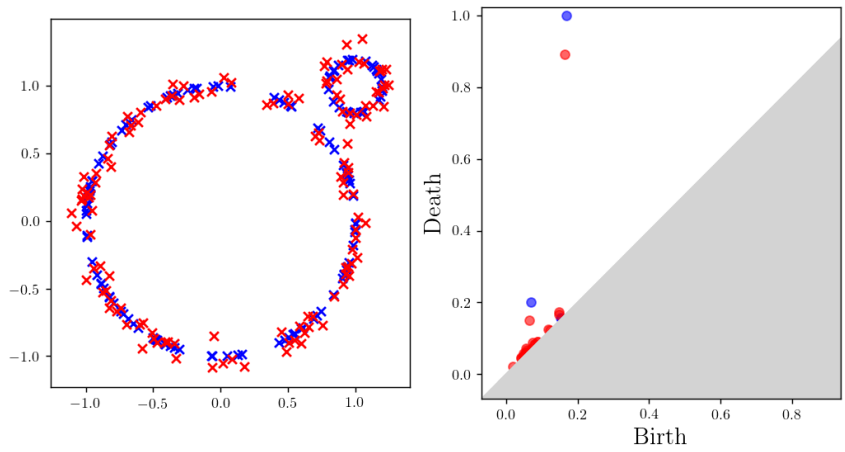


Figure 1.10: Stability of persistence diagrams for the bottleneck distance. Here, the bottleneck distance between the two diagrams is controlled by the Hausdorff distance between the point clouds.

Limitations. Given a collection of diagrams, and thanks to the bottleneck distance, it becomes possible to compute pairwise distances between diagrams. It makes it possible to use these diagrams as the input of some simple machine learning algorithms, such as the *nearest neighbor* algorithm or the metric *multi-dimensional scaling* (MDS) one, as these algorithms only require to be able to compute distances between our observations (here, the persistence diagrams).

However, most modern machine learning algorithms ask for a *linear* structure on the space in which the data live. Unfortunately, the *space* of persistence diagrams is not linear, which essentially means that there is no simple way[12] to define the sum of two diagrams, or the multiplication of a diagram by a real number. Therefore, it is not possible to incorporate faithfully persistence diagrams in a modern machine learning pipeline, compromising their use in practical applications.

An alternative is to *embed* the space of persistence diagrams in a linear space, on which it will be possible to easily perform standard learning tasks. It means that we look for a map Φ which transforms the diagrams into *vectors*, namely an element of a linear space. Such a map is called a *vectorization*.

Various vectorizations have been proposed, with significant success in

[12]That will be compatible with the bottleneck distance

applications [Bub15, AEK+17, CCO17]. Nonetheless, using a vectorization raises some important questions:

- The choice of the vectorization is arbitrary. There is currently no heuristic available to choose, among all the possible vectorizations, which one will be adapted to perform well on a given learning task.

- It seems impossible[13] to have, in the greatest generality, a *coarse* embedding, that is an embedding for which it holds, for any pair of diagrams X, Y, and two non-decreasing map $\rho_1, \rho_2 : [0, +\infty) \to [0, +\infty)$ with $\rho_1(t) \to +\infty$ when $t \to +\infty$,

$$\rho_1(\mathbf{d}_\infty(X, Y)) \leq \|\Phi(X) - \Phi(Y)\| \leq \rho_2(\mathbf{d}_\infty(X, Y)),$$

  which would ensure that close diagrams have close representations, and vice-versa.

- Finally, let us recall that diagrams benefit from their interpretability. Using a vectorization, and doing the computations in the corresponding linear space, would lose this interpretability. For instance, given a map Φ and a set of diagrams $X_1 \dots X_n$, one can compute a mean in the linear space by setting $M = \frac{1}{n} \sum_{i=1}^{n} \Phi(X_i)$. However, it is important to note that M has, *a priori*, nothing to do with a notion of mean in the space of persistence diagrams as it may not even be in the image of Φ. This holds more generally for any computation performed in the linear space (which could lead to clustering, etc.).

Challenges. While studying vectorizations of the space of persistence diagrams remains an important research topic, we propose in this manuscript to work *in the space of persistence diagrams directly*. The goal is thus to answer the following question:

"How can we do statistics in the space of persistence diagrams?"

To answer this question, it is important to understand the geometric properties of this space. This should allow us to provide a solid theoretical framework in which statistical tools will be properly defined. Finally, these statistical tools must be usable in practice, that is we must be able to compute (or at least estimate) them.

[13] This problem is still open, but all recent results seem to point this way [BV18, CB18, BLPY19, BW19, Wag19].

1.2.4 Contributions and outline

To provide some solutions to this problem, we will rely on another field of applied mathematics: *optimal transport* theory, presented in Section 2.2. It is one of the most important contributions of this manuscript: establishing a formal connection between the geometry of the space of persistence diagrams and models used in optimal transport. This connection between these two fields is highly useful: optimal transport is a well-developed domain both from a theoretical and computational perspective and has many applications in statistical and machine learning. It turns out that most of its tools can be transposed or adapted to deal with persistence diagrams, which will allow us to obtain new theoretical results regarding the use of persistence diagrams in statistics and learning, but also to provide various algorithms and computational tools that allow us to address some difficult problems in topological data analysis, such as the estimation of Fréchet means, quantization, vectorization learning, to name a few.

In general and unless stated differently, results (theorems, propositions, etc.) for which we provide a proof are consequences of my work or joint work with my Ph.D. advisors or collaborators: Mathieu Carrière, Frédéric Chazal, Marco Cuturi, Vincent Divol, Yuichi Ike, Steve Oudot, Martin Royer, and Yuhei Umeda.

Outline and relations with the research productions during the Ph.D. thesis.

Chapter 2: Background. This chapter presents the two fields involved in this work: topological data analysis (Section 2.1) and optimal transport (Section 2.2). It aims to provide a concise introduction to both domains and a presentation of the different problems tackled in this manuscript. Note that a presentation of the algebraic construction of persistence diagrams can be found in Appendix A.

The contributions are organized into two separate parts, according to their theoretical or applied (roughly speaking, results leading to an implementation) nature.

Part I — Theory.

Chapter 3: an optimal transport framework for persistence diagrams. This chapter presents the theoretical foundations of the manuscript. It essentially corresponds to Section 3 in [DL19], under minor revisions for the *Journal of Applied and Computational Topology*. We present here how

the standard metrics used to compare persistence diagrams can be reformulated as optimal *partial* transport problems, and the new theoretical results that come out of it.

Chapter 4: A theoretical study of Fréchet means for persistence diagrams. This chapter is dedicated to the study of Fréchet means (or barycenters) of persistence diagrams from a theoretical perspective. It essentially corresponds to Section 4 in [DL19]. We prove the existence of Fréchet means in great generality, and establish a close link between those and the "Wasserstein barycenters", their counterparts in standard optimal transport theory.

Part II — Applications.

Chapter 5: Fast algorithms for the estimation of Fréchet means of persistence diagrams. We propose an algorithm to estimate these means, which is useful especially on large scale problems. This approach has been published in the proceedings of the international conference on *Neural Information Processing Systems, 2018* [LCO18].

Chapter 6: Linear representation of persistence diagrams. This is dedicated to the study of *linear* vectorizations of the space of persistence diagrams, for which we provide an exhaustive characterization. The theoretical part (Section 6.1) comes from [DL19, §5.1]. We then provide an application in machine learning, where we introduce PersLay, a neural network layer[14] devoted to learning optimal linear vectorizations to solve a given learning task. This work has been published in the proceedings of the international conference on *Artificial Intelligence and Statistics, 2020* [CCI+19].

Chapter 7: Complementary examples. This last chapter gathers other algorithms developed for persistence diagrams. In Section 7.1, we showcase the strength of the formalism developed in Chapter 3 in the context of random persistence diagrams. We prove convergence and stability results of the probabilistic counterpart of persistence diagrams. Section 7.2 is dedicated to the *quantization* of persistence diagrams, a useful tool that has benefits when incorporating persistence diagrams in machine learning

[14] including a publicly available implementation and incorporation to the `Gudhi` library.

pipelines. Finally, Section 7.3 proposes a simple and efficient way to estimate a shift-invariant distance between persistence diagrams.

Code and contributions to the `Gudhi` library. Most of the methods presented in Part II have been—or will be—incorporated into the topological data analysis library `Gudhi`.

Chapter 2

Background

Abstract

This chapter presents some background material coming from the two main research fields involved in this work: topological data analysis (Section 2.1) and optimal transport (Section 2.2). They are essentially presented in a way that serves the contents of Part I and Part II. Therefore, they do not aim at being exhaustive but rather at providing a decent state-of-the-art of the topics covered by this thesis.

2.1 Topological Data Analysis

This section is dedicated to the presentation of our objects of interest: persistence diagrams. A detailed algebraic construction of persistence diagrams—which is not required to understand the vast majority of this manuscript—can be found in Appendix A. We also refer the interested reader to [EH10, Oud15] for a thorough description. Subsection 2.1.1 below gives a concise presentation of these notions. Subsections 2.1.2 and 2.1.3 introduce various metrics between persistence diagrams and some of the properties of the resulting metric spaces.

2.1.1 Persistent homology in a nutshell

Let $\mathbb{X}$ be a topological space, and $f\colon \mathbb{X} \to \mathbb{R}$ a real-valued continuous function. The *t-sublevel set* of $(\mathbb{X}, f)$ is defined as

$$\mathcal{F}_t = \{x \in \mathbb{X},\ f(x) \leq t\}.$$

Making t increase from $-\infty$ to $+\infty$ gives an increasing sequence of sublevel sets $(\mathcal{F}_t)_t$ called the *filtration* induced by f. To the increasing family of topological spaces $(\mathcal{F}_t)_t$ corresponds a *persistence module* $(\mathbb{V}_t)_t$, which is a family of $\mathbb{K}$-vector spaces (for some fixed field of coefficients $\mathbb{K}$) equipped with linear maps $v_s^t : \mathbb{V}_s \to \mathbb{V}_t$ for $s \leq t$ which are induced by the inclusion $\mathcal{F}_s \subset \mathcal{F}_t$ (see Appendix A for details). In particular, for $I = [b, d] \subset \overline{\mathbb{R}}$ an interval—with $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$—, the *interval module* $\mathbb{I}_{[b,d]}$ is the persistence module defined by $\mathbb{I}_{[b,d]_t} = \mathbb{K}$ for $t \in I$, and $\{0\}$ otherwise, while $v_s^t = \mathrm{id}_{\mathbb{K}}$ if $s, t \in I$ and $v_s^t = 0$ otherwise. Under mild assumptions (e.g. $\mathrm{rk}(v_s^t)$ is finite for any s, t), a persistence module $(\mathbb{V}_t)_t$ can be *decomposed* uniquely as a direct sum of interval modules, which reads

$$\mathbb{V} = \bigoplus_{j \in J} \mathbb{I}_{[b_j, d_j]},$$

for some locally finite family of intervals $([b_j, d_j])_{j \in J}$. Therefore, $\mathbb{V}$ is entirely described by a multiset[1] of points $\{(b, d)\}$ that belong to the extended upper half-plane

$$\overline{\mathbb{R}}^2_{\geq} := \{(b, d) \in \overline{\mathbb{R}}^2,\ d \geq b\},$$

called the barcode, or *persistence diagram*, of $(\mathbb{X}, f)$, also denoted by $\mathrm{Dgm}(\mathbb{X}, f)$ or simply $\mathrm{Dgm}(f)$ if there is no ambiguity.

Intuitively, an interval $[b, d]$ appearing in $\mathrm{Dgm}(\mathbb{X}, f)$ accounts for the presence of a topological feature (connected component, loop, cavity, etc.) that appears at scale $t = b$ in $(\mathcal{F}_t)_t$ and disappears (gets "filled") at scale $t = d$. Equivalently, a persistence diagram can be described as a *point measure* supported on $\overline{\mathbb{R}}^2_{\geq}$, that is a measure μ of the form

$$\mu = \sum_{x \in X} n_x \delta_x,$$

where X is a locally finite subset of $\overline{\mathbb{R}}^2_{\geq}$, n_x is an integer, and δ_x denotes the Dirac mass supported on $\{x\}$. This measure-theoretic perspective on persistence diagrams is motivated in [CDSGO16] (see also Theorem A.2) and has huge benefits when studying the properties of the space of persistence diagrams, as showcased in Chapter 3.

2.1.2 Metrics between persistence diagrams

In order to use persistence diagrams in practice, we must—at the very least—be able to compare them using a metric. The standard metric used

[1] Points can be repeated in the family. The number of times a point appears in a the multiset is called its *multiplicity*.

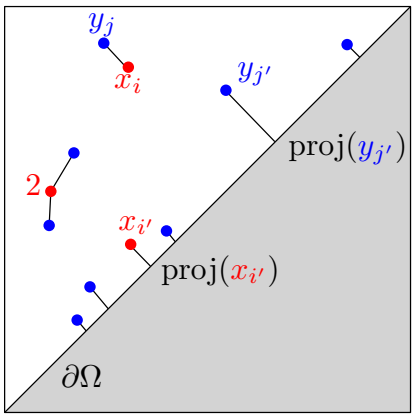


Figure 2.1: Partial matching between two diagrams. The '2' indicates that the point has multiplicity 2.

to compare persistence diagrams is the bottleneck distance, defined below. It is a *partial matching* distance between the points of the diagrams, counted with multiplicity. We recall that $\overline{\mathbb{R}}^2_{\geq}$ denotes the extended closed upper half-plane. Let also $\partial\Omega$ denote the diagonal (this notation will be clarified later on, in Definition 2.4):

$$\partial\Omega := \{(t,t),\ t \in \mathbb{R}\}.$$

Definition 2.1. Let $\mu = \sum_i \delta_{x_i}$ and $\nu = \sum_j \delta_{y_j}$ be two persistence diagrams. Let $X = (x_i)_i \subset \overline{\mathbb{R}}^2_{\geq}$ and $Y = (y_j)_j \subset \overline{\mathbb{R}}^2_{\geq}$ be the points of the diagrams, counted with multiplicity.

A *partial matching* γ is a subset of $(X \cup \partial\Omega) \times (Y \cup \partial\Omega)$ such that any $x_i \in X$ (resp. y_j in Y) appears exactly once as a first (resp. second) coordinate in γ (see Figure 2.1).

Let $\Gamma(\mu,\nu)$ denote the set of partial matchings between μ and ν.

The *cost* of a matching $\gamma \in \Gamma(\mu,\nu)$ is defined as the length of the longest edge in the matching, that is $\max\{\|x-y\|_\infty, (x,y) \in \gamma\}$.

The *bottleneck distance* between the two diagrams is then defined as the minimal cost that can be achieved by such a matching, that is

$$\mathbf{d}_\infty(\mu,\nu) = \inf_{\gamma\in\Gamma(\mu,\nu)} \max_{(x,y)\in\gamma} \|x-y\|_\infty. \tag{2.1}$$

A partial matching that realizes this infimum is said to be *optimal*.

Remark 2.1. We start by noting some important remarks about these definitions.

- In the above definition, x (resp. y) belongs either to X or to the diagonal $\partial\Omega$. If $x \in X$ has finite coordinates and is matched to $\partial\Omega$, that is it belongs to a tuple $(x, y) \in \gamma$ with $y \in \partial\Omega$, then one can always assume that y is the orthogonal projection of x onto the diagonal as this would only reduce the cost.

- Similarly, although a partial matching might theoretically contain couples $(x, y) \in \partial\Omega \times \partial\Omega$, these can be removed when looking for an optimal matching as this does not increase the cost.

- Persistence diagrams might contain points that belong to $\partial\Omega$, in which case they can always be matched to the diagonal (with themselves) with a null cost. In particular, the bottleneck distance is only a pseudo-metric at this stage. This aspect will be clarified in the following subsection.

- In the standard definition of the bottleneck distance, $\|\cdot\|_\infty$ stands for the L^∞ norm in $\mathbb{R}^2$, although it can be replaced with any norm $\|\cdot\|_q$ with $1 \leq q \leq \infty$. Adapting the following results is straightforward due to the equivalence of norms in $\mathbb{R}^2$. When q does not play any role, we simply use the notation $\|\cdot\|$.

Remark 2.2. Persistence diagrams may contain points with coordinates of the form $x = (b, +\infty)$ (resp. $(-\infty, d)$). The set of such points is called the *essential part* of the diagram. When comparing two persistence diagrams, if the cardinalities of their essential parts differ, the cost of any matching (thus the bottleneck distance) is $+\infty$.

Otherwise, the points of the respective essential parts must be matched together for the cost to be finite, and it is straightforward to observe that optimally matching those points consists in sorting the points with respect to their first (resp. second) coordinate and then using the increasing matching. It mostly means that the behavior of essential parts is obvious and independent of the rest of the diagrams.

In the rest of this manuscript, we will therefore only consider persistence diagrams with empty essential parts. Most of our theoretical and practical results can be easily adapted to take the essential parts into account separately.

Using the bottleneck distance to compare persistence diagrams is motivated by algebraic considerations detailed in Appendix A. In addition, this distance is supported by the following strong stability result, due to [CSEH07, CCSG+09a], which states that two close filtrations induce close persistence diagrams in the bottleneck distance.

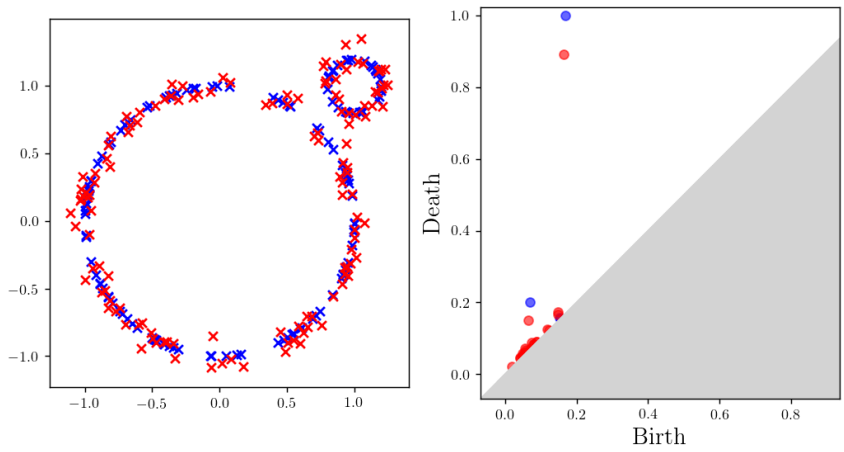


Figure 2.2: Stability of the bottleneck distance for the Čech diagrams with respect to the Hausdorff distance between the point clouds.

Theorem 2.1 (*Stability theorem*)**.** *Let $f, g : \mathbb{X} \to \mathbb{R}$ be two* q-tame *functions (they induce q-tame persistence modules, see Appendix A), and let* $\mathrm{Dgm}(f), \mathrm{Dgm}(g)$ *be their respective persistence diagrams. Then,*

$$\mathbf{d}_\infty(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_\infty, \tag{2.2}$$

that is the operator $f \mapsto \mathrm{Dgm}(f)$ is 1-Lipschitz, that is:

Example A.5 in Appendix A provides a simple application in the context of diagrams built on top of point clouds. Briefly, if P and P' are two finite point clouds in $\mathbb{R}^d$, one has

$$\mathbf{d}_\infty(\mathrm{Dgm}(d_P), \mathrm{Dgm}(d_{P'})) \leq d_H(P, P'),$$

where d_P (resp. $d_{P'}$) denotes the distance to P (resp. P'), and $d_H(P, P')$ is the Haussdorf distance between P and P': similar point clouds have similar persistence diagrams. This fact is illustrated in Figure 2.2.

Other metrics of interest. From a theoretical perspective, the bottleneck distance is a natural choice to compare persistence diagrams. However, when used for statistical and machine learning applications, it has some important drawbacks. Indeed, given two diagrams μ and ν and an optimal partial matching γ^*, the bottleneck distance between μ and ν only depends on the longest edge in γ^*. It means that

- Any matching, as long as it does not change the length of the longest edge in γ^*, will also be optimal. There are therefore plenty of optimal matchings between two diagrams in general.
- Similarly, a slight perturbation in the points in μ, as long as it does not change the length of the longest edge of an optimal matching, will not be detected in the bottleneck distance.

In statistical and machine learning applications, discriminating information might lie in these undetected properties. Therefore, in practice, a variation of the bottleneck distance is usually preferred.

Definition 2.2. Let $1 \leq p < \infty$. The *p-th distance* between two diagrams μ and ν is defined as

$$\mathbf{d}_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \sum_{(x,y)\in\gamma} \|x - y\|_q^p \right)^{\frac{1}{p}}. \tag{2.3}$$

In (2.3), $\|\cdot\|_q$ stands for any q-norm, with $1 \leq q \leq \infty$. Natural choices are $q = 2$ (Euclidean ground metric), $q = p$, and $q = \infty$. This choice has barely any impact on the theoretical results presented in the manuscript. The choice $q = p$ might have benefits in numerical problems (see Chapter 5) as it leads to a *separable cost*.

While $\mathbf{d}_p \to \mathbf{d}_\infty$ as $p \to \infty$, it improves on the drawbacks mentioned above: the optimal partial matching between two diagrams (for the $\mathbf{d}_p$ metric) is generically unique, its cost involves *all* the edges of the matching (making this distance able to account for small perturbations), etc. The p-th diagram distance also satisfies a (weaker) stability property.

Definition 2.3. A metric space $\mathbb{X}$ is said to imply a bounded degree-q *total persistence* if there exists a constant $C_{\mathbb{X}}$ such that for every tame Lipschitz function $f : \mathbb{X} \to \mathbb{R}$, with $\mathrm{Lip}(f) \leq 1$, one has $\mathrm{Pers}_q(\mathrm{Dgm}(f)) \leq C_{\mathbb{X}}$, where

$$\mathrm{Pers}_q(\mathrm{Dgm}(f)) := d_q(\mathrm{Dgm}(f), \mathbf{0}) = \left(\sum_{x\in\mathrm{Dgm}(f)} \|x - \partial\Omega\|^q \right)^{\frac{1}{q}}, \tag{2.4}$$

with $\mathbf{0}$ denoting the empty diagram.

Theorem 2.2 ([CSEHM10])**.** *Let $\mathbb{X}$ be a triangulable, compact metric space that implies bounded degree-q total persistence for some $q \geq 1$, and*

$f, g : \mathbb{X} \to \mathbb{R}$ two tame Lipschitz functions. There exists a constant $C_{\mathbb{X}}$ that depends on $\mathbb{X}$ such that for all $p \geq q$,

$$\mathbf{d}_p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq C_{\mathbb{X}}(\max \mathrm{Lip}(f)^q, \mathrm{Lip}(g)^q)^{\frac{1}{p}} \|f - g\|_\infty^{1-\frac{q}{p}}. \quad (2.5)$$

Remark 2.3. The p-th distances between persistence diagrams are generally referred to as "Wasserstein distances" between persistence diagrams, a terminology that comes from optimal transport theory, see Section 2.2 below. As optimal transport will take a big place in this manuscript, we use the word "Wasserstein" to denote standard optimal transport metrics to avoid confusion.

2.1.3 Statistics and machine learning with persistence diagrams

Let us end this section by introducing important notions for the development of this manuscript. Namely, we start by defining and stating the main properties of the *space of persistence diagrams*. We then present some statistical and machine learning problems that take place in this space.

Recall (Remark 2.2) that we only consider diagrams whose points have finite coordinates.

The space of persistence diagrams

We can now compare persistence diagrams using either the bottleneck or the p-th distances (for the sake of simplicity, we use the notation $\mathbf{d}_p$ for both metrics). It invites us to consider the space of all possible diagrams equipped with such metrics.

First, note that the wording "metrics" is actually improper as the bottleneck and the p-th distances are only *pseudo-metrics* in the sense that one can have $\mathbf{d}_p(\mu, \nu) = 0$ while $\mu \neq \nu$. This is due to the points in the diagrams that lie on the diagonal $\partial\Omega$, as those can be matched to the diagonal with a cost of 0. This leads to a natural equivalence between diagrams: two persistence diagrams are equivalent if their off-diagonal parts coincide. The resulting equivalence classes will still be referred to as persistence diagrams. From a practical perspective, this is equivalent to simply removing the points lying on the diagonal. We can now give our core definitions.

Definition 2.4. Let us define the open upper half-plane

$$\Omega := \{(b, d) \in \mathbb{R}^2, b < d\}, \quad (2.6)$$

and note that the diagonal $\partial\Omega := \{(t,t), t \in \mathbb{R}\}$ is its boundary.

A persistence diagram is a *point measure* supported on Ω, that is a measure of the form

$$\mu = \sum_{x \in X} n_x \delta_x,$$

where X is a locally finite subset of Ω, and $n_x \in \mathbb{N}$.

The set of all persistence diagrams will be denoted by $\mathcal{D}$.

Essentially, this definition is consistent with the algebraic one developed in the previous subsection, where diagrams are understood without their essential parts (as it is treated separately in a trivial manner) and points on the diagonal are removed (as those cannot be detected by the $\mathbf{d}_p$ distances). These assumptions allow recovering true metric spaces on which we can provide a relevant analysis.

Definition 2.5. Let $\mu \in \mathcal{D}$, and $1 \leq p \leq \infty$. The *total persistence* of μ of parameter p is

$$\mathrm{Pers}_p(\mu) := \mathbf{d}_p(\mu, \mathbf{0}) \in [0, +\infty], \tag{2.7}$$

where $\mathbf{0}$ denotes the empty diagram.

The *space of persistence diagrams* of parameter p is

$$\mathcal{D}^p := \{\mu \in \mathcal{D}, \mathrm{Pers}_p(\mu) < \infty\}. \tag{2.8}$$

This definition was introduced in the seminal papers [MMH11, TMMH14]. The assumption $\mathrm{Pers}_p(\mu) < \infty$ in (2.8) ensures that for any $\mu, \nu \in \mathcal{D}^p$, the quantity $\mathbf{d}_p(\mu, \nu)$ is finite. It makes $(\mathcal{D}^p, \mathbf{d}_p)$ a metric space. Note the convenient expression, for $1 \leq p < \infty$,

$$\mathrm{Pers}_p(\mu) = \sum_{x \in X} n_x \|x - \partial\Omega\|^p = \mu(\|\cdot - \partial\Omega\|^p), \tag{2.9}$$

where $\|x - \partial\Omega\|$ denotes the distance between a point $x \in \Omega$ and (its orthogonal projection onto) the diagonal $\partial\Omega$. When $p = \infty$, note that $\mathrm{Pers}_\infty(\mu) = \sup_{x \in X} \|x - \partial\Omega\|$.

The space of persistence diagrams, at least for finite p, satisfies some fundamental properties.

Proposition 2.3 ([MMH11, Theorems 6 and 12])**.** *Let $1 \leq p < \infty$. The space $\mathcal{D}^p$ is a Polish metric space (i.e. is complete and separable in the metric $\mathbf{d}_p$).*

It implies that one can define probability distributions supported on $\mathcal{D}^p$, the starting point of any statistical analysis.

Remark 2.4. It is important to keep in mind that we allow for diagrams with possibly infinitely many points (although they must remain *locally* finite). This assumption is required for $\mathcal{D}^p$ to be complete (see Example 2.1). Note that one could also recover a complete space by enforcing a uniform bound N on the number of points (mass) on the diagrams (i.e. considering the space of diagrams with at most N points). This assumption is however not suitable nor realistic in some situations. For instance,

- The number of points in a diagram is not a stable quantity in applications: slight modifications in the input might drastically change the number of points (by creating many points close to the diagonal, see for instance Figure 2.2). This phenomenon also occurs when sampling points (uniformly, iid) on a (unknown) shape X and looking at the Čech diagram of the sample. The Hausdorff distance between the sample and the shape X converges to 0 as the sample size $n \to \infty$, so does the diagram[2] thanks to the stability theorem. However the number of points in the diagram reflecting homology of dimension 1 (loops) explodes (see Figure 2.3): in this case, working with diagrams with at most N points would refrain us from producing a relevant analysis of convergence.

- In theory, some objects might have diagrams with infinitely many points. For instance, the (random) persistence diagram μ built from the sublevel sets of a Brownian motion has (almost surely) infinitely many points, while for any $\varepsilon > 0$, we have $\mathrm{Pers}_{2+\varepsilon}(\mu) < \infty$ thus $\mu \in \mathcal{D}^{2+\varepsilon}$. In practice, if we want to compute the diagrams of approximated counterparts of such objects, we might expect the number of points in our diagrams to explode as the approximation error goes to 0.

Nonetheless, the space of persistence diagrams with uniformly bounded mass has some interesting structure and can provide a relevant framework in some cases. It will be studied in detail in Section 3.2.

Example 2.1. Let $\mu_n = \sum_{k=1}^{n} \delta_{x_k}$, where $x_k = \left(0, 2^{-k}\right)$. The sequence $(\mu_n)_n$ is a Cauchy sequence for the $\mathbf{d}_p$ metric, as for $n < m$ we have $\mathbf{d}_p(\mu_n, \mu_m) \leq 2^{(m-n)}$. However, the limit object is $\sum_{k \geq 1} \delta_{x_k}$, which has infinitely many points. See Figure 2.4 for an illustration.

Remark 2.5. The space of persistence diagrams equipped with the bottleneck metric ($p = \infty$) is more intricate. Indeed, Eq. (2.8) does not define

[2] for the bottleneck distance. A slightly more sophisticated analysis is required for the $\mathbf{d}_p$ metric with $p < \infty$.

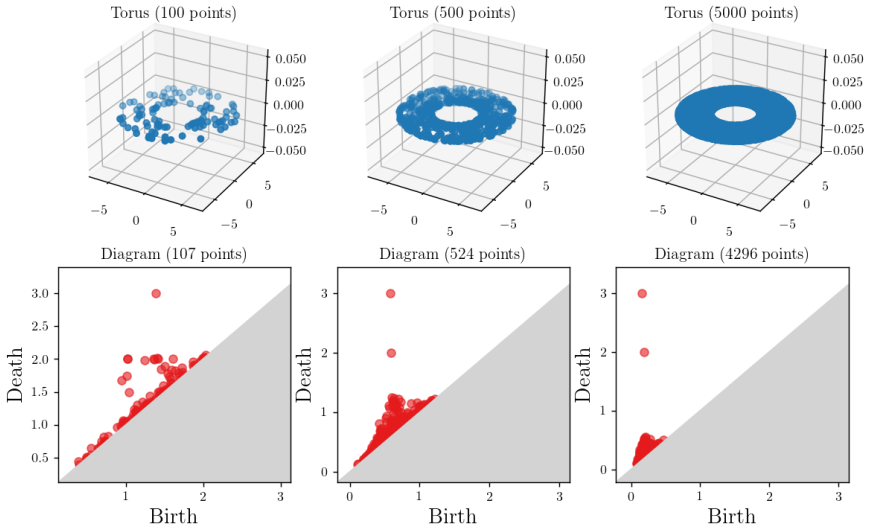


Figure 2.3: *(Top)* A n-sample on a Torus for $n = 100, 500, 5000$. *(Bottom)* The corresponding persistence diagrams in H_1; the title of each subplot indicates the number of points in the corresponding diagram. Although the diagram converges (in the bottleneck distance) to the diagram of the underlying torus (which has 2 points, one for each generating loop), the number of points in the diagram diverges.

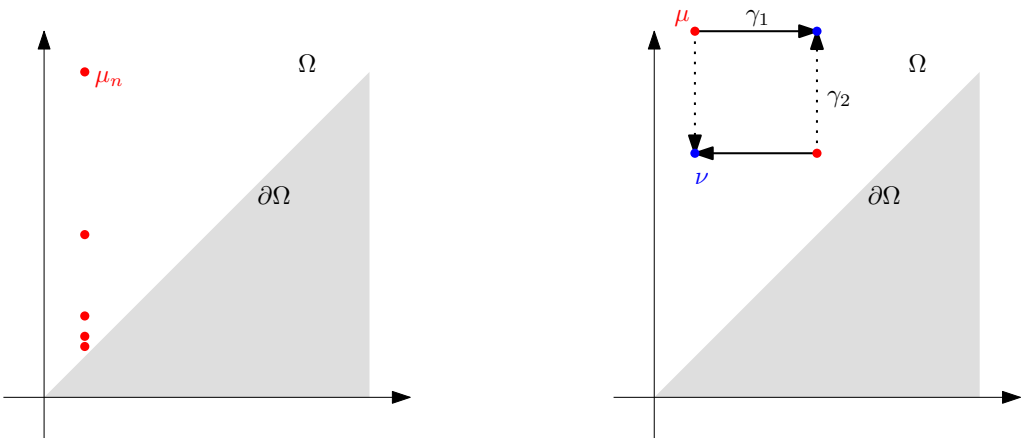


Figure 2.4: *(Left)* A Cauchy sequence of finite diagrams that converges to a diagram with infinitely many points. *(Right)* An example where the geodesic between two diagrams is not unique.

a separable metric space in this context (see Remark 3.8 in Chapter 3 or [BV18, Thm. 4.20]). A careful study of this space is provided in Section 3.3.

Thanks to the work [TMMH14], we know more about the geometric structure of $\mathcal{D}^p$ in the case $p = 2$ (although their analysis easily extends to any finite p). See appendix C for a reminder of important notions in metric geometry (*geodesic* curves, *curvature*, etc.).

Proposition 2.4. *The space $\mathcal{D}^2$ is a geodesic space. It is a non-negatively curved Alexandrov space.*

In Chapter 3, we will provide a new analysis of this space based on optimal transport theory. We will recover some results of [MMH11, TMMH14] using different tools and prove new properties of the space by adopting a measure-based formalism. This formalism will allow us to address some of the problems we present below.

Fréchet means of persistence diagrams

Assume we have a set of observations $X_1 \dots X_n$ from which we compute (for some filtration, etc.) a set of diagrams $\mu_1 \dots \mu_n \in \mathcal{D}^2$ (i.e. with $p = 2$). In order to summarize the topological information contained in the $(\mu_i)_i$s, the most natural attempt is to define and compute a notion of mean, or *barycenter*. However, the space of persistence diagrams, equipped with any of the $\mathbf{d}_p$ metrics, does not have a linear structure. Therefore, the notion of arithmetic mean used routinely in Euclidean spaces does not make sense in this space.

However, a characterization of the arithmetic mean in metric spaces is that it is the unique minimizer of $x \mapsto \sum_{i=1}^n \|x - x_i\|^2$. The latter formula only involves metrics and thus still makes sense in the context of metric spaces, in particular $\mathcal{D}^2$.

Definition 2.6. Let $\mu_1 \dots \mu_n$ be n persistence diagrams. A *Fréchet mean* of $\mu_1 \dots \mu_n$ is any minimizer, should it exist, of the *energy functional*

$$\mathcal{E} : \nu \mapsto \frac{1}{n} \sum_{i=1}^{n} \mathbf{d}_2(\nu, \mu_i)^2. \tag{2.10}$$

More generally, if P is a probability distribution supported on $\mathcal{D}^2$, the energy functional reads

$$\mathcal{E} : \nu \mapsto \int \mathbf{d}_2(\nu, \mu)^2 \mathrm{d}P(\mu). \tag{2.11}$$

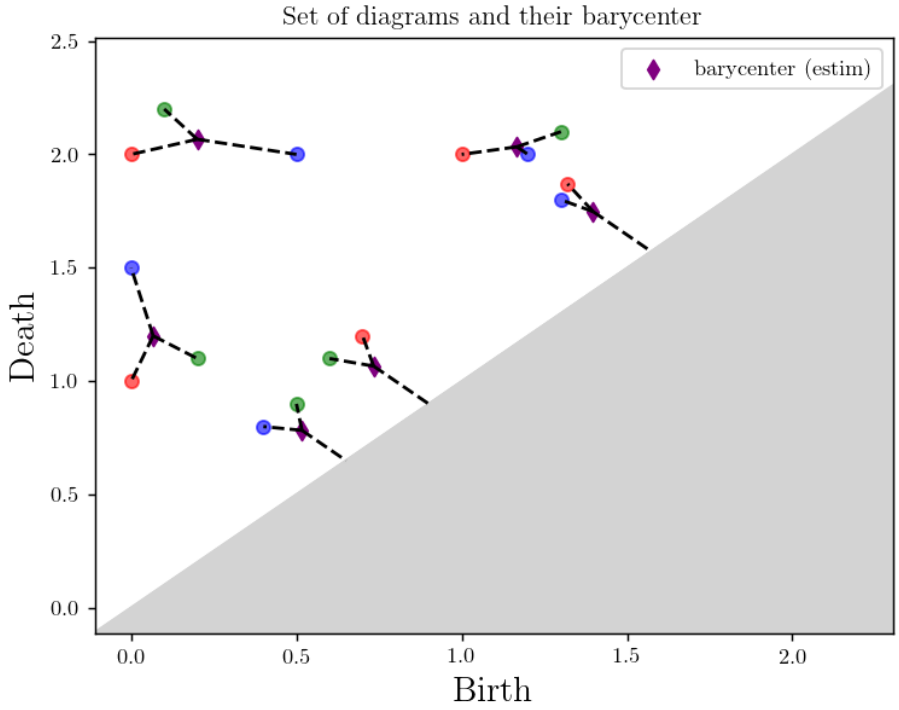


Figure 2.5: An example of Fréchet mean of three persistence diagrams (red, blue, green) with uniform weights.

We commonly add the assumption $\mathcal{E}(\mathbf{0}) < \infty$, where $\mathbf{0}$ denotes the empty diagram.

The study of Fréchet means in $\mathcal{D}^2$ was initiated in [MMH11, TMMH14], where authors proved the existence of Fréchet means in the finite case (2.10) and in the general case (2.11) under some assumptions on the distribution P. Note that Fréchet means in $\mathcal{D}^2$ might not be unique, but generically are [Tur13]. From a computational perspective, [TMMH14] provided an algorithm to estimate a minimizer of (2.10), see Figure 2.5 for an illustration. However, this algorithm can return possibly bad estimates (see Chapter 5) and does not scale up on large samples.

Once again, by making use of the optimal transport formalism developed in Chapter 3, we will show how we can improve on this result. From a theoretical perspective, we will prove in Chapter 4 the existence of minmizers of (2.11) for any distribution P (satisfying a necessary finite moment assumption), extending the previous result of [MMH11]. From a practical perspective, making use of recent progress in computational optimal trans-

port [PC17], we will propose in Chapter 5 a *convex* and *scalable* relaxation strategy to estimate minimizers of (2.10).

Persistence diagrams in random settings

A natural approach in statistics is to assume that our observations are generated, independently, by some underlying random process. In the context of topological data analysis, it leads to considering the *random* topological properties of this process.

For instance, consider the following: let ξ be a probability distribution supported on $\mathbb{R}^d$. Fix $n \in \mathbb{N}$. Let $\mathbb{X}_n = (X_1 \dots X_n)$ be a n-sample of law ξ (that is, $\forall i, X_i \sim \xi$, i.i.d.). One can build, using the Čech (or Rips) filtration (Example A.2) a *random persistence diagram* $\mathrm{Dgm}(\mathbb{X}_n)$.

The topology of this sampling process is encoded by the (random) points of $\mathrm{Dgm}(\mathbb{X}_n)$. As such, for any compact set $K \subset \Omega$, the quantity $\mathrm{Dgm}(\mathbb{X}_n)(K)$ encodes the (random) number of points of $\mathrm{Dgm}(\mathbb{X}_n)$ that fall into K; it is an integer-valued random variable. A natural *deterministic* quantity to consider is then the expectation of this random variable. It defines the *expected persistence diagram* (EPD) $\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)]$ of the random process $\mathbb{X}_n$:

$$\text{For all } K \subset \Omega, \quad \mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)](K) := \mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)(K)]. \tag{2.12}$$

Obviously, $\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)]$ is not integer-valued. These objects were introduced by Chazal and Divol in [CD18], where the authors proved that under regularity condition on the sampling process ξ, the EPD is a Radon measure supported on Ω that admits a Lebesgue density supported on Ω.

How to study such objects? EPDs can be interpreted in a similar way to classical persistence diagrams: areas of high density indicate that, on average, random realizations of $\mathrm{Dgm}(\mathbb{X}_n)$ have points in this area. However, the combinatorial definition of the diagram metrics (2.1) and (2.3) makes them unable to handle measures with continuous support. Being able to compute distances involving expected persistence diagrams (or other measures with a continuous support) is of interest, as it would allow us to address interesting questions from a statistical viewpoint, such as

- How does the expected persistence diagram depend on the underlying law ξ generating the points?
- Given m realizations of $\mathbb{X}_n$, denoted by $(X_n^{(i)})_{i=1}^m$, a natural estimator of $\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)]$ is $\mu_n^m := \frac{1}{n}\sum_{i=1}^m \mathrm{Dgm}(X_n^{(i)})$. Indeed, for any compact

$K \subset \Omega$, $\mu_n^m(K) = \frac{1}{n}\sum_i \mathrm{Dgm}(X_n^{(i)})(K) \to \mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)(K)]$ as $m \to \infty$ by the law of large numbers. However, can we quantify how fast μ_n^m converges?

In Chapter 3, we show how the metrics $\mathbf{d}_p$ can be extended to handle measures with continuous support. It will allow us to address the above questions in Section 7.1.

Learning with persistence diagrams

We end this section by mentioning some interesting questions that arise when one attempts to perform machine learning with persistence diagrams. First, note that as mentioned above, the lack of linear structure in the space of persistence diagrams prevents from using simple statistical descriptors faithfully, such as arithmetic means. *A fortiori*, some sophisticated machine learning pipelines, think for instance of deep neural networks, cannot be used as they generally take elements of a finite-dimensional vector space as input.

To circumvent this issue, various approaches have been proposed to embed the space of persistence diagrams into a linear space, essentially making use of *vectorizations*.

Definition 2.7. A *vectorization* is a map $\Phi : \mathcal{D} \to \mathcal{B}$ for some Banach space $\mathcal{B}$.

A *linear representation* is a vectorization of the form $\Phi : \mu \mapsto \mu(f)$ for some fixed function $f : \Omega \to \mathcal{B}$, where $\mathcal{B}$ is a Banach space.

In practice, essentially two types of vectorizations have been used:

- Through the use of kernel methods [KFH17, LY18, CCO17, RHBK15]. It consists in proposing a map $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ which is symmetric, positive definite. In this case, Aronszajn's theorem states that there exists a map $\phi : \mathcal{D} \to \mathcal{H}$ to some Hilbert space $\mathcal{H}$ (called a *reproducing kernel Hilbert space*) such that $k(\mu, \nu) = \langle \phi(\mu), \phi(\nu) \rangle_{\mathcal{H}}$. The map ϕ is implicit in general but various machine learning algorithms only make use of (expressions based on) the computable quantity $k(\cdot, \cdot)$. This approach will not be covered in detail in this manuscript.

- Through explicit mappings into finite-dimensional spaces [Bub15, AEK+17, BM19]. In practice, it turns out that many of these are linear representations.

The formalism we develop in this manuscript is especially adapted to study linear representations of persistence diagrams, as detailed in Chapter 6. We characterize exhaustively *continuous* linear representations, and detail in Section 6.2 how one can design a neural network layer that will learn a task-specific linear representation based on this formalism.

2.2 Optimal Transport

This section is addressed to the reader who is new to the optimal transport (OT) literature. This brief introduction will essentially focus on the important aspects of OT that will be used through the manuscript. We indicate to the interested reader three books that provide a significantly more exhaustive presentation of the field, ranked from the most theory-heavy to the more computation-oriented: [Vil08], [San15], and [PC17].

2.2.1 Monge and Kantorovich problems

We present here the optimal transport problem in its historical formulation as introduced by Gaspard Monge [Mon84], and its extension due to Kantorovich [Kan42].

Motivation and intuition

The initial motivations behind the development of the optimal transport theory were practical ones. To get the first intuition, one could consider the following problem: consider a pile of sand in some starting configuration, and suppose that you want to move it in order to reach some target configuration (for instance, to fill a hole). Of course, moving the sand is exhausting, and you may want to minimize the amount of effort put to reach your goal.

Without being completely formal yet, we can model this problem in the following way: initially, the sand is distributed in some manner, denoted by μ. The notation $\mathrm{d}\mu(x)$ will denote the quantity (say, mass) of sand located at a given position x. Similarly, the target distribution (i.e. configuration) will be denoted by ν. *Transporting* μ onto ν, in the sense of Monge, means finding a map $T : x \mapsto T(x)$ such that for any set of locations A, one has $\nu(A) = \mu(T^{-1}(A))$, a relation which is denoted by $\nu = T_{\#}\mu$, called a *push-forward*. Note that this implicitly assumes that the distribution μ and ν must have the same total mass: at this stage, we cannot create sand *ex-nihilo*, nor make some sand disappear.

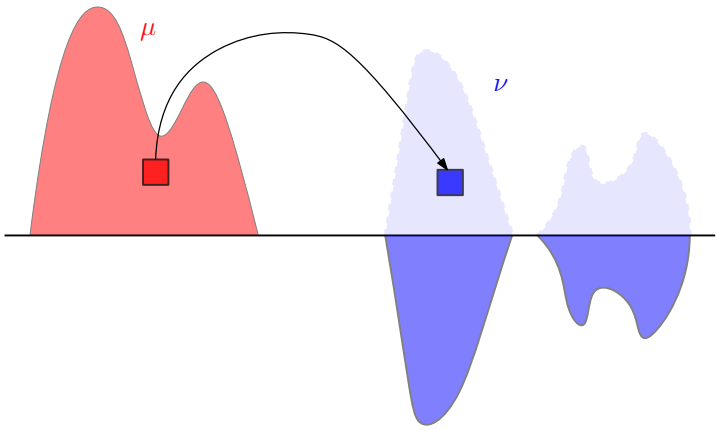


Figure 2.6: Intuitive illustration of the optimal transportation problem in 1D.

For each initial location x, going from some position x to a target position $T(x)$ has some cost, denoted by $c(x, T(x))$. For instance, c could be the distance traveled to reach x from $T(x)$, the time taken to travel, or some price. A natural assumption is that the practical cost for the user is actually proportional to the quantity of mass transported $\mathrm{d}\mu(x)$, that is we actually pay $c(x, T(x))\mathrm{d}\mu(x)$ when transporting a quantity of mass $\mathrm{d}\mu(x)$ located at x onto $T(x)$. As this holds for any location x, the total cost induced by the *transport map* T when transporting μ onto ν is given by the sum of each individual cost $c(x, T(x))\mathrm{d}\mu(x)$, yielding

$$\int_x c(x, T(x))\mathrm{d}\mu(x).$$

Of course, one might have various ways of transporting μ onto ν, that is to pick a map T satisfying $T_{\#}\mu = \nu$, and one might pick an *optimal* one. Namely, we are looking for a map T^* that would realize

$$\inf_{T:T_{\#}\mu=\nu} \int_x c(x, T(x))\mathrm{d}\mu(x).$$

The Monge problem

We are now ready to give the formulation of the Monge problem. First, we consider two Polish spaces $\mathcal{X}$ and $\mathcal{Y}$, and two non-negative finite measures μ and ν supported on $\mathcal{X}$ and $\mathcal{Y}$ respectively. Let $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a lower semi-continuous non-negatively valued cost function.

The *Monge problem* between μ and ν is

$$\underset{T}{\text{minimize}} \left\{ \int_{x\in\mathcal{X}} c(x, T(x))\mathrm{d}\mu(x) : T_{\#}\mu = \nu \right\}. \tag{2.13}$$

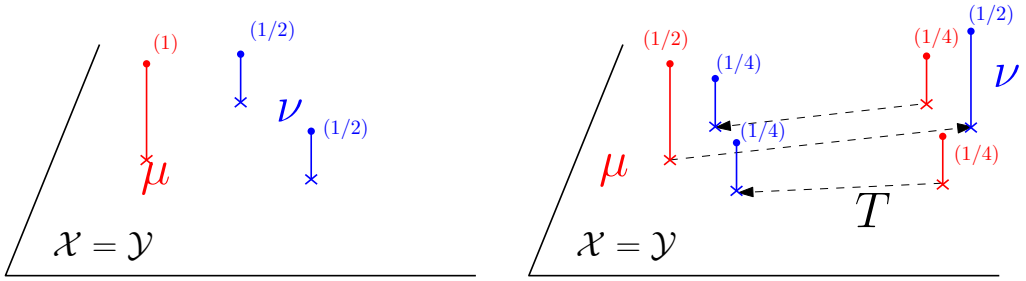


Figure 2.7: *(Left)* A basic example where there exists no transport map between μ and ν. *(Right)* An example where there exists an optimal Monge map (dashed), but performs poorly compared to an optimal transport plan (that would be able to split the (1/2) red mass and transport it onto the (1/4) blue masses).

A map T^* minimizing (2.13) is an *optimal transport map*, and its cost $\int_x c(x, T^*(x))\mathrm{d}\mu(x)$ is the (Monge-)*optimal transport cost* between μ and ν, denoted by $\mathrm{OT}_c(\mu, \nu)$.

First, observe that if μ and ν do not have the same total mass, there is no transport map satisfying $T_{\#}\mu = \nu$. We will thus make this assumption in the remainder of this section. Note also that the problem is linear in the mass, that is for any $\lambda > 0$, we have $\mathrm{OT}_c(\lambda\mu, \lambda\nu) = \lambda\mathrm{OT}_c(\mu, \nu)$. As such, it is common (if not systematic) in the optimal transport literature to consider that μ, ν are *probability measures* on their respective spaces, that is non-negative Radon measures of mass 1.

Unfortunately, in many situations, looking for transport maps from μ to ν is too restrictive. Indeed, this formalism enforces that the *whole* mass located at $x \in \mathcal{X}$ is transported to a *single* target location $T(x) \in \mathcal{Y}$: splitting the mass is not allowed. As such, there are simple situations where there is no transport map. For instance, fix $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ and consider $\mu = \delta_0$ the Dirac mass located at 0, and $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$. In this context, there is no T satisfying $T_{\#}\mu = \nu$; see also Figure 2.7 for an illustration. Actually, even when the set of transport maps is not empty, it might happen that no transport map is optimal, see [Vil08, Example 4.9] for instance.

To overcome this limitation, one resorts to the *Kantorovich* formulation of the transport problem. Essentially, this extension of the Monge problem consists in optimizing on *transport plans* instead of transport maps. A transport plan is a measure π supported on the product space $\mathcal{X} \times \mathcal{Y}$ whose natural interpretation is the following: $\mathrm{d}\pi(x, y)$ represents the fraction of mass initially located at x that is transported to y. In addition, we ask the transport plan to actually transport μ onto ν, which is encoded by the

marginal constraints that read, for all Borel subsets $A \subset \mathcal{X}, B \subset \mathcal{Y}$,

$$\mu(A) = \int_{x\in A}\int_{y\in\mathcal{Y}} \mathrm{d}\pi(x,y), \quad \nu(B) = \int_{y\in B}\int_{x\in\mathcal{X}} \mathrm{d}\pi(x,y). \tag{2.14}$$

The set of transport plans between μ and ν is called the *transportation polytope* between μ and ν and will be denoted by $\Pi(\mu,\nu)$. The Kantorovich formulation of the optimal transport problem finally reads

$$\underset{\pi\in\Pi(\mu,\nu)}{\text{minimize}} \iint_{\mathcal{X}\times\mathcal{Y}} c(x,y)\mathrm{d}\pi(x,y). \tag{2.15}$$

As in the Monge transportation problem, a minimizer of (2.15) is an *optimal transport plan*, and the infimum value is the (Kantorovich-)*optimal transport cost* between μ and ν for the cost c, also denoted by $\mathrm{OT}_c(\mu,\nu)$.

Remark 2.6. Unlike the Monge problem, optimal transport plans for the Kantorovich problem always exist [Vil08, Theorem 4.1]. In general, it might happen that an optimal transport map exists but achieves a much larger cost than an optimal transport plan (see Figure 2.7). However, if μ is atomless, although an optimal transport map might not exist, the infimum value in the Monge problems coincide with the minimum value of the Kantorovich problems [Pra07, Theorem B].

Finally, let us note two important situations where optimal transport maps and optimal transport plans do coincide.

Proposition 2.5 ([Vil08, Theorem 9.4, Theorem 10.41])**.** *If $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ and $c(x,y) = \|x-y\|^2$ (the so-called* quadratic cost*), and if μ admits a density with respect to the Lebesgue measure, then there exists a unique optimal transport plan, which is actually an optimal transport map, which eventually turns out to be the (sub)gradient of a convex function.*

Proposition 2.6 (Birkhoff-von Neumann)**.** *If μ and ν are discrete measures with* uniform *weights, that is of the form $\mu = \sum_{i=1}^n \delta_{x_i}$ and $\nu = \sum_{i=1}^n \delta_{y_j}$ for some integer n, then there exists an optimal transport plan between μ and ν which is a transport map; that is a bijection between the supports of μ and ν, also known as an* optimal matching*.*

2.2.2 Wasserstein distances and Wasserstein spaces

A natural context where one can apply optimal transport is the following: assume that the (Polish) space $\mathcal{X}$ is equipped with a metric d, and consider some $1 \leq p < \infty$. The *p-Wasserstein distance* between μ and ν is defined

as the optimal transport cost (to the power $1/p$) of transporting μ onto ν for the cost $c(x,y) = d(x,y)^p$, that is

$$W_{p,d}(\mu,\nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}, \tag{2.16}$$

When there is no ambiguity on the distance d used, we simply write W_p instead of $W_{p,d}$. In order to have W_p finite, μ and ν are required to have a finite p-th moment, that is there exists $x_0 \in \mathcal{X}$ such that $\int_{\mathcal{X}} d(x,x_0)^p \mathrm{d}\mu(x)$ (resp. $\mathrm{d}\nu$) is finite. A nice observation is that W_p is a metric over the set of such probability measures, referred to as $\mathcal{W}^p(\mathcal{X})$ and called the *p-Wasserstein space* over $(\mathcal{X}, d)$. A first interesting remark that can be done is the simple equality

$$W_{p,d}(\delta_x, \delta_y) = d(x,y),$$

for any $x, y \in \mathcal{X}$. Thus, the Wasserstein space $\mathcal{W}^p(\mathcal{X})$ contains an isometric copy of the ground space $\mathcal{X}$ by simply looking at Dirac masses. From a theoretical viewpoint, the Wasserstein space over $(\mathcal{X}, d)$ is of interest as many geometric properties of $(\mathcal{X}, d)$ are reflected in $\mathcal{W}^p(\mathcal{X})$. For instance, if $(\mathcal{X}, d)$ is a Polish [Vil08, Theorem 6.18], or a compact [Vil08, Remark 6.19], or a geodesic [Vil08, Corollary 7.22] space, so is $\mathcal{W}^p(\mathcal{X})$. Therefore, when dealing with a collection of probability measures $\mu_1 \dots \mu_N$ supported on some "nice" ground space $(\mathcal{X}, d)$, think for instance of $\mathbb{R}^d$ equipped with the Euclidean metric, one can build statistical summaries of the $(\mu_i)_i$s by working in the Wasserstein space. Such statistics are likely to reflect the geometry of the ground space in some sense (see Example 2.2).

Example 2.2. Let $\mu_1 \dots \mu_N$ be a collection of probability measures supported on $\mathbb{R}^d$ with finite second moment. A *Wasserstein barycenter* of $\mu_1 \dots \mu_N$ (also called a *Fréchet mean*) is any minimizer (in $\mathcal{W}^2(\mathbb{R}^d)$) of the *energy functional*

$$\mathcal{E} : \mu \mapsto \sum_{i=1}^{N} \lambda_i W_2^2(\mu, \mu_i), \tag{2.17}$$

where the λ_is are non-negative weights that sum to 1. This problem was studied in the seminal work [AC11] where authors proved the existence of solutions. From a theoretical viewpoint, this result was extended in [LGL16] to the case of averaging infinitely many probability measures on $\mathcal{W}^p(\mathbb{R}^d)$. Namely, given a probability distribution P supported on $\mathcal{W}^p(\mathbb{R}^d)$ (a distribution of probability measures), they proved the existence of solutions to

$$\mathcal{E} : \mu \mapsto \int W_2^2(\mu, \nu) \mathrm{d}P(\nu), \tag{2.18}$$

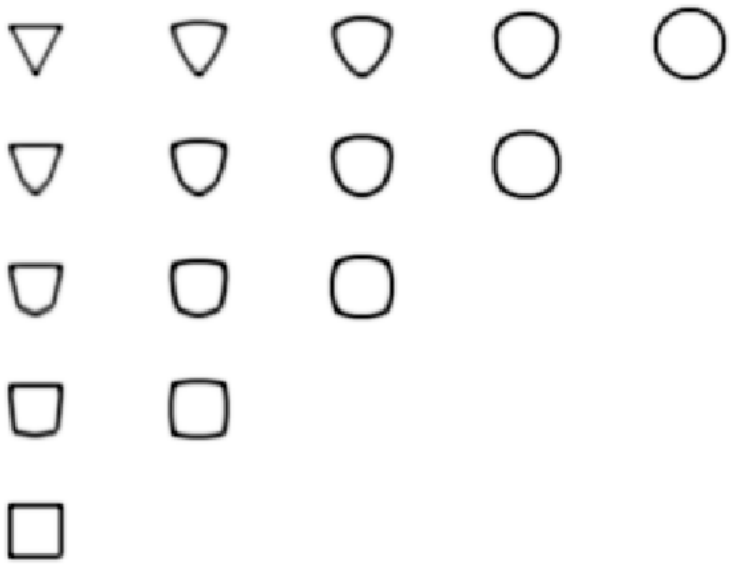


Figure 2.8: Wasserstein barycenters of three measures for different weights.

and they also proved that minimizers of (2.17) must converge to minimizers of (2.18) as $N \to \infty$—under the assumption that the $\mu_i \sim P$, i.i.d.

From a computational perspective, based on entropic regularization (see Section 2.2.4 below) of the Wasserstein metrics, efficient methods to estimate solutions of (2.17) have been proposed in a series of works in the mid 2010's [CD14, SDGP+15, CP16], see Figure 2.8 for an illustration. Wasserstein barycenters are still actively studied; among interesting research branches, we point out the work [GPRS19] where authors study the convergence rates of minimizers of (2.17) to minimizers of (2.18) as $N \to \infty$ (where $\mu_i \sim P$).

2.2.3 Unbalanced optimal transport

Wasserstein distances and $\mathbf{d}_p$ metrics between persistence diagrams defined in Eq. (2.3) share the key idea of defining a distance by minimizing a cost over some matchings. However, the set of transport plans $\Pi(\mu, \nu)$ between two measures is non-empty if and only if the two measures have the same

mass, while persistence diagrams with different masses (number of points) can be compared, making a crucial difference between the W_p and $\mathbf{d}_p$ metrics.

Extending optimal transport to measures of different masses, generally referred to as *unbalanced optimal transport*, has been addressed by different authors [Fig10, CPSV15, KMV+16], and we refer to the PhD thesis manuscript of L.Chizat [Chi17] for a general presentation. We present in the following two approaches to deal with unbalanced measures that will be of interest to handle persistence diagrams.

Regularization of marginal constraints

Consider a Polish space $\mathcal{X}$ and two non-negative finite Radon measures μ, ν supported on $\mathcal{X}$. As mentioned above, the transportation polytope $\Pi(\mu, \nu)$ is empty if (and only if) μ and ν do not have the same masses. Intuitively, it means that in this context, satisfying the marginal constraints (2.14) is too strong of a requirement. A natural workaround is then to relax this requirement: instead of asking the marginals of a measure π on $\mathcal{X} \times \mathcal{X}$ to be *exactly* μ and ν, we penalize the distance between the marginals and the "targets" μ and ν. Formally, this is done using a *Csiszár-divergence* (or *φ-divergence*).

Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be an *entropy function*, that is a convex lower semi-continuous function, satisfying $\varphi(p) = +\infty$ if $p < 0$ and $\varphi(1) = 0$. Define $\varphi'_\infty = \lim_{p\to\infty} \varphi(p)/p$. The divergence associated to φ between two non-negative Radon measures is defined as

$$D_\varphi(\alpha|\beta) := \int_{\mathcal{X}} \varphi\left(\frac{\mathrm{d}\alpha}{\mathrm{d}\beta}\right) \mathrm{d}\beta(x) + \varphi'_\infty \int_{\mathcal{X}} \mathrm{d}\alpha', \tag{2.19}$$

where $\frac{\mathrm{d}\alpha}{\mathrm{d}\beta}$ is the *Radon-Nikodym* decomposition of α with respect to β (see Appendix B), with the standard convention $\infty \times 0 = 0$ if $\varphi'_\infty = \infty$. The unbalanced transport problem between μ and ν is then defined as

$$\underset{\pi\in\mathcal{M}(\mathcal{X}\times\mathcal{X})}{\text{minimize}} \iint c(x,y)\mathrm{d}\pi + \rho D_\varphi(\pi^{(1)}|\mu) + \rho D_\varphi(\pi^{(2)}|\nu), \tag{2.20}$$

where $\pi^{(1)}, \pi^{(2)}$ denote the first and second marginals of π respectively, and $\rho > 0$ is a parameter controlling the strength of the (regularized) marginal constraints.

Example 2.3. Some choices of φ allow recovering well-known notions of divergence between measures. For instance,

- $\varphi : z \mapsto z\log(z) - z + 1$ defines $D_\varphi = \mathrm{KL}$, the *Kullback-Leibler divergence*.
- $\varphi : z \mapsto |1 - z|$ defines the *total variation* between measures.

Note also that taking $\varphi : z \mapsto 0$ if $z = 1$, and $\varphi(z) = +\infty$ otherwise, define the *convex indicator function* of 1, in which case we recover the balanced OT problem (2.15) as a particular case of (2.20).

This formalism can be appealing when attempting to provide an optimal transport formulation of metrics between persistence diagrams. However, it is important to observe that the penalization $D_\varphi(\pi^{(1)}|\mu)$ is *independent of the geometry of the ground space*. When comparing persistence diagrams (2.3), in contrast, we considered that unmatched points were transported to their orthogonal projections onto the diagonal; so the cost of not transporting a point depends on its location.

Optimal transport with boundary

An alternative approach, that is not encompassed by (2.20), was proposed by Figalli and Gigli in [FG10]. In this work, authors propose to extend Wasserstein distances to Radon measures supported on a bounded open proper subset $\mathcal{X}$ of $\mathbb{R}^d$, whose boundary is denoted by $\partial\mathcal{X}$ (and $\overline{\mathcal{X}} := \mathcal{X} \sqcup \partial\mathcal{X}$).

Informally, they propose to use the boundary $\partial\mathcal{X}$ as an infinite reservoir of mass. To do that, they relax the marginal constraints, proposing to match an element of mass $\mathrm{d}\mu(x)$ either to some $\mathrm{d}\nu(y)$, with cost $d(x,y)^p$, or to the boundary of the space $\partial\mathcal{X}$, with cost $d(x,\partial\mathcal{X}) := \min_{x'\in\partial\mathcal{X}} d(x,x')^p$ (and similarly for ν). This is formalized by the following definition.

Definition 2.8. [FG10, Problem 1.1] Let $p \in [1,+\infty)$. Let μ,ν be two Radon measures supported on $\mathcal{X}$ satisfying

$$\int_{\mathcal{X}} d(x,\partial\mathcal{X})^p \mathrm{d}\mu(x) < +\infty, \quad \int_{\mathcal{X}} d(x,\partial\mathcal{X})^p \mathrm{d}\nu(x) < +\infty. \tag{2.21}$$

The set of *admissible transport plans* (or couplings) $\mathrm{Adm}(\mu,\nu)$ is defined as the set of Radon measures π on $\overline{\mathcal{X}}\times\overline{\mathcal{X}}$ satisfying for all Borel sets $A, B \subset \mathcal{X}$,

$$\pi(A \times \overline{\mathcal{X}}) = \mu(A) \quad \text{and} \quad \pi(\overline{\mathcal{X}} \times B) = \nu(B).$$

The cost of $\pi \in \mathrm{Adm}(\mu,\nu)$ is defined as

$$C_p(\pi) := \iint_{\overline{\mathcal{X}}\times\overline{\mathcal{X}}} d(x,y)^p \mathrm{d}\pi(x,y). \tag{2.22}$$

The Optimal Transport (with boundary) distance $\mathrm{OT}_p(\mu, \nu)$ is then defined as

$$\mathrm{OT}_p(\mu, \nu) := \left(\inf_{\pi \in \mathrm{Adm}(\mu,\nu)} C_p(\pi) \right)^{1/p}. \tag{2.23}$$

As in the balanced OT formalism (2.15), plans $\pi \in \mathrm{Adm}(\mu, \nu)$ realizing the infimum in (2.23) are called *optimal*. The set of optimal transport plans between μ and ν for the cost $(x, y) \mapsto d(x, y)^p$ is denoted by $\mathrm{Opt}_p(\mu, \nu)$.

As we will see later in Chapter 3, this formulation is exactly the one we need to bridge the gap between persistence diagram metrics and optimal transport formulations. At this point, we simply put the emphasis on some important facts:

- OT_p is well-defined even between Radon measures of infinite mass, as long as they satisfy (2.21). This property can be seen as an analog to the finite moment condition involved in the definition of the Wasserstein distance in Section 2.2.2.

- In their work, Figalli and Gigli supposed the ground space $\mathcal{X}$ to be bounded. Of course, the upper half-plane Ω on which persistence diagrams are supported is not bounded. This does not prevent us from defining OT_p in this context but will require additional care in the proofs.

We end this section by introducing a useful notion that will be used in Chapter 3, which shows how to build an element of $\mathrm{Adm}(\mu, \nu)$ given a map $f : \overline{\mathcal{X}} \to \overline{\mathcal{X}}$ satisfying some balance condition (see Figure 2.9).

Definition 2.9. Let $\mu, \nu \in \mathcal{M}$. Consider $f : \overline{\mathcal{X}} \to \overline{\mathcal{X}}$ a measurable function satisfying for all Borel set $B \subset \mathcal{X}$

$$\mu(f^{-1}(B) \cap \mathcal{X}) + \nu(B \cap f(\partial\mathcal{X})) = \nu(B). \tag{2.24}$$

Define for all Borel sets $A, B \subset \overline{\mathcal{X}}$,

$$\pi(A \times B) = \mu(f^{-1}(B) \cap \mathcal{X} \cap A) + \nu(\mathcal{X} \cap B \cap f(A \cap \partial\mathcal{X})). \tag{2.25}$$

π is called the transport plan *induced* by the *transport map* f.

One can easily check that we have indeed $\pi(A \times \overline{\mathcal{X}}) = \mu(A)$ and $\pi(\overline{\mathcal{X}} \times B) = \nu(B)$ for any Borel sets $A, B \subset \mathcal{X}$, so that $\pi \in \mathrm{Adm}(\mu, \nu)$ (see Figure 2.9).

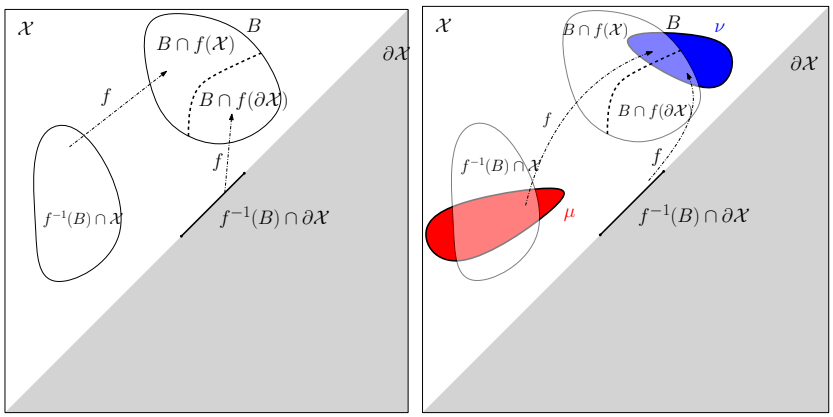


Figure 2.9: A transport map f must satisfy that the mass $\nu(B)$ (light blue) is the sum of the mass $\mu(f^{-1}(B) \cap \mathcal{X})$ given by μ that is transported by f onto B (light red) and the mass $\nu(B \cap f(\partial\mathcal{X}))$ coming from $\partial\mathcal{X}$ and transported by f onto B.

2.2.4 Entropic regularization of optimal transport

The last variation of the optimal transport problem we need to introduce in this background section is the *entropic regularization* of optimal transport, an idea that can be traced back to the Schrödinger problem [Sch31, Léo13]. This approach was highly popularized by Cuturi [Cut13], showing that the (unique!) solution of the regularized problem introduced below can be approximated at a lightspeed rate.

Definition

Consider μ, ν, two probability measures supported on some space $\mathcal{X}$ equipped with some cost function c (typically $c(x, y) = d(x, y)^p$, for $p \geq 1$). The transportation polytope $\Pi(\mu, \nu)$ is not empty: it contains at least the *product measure* $\mu \otimes \nu$, defined as $\mu \otimes \nu(A, B) := \mu(A)\nu(B)$ for all Borel $A, B \subset \mathcal{X}$. Interestingly, it turns out that $\mu \otimes \nu$ is the element of $\Pi(\mu, \nu)$ that has minimal *negentropy* h, defined as

$$h : \pi \mapsto \iint (\log(\pi) - 1)\mathrm{d}\pi,$$

with the standard convention $0 \times \log(0) = 0$. As such, an intuitive understanding one can have about the entropic regularization of optimal transport is the following. On the one hand, there is (at least) one optimal transport plan π^*, but this might be "hard" to find. On the other hand, there is the trivial transport plan $\mu \otimes \nu$, which is straightforward to compute, but which

will be suboptimal in general. Why not interpolate between both? It turns out that $\mu \otimes \nu$ is the element of $\Pi(\mu, \nu)$ that has maximal entropy (minimal negentropy). It invites us to consider, for some regularization parameter $\varepsilon > 0$,

$$W_c^\varepsilon(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \iint c(x, y) \mathrm{d}\pi + \varepsilon h(\pi). \tag{2.26}$$

A slight variation of this definition, used in recent works in optimal transport [RTC17, GPC18, FSV+18, SFV+19], consists in changing the negentropy h for the *Kullback-Leibler divergence* to the product measure $\mu \otimes \nu$, reading

$$\mathrm{OT}_c^\varepsilon(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \iint c(x, y) \mathrm{d}\pi + \varepsilon \mathrm{KL}(\pi | \mu \otimes \nu), \tag{2.27}$$

where we recall that, as μ, ν have mass 1,

$$\mathrm{KL}(\pi | \mu \otimes \nu) = \iint \log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu \otimes \nu}\right) \mathrm{d}\pi.$$

In both cases, the first consequence of introducing the regularization term is that it makes the minimization problem *strictly* convex, ensuring that there is a *unique* solution π^ε. Furthermore, $\pi^\varepsilon \to \pi^*$ as $\varepsilon \to 0$, where π^* is the optimal transport plan between μ and ν (for the non-regularized problem) that has maximal entropy.

Remark 2.7. As explained in [PC17, Remark 4.2], using the Kullback-Leibler divergence instead of the negentropy is nothing but a change of "reference measure". Indeed, the negentropy h is (up to a constant factor) the Kullback-Leibler divergence with respect to the Lebesgue measure. Using $\mu \otimes \nu$ as a reference measure is motivated by the following fact: as the regularization parameter ε goes to ∞, the unique solution π^ε (either of (2.26) or (2.27)) is expected to get closer to $\mu \otimes \nu$ (as it is the measure in $\Pi(\mu, \nu)$ of largest entropy). Using any other reference measure would make the regularized transport cost to diverge as $\varepsilon \to \infty$. In contrast, one has

$$\mathrm{OT}_c^\varepsilon(\mu, \nu) \to \iint c(x, y) \mathrm{d}\mu(x) \mathrm{d}\nu(y)$$

as $\varepsilon \to \infty$. A proof of this fact can be found in the Ph.D. thesis of Aude Genevay (see [Gen19, Theorem 10]).

Eventually, it turns out that the map $(\mu, \nu) \mapsto \pi^\varepsilon$ is smooth. This smoothness is at the core of some variational algorithms to solve optimization problems *in the Wasserstein space*, such as computing Wasserstein barycenters. We will present in Chapter 5 an application of this property in the context of persistence diagrams.

Computational aspect

Nonetheless, what makes the entropic regularization so popular in computational optimal transport nowadays lies in its computational strengths. For the sake of exposition, and as we target numerical applications here, let us expose these strengths in the *finite-dimensional setting*, that is when the measures of interest have finite support. Namely, let $\mu = \sum_{i=1}^{n} a_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{m} b_j \delta_{y_j}$, where the $a = (a_i)_i \in \mathbb{R}_+^n$ and $b = (b_j)_j \in \mathbb{R}_+^m$ are non-negative weights that sum to 1, and $(x_i)_i, (y_j)_j$ are points in $\mathbb{R}^d$. Finally, let $C \in \mathbb{R}^{n \times m}$ denote the matrix $(c(x_i, y_j))_{ij}$ (typically $c(x_i, y_j) = \|x_i - y_j\|_2^2$, but any choice would work in the following).

In this context, (2.26) reads[3]

$$W_c^\varepsilon(\mu, \nu) = \inf_{P \in \mathbb{R}^{n \times m}} \{ \langle P, C \rangle + \varepsilon \sum_{ij} P_{ij}(\log(P_{ij}) - 1), \\ P\mathbf{1}_m = a, P^T \mathbf{1}_n = b \}, \tag{2.28}$$

where $\langle P, C \rangle = \sum_{ij} P_{ij} C_{ij}$ denotes the Frobenius dot product between P and C, and $\mathbf{1}_k$ denotes the vector $(1 \dots 1) \in \mathbb{R}^k$. Using standard duality results, this problem is equivalent to the following unconstrained maximization problem

$$W_c^\varepsilon(\mu, \nu) = \sup_{f \in \mathbb{R}^n, g \in \mathbb{R}^m} \langle a, f \rangle + \langle b, g \rangle - \varepsilon \sum_{ij} \exp\left(\frac{f_i + g_j - C_{ij}}{\varepsilon} \right), \tag{2.29}$$

with the primal-dual relation between a minimizer P^* in (2.28) and maximizers (f^*, g^*) of (2.29):

$$P^* = \exp\left(\frac{f_i + g_j - C_{ij}}{\varepsilon} \right). \tag{2.30}$$

Introducing the variables $u = \exp(f/\varepsilon), v = \exp(g/\varepsilon)$ and $K = \exp(-C/\varepsilon)$ (where exponentiation must be understood term-wise), the first order conditions read (term-wise division)

$$\begin{cases} u = \frac{a}{Kv} \\ v = \frac{b}{K^T u} \end{cases}. \tag{2.31}$$

Said differently, an optimal pair (u, v) is a fixed point of the *Sinkhorn map*

$$S : (u, v) \mapsto \left(\frac{a}{Kv}, \frac{b}{K^T u} \right). \tag{2.32}$$

[3] Similar computational results exist for the variation $\mathrm{OT}_c^\varepsilon$, not detailed here for the sake of concision. See [FSV+18] for details.

Furthermore, the sequence $(u_t, v_t)_t$ defined for arbitrary (u_0, v_0) as

$$\begin{cases} u_{t+1} \leftarrow \frac{a}{Kv_t} \\ v_{t+1} \leftarrow \frac{b}{K^T u_{t+1}} \end{cases} \tag{2.33}$$

converges toward a such a fixed point (u, v). This gives a fairly simple algorithm to approximate a solution of (2.28) (both in terms of minimum and minimizer). The iterations involved in (2.33) only involve basic matrix manipulations and are thus especially suited to be run on modern hardware such as a GPU.

This algorithm can have even more computational benefits when dealing with *histograms* (think of images for instance), that is measures supported on a common finite set of cardinality n. A histogram is thus a vector $\mathbb{R}^n_+$ whose coordinates sum to one. Assume that one wants to solve N instances of (2.28). It involves N pairs of variables $(u^{(1)}, v^{(1)}) \dots (u^{(N)}, v^{(N)})$ but a *single* cost matrix $C \in \mathbb{R}^{n \times n}$. As such, one can run the iterations (2.33) only once, using $u^{(1\dots N)} := (u^{(1)} \dots u^{(N)}) \in \mathbb{R}^{n \times N}$ and $v^{(1\dots N)} = (v^{(1)} \dots v^{(N)})$. This makes the Sinkhorn algorithm especially efficient when one wants to estimate a large number of transport distances between histograms.

Sinkhorn divergences

Although regularized optimal transport was initially popularized for its computational properties, some effort has been recently dedicated to providing a better understanding of its theoretical behavior. In [GPC18], Genevay et al. introduced the notion of *Sinkhorn divergence*, namely they define:

$$S_c^\varepsilon(\mu, \nu) = \mathrm{OT}_c^\varepsilon(\mu, \nu) - \frac{1}{2}\mathrm{OT}_c^\varepsilon(\mu, \mu) - \frac{1}{2}\mathrm{OT}_c^\varepsilon(\nu, \nu). \tag{2.34}$$

The initial idea behind this definition is that $\mathrm{OT}_c^\varepsilon(\mu, \mu)$ does not necessarily equal 0, while $S_c^\varepsilon(\mu, \mu)$ obviously does. Furthermore, while it was well-known that $\mathrm{OT}_c^\varepsilon(\mu, \nu) \to \mathrm{OT}_c(\mu, \nu)$ as $\varepsilon \to 0$ (the same hold with W_c^ε or S_c^ε), authors in [RTC17] and [GPC18] observe that

$$\begin{aligned} \lim_{\varepsilon \to +\infty} S_c^\varepsilon(\mu, \nu) &= -\iint c(x, y)\mathrm{d}(\mu - \nu)(x)\mathrm{d}(\mu - \nu)(y) \\ &=: \mathrm{MMD}(\mu, \nu), \end{aligned}$$

where MMD stands for the *maximum mean discrepancy* [GBR+12], another notion of distance between probability measures, making the Sinkhorn divergence an interpolation between two common ways to compare probability measures. Feydy et al. [FSV+18, Theorem 1] strengthened this fact

by showing, under some assumptions on c and for measures supported on a common compact domain, that although S_c^ε is not a distance (it does not satisfy the triangle inequality), one has for all probability measures $(\mu_n)_n, \mu, \nu$,

$$\begin{aligned} S_c^\varepsilon(\mu,\nu) \geq S_c^\varepsilon(\mu,\mu) \geq 0, \\ S_c^\varepsilon(\mu,\nu) = 0 \Leftrightarrow \mu = \nu, \\ S_c^\varepsilon(\mu_n,\mu) \to 0 \Leftrightarrow \mathrm{OT}_c(\mu_n,\mu) \to 0. \end{aligned}$$

Thus, the Sinkhorn divergence induces a good notion of *divergence* between probability measures. Finally, let us mention that Sinkhorn divergences have been extended to the unbalanced case (Section 2.2.3) in the recent work of Séjourné et al. [SFV+19].

2.3 Notations

We will use the following notations throughout the manuscript.

Let Ω denote the open half-plane

$$\Omega := \{(b, d) \in \mathbb{R}^2,\ b < d\}, \tag{2.35}$$

and let $\partial\Omega$ denote its boundary, namely the diagonal

$$\partial\Omega := \{(t, t),\ t \in \mathbb{R}\}. \tag{2.36}$$

For $x \in \Omega$, $\mathbf{p}_{\partial\Omega}(x)$ denotes its orthogonal projection onto the diagonal $\partial\Omega$. Let $\overline{\Omega}$ denote $\Omega \sqcup \partial\Omega$. The function $d : \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$ denotes the *ground metric* on $\overline{\Omega}$, typically $d(x, y) := \|x - y\|_q$ for some $q \in [1, +\infty]$. We will also use the notation

$$E_\Omega := \overline{\Omega} \times \overline{\Omega} \backslash \partial\Omega \times \partial\Omega.$$

The *space of persistence measures* of parameter $1 \leq p \leq \infty$ is the space of non-negative Radon measures supported on Ω that have finite *total persistence-p*, that is

$$\mathcal{M}^p := \left\{ \mu,\ \mathrm{Pers}_p(\mu) := \int_\Omega d(x, \partial\Omega)^p \mathrm{d}\mu(x) < \infty \right\}, \tag{2.37}$$

where in the case $p = \infty$ the total persistence reads

$$\mathrm{Pers}_\infty(\mu) = \sup_{x \in \mathrm{spt}(\mu)} d(x, \partial\Omega),$$

also known as μ-essential supremum of the function $x \mapsto d(x, \partial\Omega)$. It is equipped with an optimal partial transport metric denoted by OT_p.

The *space of persistence diagrams* of parameter p is the subset $\mathcal{D}^p \subset \mathcal{M}^p$ consisting only of *point measures*, that is measures μ of the form

$$\mu = \sum_{x \in X} n_x \delta_x \tag{2.38}$$

where $X \subset \Omega$ is a locally finite set, and $n_x \in \mathbb{N}$. It is equipped with a partial matching metric $\mathbf{d}_p$, although we show that $\mathbf{d}_p = \mathrm{OT}_p$ over $\mathcal{D}^p$.

Given a Polish metric space $(\mathcal{X}, d)$, and a parameter $1 \leq p \leq \infty$, let $W_{p,d}$ (or W_p if there is no ambiguity on the ground metric d) denote the p-Wasserstein distance between probability measures (of finite p-th moment) supported on $\mathcal{X}$; and $\mathcal{W}^p(\mathcal{X}, d)$ or simply $\mathcal{W}^p(\mathcal{X})$ the space of such probability measures equipped with the metric $W_{p,d}$.

Part I

Theory

Chapter 3

Persistence diagrams and measures, an Optimal Transport Viewpoint

Abstract

This chapter is dedicated to a theoretical study of spaces of persistence diagrams by adopting an optimal transport viewpoint. By establishing a strong connection between diagram metrics and optimal transport ones, we obtain new results on these spaces, in particular, a very useful characterization of convergence of persistence diagrams in terms of measure theory. Furthermore, this formalism allows us to define an extension of the diagram metrics to any measure supported on the open half-plane, not only those with discrete support such as diagrams, offering the possibility to deal with more complicated objects such as random persistence diagrams (see Section 2.1.3). The strength of this formalism will be showcased through various applications in Part II.

3.1 General properties of $\mathcal{M}^p$

Let Ω denote the open half-plane

$$\Omega := \{(b, d) \in \mathbb{R}^2,\ b > d\}, \tag{3.1}$$

and let $\partial\Omega$ denote its boundary, namely the diagonal

$$\partial\Omega := \{(t, t),\ t \in \mathbb{R}\}. \tag{3.2}$$

Eventually, let $\overline{\Omega}$ denote $\Omega \sqcup \partial\Omega$.

For now, we fix a parameter $p \in [1, +\infty)$. The particular case $p = \infty$ is studied in Section 3.3. Recall from Definition 2.4 in Section 2.1 that persistence diagrams are defined as locally finite point measures supported on Ω, that is of the form $\sum_{x \in X} n_x \delta_x$, where $X \subset \Omega$ is locally finite, and $n_x \in \mathbb{N}$ for all $x \in X$. The space of persistence diagrams is equipped with metrics $\mathbf{d}_p$, defining a metric space $(\mathcal{D}^p, \mathbf{d}_p)$. A diagram $\mu \in \mathcal{D}^p$ might have infinitely many points, but we request that $\mathrm{Pers}_p(\mu) := \sum_{x \in X} n_x d(x, \partial\Omega)^p < \infty$, where $d(\cdot, \cdot)$ denotes the ground distance on $\mathbb{R}^2$, typically $\| \cdot - \cdot \|_q$ for $q \in [1, +\infty]$ (the choice of q does not play any important role). This condition is equivalent to saying that μ must be at a finite distance from the empty diagram, i.e. $\mathbf{d}_p(\mu, \mathbf{0}) < \infty$.

Considering diagrams as measures invites us to consider more general measures supported on Ω. Let thus $\mathcal{M}(\Omega)$ denote the set of non-negative Radon measures supported on Ω (see Appendix B for details in measure theory). Without further ado, we introduce the space of *persistence measures*

$$\mathcal{M}^p := \left\{ \mu \in \mathcal{M}(\Omega),\ \mathrm{Pers}_p(\mu) := \int_x d(x, \partial\Omega)^p \mathrm{d}\mu(x) < \infty \right\}. \tag{3.3}$$

We equip $\mathcal{M}^p$ with the metric OT_p (see Section 2.2.3 for a more detailed presentation) defined as

$$\mathrm{OT}_p(\mu, \nu) = \inf_{\pi \in \mathrm{Adm}(\mu,\nu)} \iint_{\overline{\Omega} \times \overline{\Omega}} d(x, y)^p \mathrm{d}\pi(x, y), \tag{3.4}$$

where $\mathrm{Adm}(\mu, \nu)$ is the set of *admissible plans* between μ and ν, that is the set of Radon measures supported on $\overline{\Omega} \times \overline{\Omega}$ satisfying:

$$\text{for all Borel sets } A, B \subset \Omega,\ \pi(A, \overline{\Omega}) = \mu(A),\ \pi(\overline{\Omega}, B) = \nu(B). \tag{3.5}$$

Remark 3.1. These metrics have already been introduced by Figalli and Gigli in [FG10]. In their work, authors however consider the ground space (here, the open half-plane Ω) to be compact (which is obviously not the case here). Some of the results presented below have counterparts in the work of Figalli and Gigli. Sometimes, the corresponding proofs could be immediately adapted to our framework, sometimes it required to develop new techniques (and, of course, some results just do not hold on non-compact domains).

For the ease of readability, we have delayed some proofs to a separate section (§3.5). We explicitly mention for each result if its proof straightforwardly adapts from one of [FG10] or if it can be considered as a new result or a new proof.

Proposition 3.1. *Let $\mu, \nu \in \mathcal{M}$. The set of transport plans* $\mathrm{Adm}(\mu, \nu)$ *is sequentially compact for the vague topology on $E_\Omega := \overline{\Omega} \times \overline{\Omega} \backslash \partial\Omega \times \partial\Omega$. Moreover, if $\mu, \nu \in \mathcal{M}^p$, for this topology,*

- $\pi \in \mathrm{Adm}(\mu, \nu) \mapsto C_p(\pi)$ *is lower semi-continuous.*
- $\mathrm{Opt}_p(\mu, \nu)$ *is a non-empty sequentially compact set.*
- OT_p *is lower semi-continuous, in the sense that for sequences $(\mu_n)_n, (\nu_n)_n$ in $\mathcal{M}^p$ satisfying $\mu_n \xrightarrow{v} \mu$ and $\nu_n \xrightarrow{v} \nu$, we have*

$$\mathrm{OT}_p(\mu, \nu) \leq \liminf_{n \to \infty} \mathrm{OT}_p(\mu_n, \nu_n).$$

Moreover, OT_p *is a distance on $\mathcal{M}^p$.*

These properties are mentioned in [FG10, pages 4-5] in the bounded case, and the corresponding proofs adapt straightforwardly to our framework (see Section 3.5).

Remark 3.2. If a (Borel) measure μ satisfies $\mathrm{Pers}_p(\mu) < \infty$, then for any Borel set $A \subset \Omega$ satisfying $d(A, \partial\Omega) := \inf_{x \in A} d(x, \partial\Omega) > 0$, we have:

$$\begin{aligned} \mu(A) d(A, \partial\Omega)^p &\leq \int_A d(x, \partial\Omega)^p \mathrm{d}\mu(x) \\ &\leq \int_\Omega d(x, \partial\Omega)^p \mathrm{d}\mu(x) = \mathrm{Pers}_p(\mu) < \infty, \end{aligned} \tag{3.6}$$

so that $\mu(A) < \infty$. In particular, μ is automatically a Radon measure.

The following lemma gives a simple way to approximate a persistence measure (resp. diagram) with ones of finite masses.

Lemma 3.2. *Let $\mu \in \mathcal{M}^p$. Fix $r > 0$, and let $A_r := \{x \in \Omega, d(x, \partial\Omega) \leq r\}$. Let $\mu^{(r)}$ be the restriction of μ to $\Omega \backslash A_r$. Then $\mathrm{OT}_p(\mu^{(r)}, \mu) \to 0$ when $r \to 0$. Similarly, if $a \in \mathcal{D}^p$, we have $\mathbf{d}_p(a^{(r)}, a) \to 0$.*

Proof. Let $\pi \in \mathrm{Adm}(\mu, \mu^{(r)})$ be the transport plan induced by the identity map on $\overline{\Omega} \backslash A_r$, and the projection onto $\partial\Omega$ on A_r. As π is sub-optimal, one has:

$$\mathrm{OT}_p^p(\mu, \mu^{(r)}) \leq C_p(\pi) = \int_{A_r} d(x, \partial\Omega)^p \mathrm{d}\mu(x) = \mathrm{Pers}_p(\mu) - \mathrm{Pers}_p(\mu^{(r)}).$$

Thus, by the monotone convergence theorem applied to μ with the functions $f_r : x \mapsto d(x, \partial\Omega)^p \cdot \mathbf{1}_{\Omega \backslash A_r}(x)$, we obtain that $\mathrm{OT}_p(\mu, \mu^{(r)}) \to 0$ as $r \to 0$. Similar arguments show that $\mathbf{d}_p(a^{(r)}, a) \to 0$ as $r \to 0$. □

The following result is central in this work: it shows that the metrics OT_p are extensions of the metrics $\mathbf{d}_p$.

Proposition 3.3. *For $a, b \in \mathcal{D}^p$, $\mathrm{OT}_p(a,b) = \mathbf{d}_p(a,b)$.*

Proof. Let $a, b \in \mathcal{D}^p$ be two persistence diagrams. Assume first that a, b have finitely many points $\{x_1 \dots x_n\}$ and $\{y_1 \dots y_m\}$ respectively. For the sake of simplicity, let us assume that each x_i and y_j has multiplicity one. Let γ be an optimal partial matching that achieves $\mathbf{d}_p(a,b)$ (see Section 2.1). It induces an admissible transport plan π by defining $\pi(x,y) = 1$ if $(x,y) \in \gamma$, and 0 otherwise. Furthermore,

$$\sum_{(x,y)\in\gamma} d(x,y)^p = \iint d(x,y)^p \mathrm{d}\pi(x,y).$$

It yields $\mathbf{d}_p(a,b) \geq \mathrm{OT}_p(a,b)$.

Conversely, an element $\pi \in \mathrm{Adm}(a,b)$ that would achieve $\mathrm{OT}_p(a,b)$ must be supported on $\{x_1 \dots x_n\} \cup \{\mathbf{p}_{\partial\Omega}(y_1) \dots \mathbf{p}_{\partial\Omega}(y_m)\} \times \{y_1 \dots y_m\} \cup \{\mathbf{p}_{\partial\Omega}(x_1) \dots \mathbf{p}_{\partial\Omega}(x_n)\}$, where $\mathbf{p}_{\partial\Omega}(x)$ denotes the orthogonal projection of $x \in \Omega$ onto the diagonal $\partial\Omega$. It can thus be described as a bi-stochastic matrix $P \in \mathbb{R}_+^{(n+m)\times(m+n)}$. If we denote by $\mathcal{S}_{n+m}$ this set, it yields

$$\mathrm{OT}_p(a,b) = \inf_{P\in\mathcal{S}_{n+m}} \langle C, P\rangle,$$

where $\langle C, P\rangle = \sum_{ij} C_{ij}P_{ij}$ (Frobenius dot product) and $C \in \mathbb{R}^{(n+m)\times(m+n)}$ is defined by

$$\begin{aligned}
&C_{ij} = d(x_i, y_j) \text{ if } 1 \leq i \leq n, 1 \leq j \leq m,\\
&C_{ij} = d(x_i, \mathbf{p}_{\partial\Omega}(x_i)) \text{ if } 1 \leq i \leq n, j > m,\\
&C_{ij} = d(\mathbf{p}_{\partial\Omega}(y_j), y_j) \text{ if } i > n, 1 \leq j \leq m,\\
&C_{ij} = 0 \text{ if } i > n, j > m.
\end{aligned}$$

By Birkhoff-Von Neumann theorem (Proposition 2.6), it follows that among the minimizers of $\langle C, P\rangle$ in $\mathcal{S}_{n+m}$, there must be a permutation matrix P^*. Such a permutation matrix induces a partial matching γ defined by $(x_i, y_j) \in \gamma \Leftrightarrow P_{ij} = 1$, while we have

$$\mathrm{OT}_p(a,b) = \langle C, P^*\rangle = \sum_{(x,y)\in\gamma} d(x,y)^p \geq \mathbf{d}_p(a,b),$$

so that eventually $\mathrm{OT}_p(a,b) = \mathbf{d}_p(a,b)$.

In the general case, let $r > 0$. Due to (3.6), the diagrams $a^{(r)}$ and $b^{(r)}$ defined in Lemma 3.2 have a finite mass (thus finite number of points). Therefore, $\mathbf{d}_p(a^{(r)}, b^{(r)}) = \mathrm{OT}_p(a^{(r)}, b^{(r)})$. By Lemma 3.2, the former converges to $\mathbf{d}_p(a, b)$ while the latter converges to $\mathrm{OT}_p(a, b)$, giving the conclusion. □

As a consequence of this result, we will use OT_p to denote the distance between two elements of $\mathcal{D}^p$ from now on.

Proposition 3.4. *The space $(\mathcal{M}^p, \mathrm{OT}_p)$ is a Polish space.*

As for Proposition 3.1, this proposition appears in [FG10, Proposition 2.7] in the bounded case, and its proof is straightforwardly adapted to our framework (see Section 3.5).

The following is one of the most important results of this chapter: a characterization of convergence in $(\mathcal{M}^p, \mathrm{OT}_p)$.

Theorem 3.5. *Let $\mu, \mu_1, \mu_2, \ldots$ be measures in $\mathcal{M}^p$. Then,*

$$\mathrm{OT}_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \xrightarrow{v} \mu, \\ \mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu). \end{cases} \tag{3.7}$$

This result is analog to the characterization of convergence of probability measures in the Wasserstein space (see [Vil08, Theorem 6.9]) and can be found in [FG10, Proposition 2.7] in the case where the ground space is bounded. While the proof of the direct implication can be easily adapted from [FG10], a new proof is needed for the converse implication. Both can be found in Section 3.5.

Remark 3.3. The assumption $\mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu)$ is crucial to obtain OT_p-convergence assuming vague convergence. For instance, the sequence defined by $\mu_n := \delta_{(n,n+1)}$ converges vaguely to $\mu = 0$ and $(\mathrm{Pers}_p(\mu_n))_n$ does converge (it is constant), while $\mathrm{OT}_p(\mu_n, 0) \nrightarrow 0$. This does not contradict Theorem 3.5 since $\mathrm{Pers}_p(\mu) = 0 \neq \lim_n \mathrm{Pers}_p(\mu_n)$.

Theorem 3.5 implies some very useful results. First, it entails that the topology of the metric OT_p is stronger than the vague topology. As a consequence, the following corollary holds, using Proposition B.5 ($\mathcal{D}^p$ is closed in $\mathcal{M}^p$ for the vague convergence).

Corollary 3.6. *$\mathcal{D}^p$ is closed in $\mathcal{M}^p$ for the metric OT_p.*

We recover in particular that the space of persistence diagrams $(\mathcal{D}^p, \mathrm{OT}_p)$ is a Polish space (due to Proposition 3.4), a result already proved in [MMH11, Theorems 7 and 12] with a different approach.

Secondly, we show that the vague convergence of μ_n to μ along with the convergence of $\mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu)$ is equivalent to the weak convergence of a weighted measure (see Appendix B for a definition of weak convergence, denoted by $\xrightarrow{w}$ in the following). For $\mu \in \mathcal{M}^p$, let us introduce the measure with finite mass μ^p defined, for a Borel subset $A \subset \Omega$, as:

$$\mu^p(A) = \int_A d(x, \partial\Omega)^p \mathrm{d}\mu(x). \tag{3.8}$$

Corollary 3.7. *For a sequence $(\mu_n)_n$ and a persistence measure $\mu \in \mathcal{M}^p$, we have*

$$\mathrm{OT}_p(\mu_n, \mu) \to 0 \text{ if and only if } \mu_n^p \xrightarrow{w} \mu^p.$$

Proof. Consider $\mu, \mu_1, \mu_2, \cdots \in \mathcal{M}^p$ and assume that $\mathrm{OT}_p(\mu_n, \mu) \to 0$. By Theorem 3.5, this is equivalent to $\mu_n \xrightarrow{v} \mu$ and $\mu_n^p(\Omega) = \mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu) = \mu^p(\Omega)$. Since for any continuous function f compactly supported, the map $x \mapsto d(x, \partial\Omega)^p f(x)$ is also continuous and compactly supported, $\mu_n \xrightarrow{v} \mu$ is equivalent to $\mu_n^p \xrightarrow{v} \mu^p$. By Proposition B.3, the vague convergence along with the convergence of the masses is also equivalent to the weak convergence $\mu_n^p \xrightarrow{w} \mu^p$. □

We end this section with two results: a characterization of relatively compact sets in $(\mathcal{M}^p, \mathrm{OT}_p)$, and a convexity property of OT_p^p that will be useful later on.

Proposition 3.8. *A set F is relatively compact in $(\mathcal{M}^p, \mathrm{OT}_p)$ if and only if the set $\{\mu^p, \mu \in F\}$ is tight and $\sup_{\mu \in F} \mathrm{Pers}_p(\mu) < \infty$.*

Proof. From Corollary 3.7, the relative compactness of a set $F \subset \mathcal{M}^p$ for the metric OT_p is equivalent to the relative compactness of the set $\{\mu^p,\ \mu \in F\}$ for the weak convergence. Recall that all μ^p have a finite mass, as $\mu^p(\Omega) = \mathrm{Pers}_p(\mu) < \infty$. Therefore, one can use Prokhorov's theorem (Proposition B.2) to conclude. □

Remark 3.4. This characterization is equivalent to the one described in [MMH11, Theorem 21] for persistence diagrams. The notions introduced therein by the authors of off-diagonally birth-death boundedness, and uniformness are rephrased using the notion of tightness, standard in measure theory.

Lemma 3.9. *For $1 \leq p < \infty$, the function $\mathrm{OT}_p^p : \mathcal{M}^p \times \mathcal{M}^p \to \mathbb{R}$ is convex, that is for all $t \in [0,1]$ and all $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{M}^p$, we have*

$$\mathrm{OT}_p^p(t\mu_1 + (1-t)\mu_2, t\nu_1 + (1-t)\nu_2) \leq t\mathrm{OT}_p^p(\mu_1, \nu_1) + (1-t)\mathrm{OT}_p^p(\mu_2, \nu_2).$$

Proof. Fix $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{M}^p$ and $t \in [0,1]$. Let $\pi_{11} \in \mathrm{Opt}_p(\mu_1, \nu_1)$ and $\pi_{22} \in \mathrm{Opt}_p(\mu_2, \nu_2)$. It is straightforward to check that $\pi := t\pi_{11} + (1-t)\pi_{22}$ is an admissible plan between $t\mu_1 + (1-t)\mu_2$ and $t\nu_1 + (1-t)\nu_2$. The cost of this admissible plan is $t\mathrm{OT}_p^p(\mu_1, \nu_1) + (1-t)\mathrm{OT}_p^p(\mu_2, \nu_2)$, which is therefore larger than $\mathrm{OT}_p^p(t\mu_1 + (1-t)\mu_2, t\nu_1 + (1-t)\nu_2)$. □

3.2 Persistence measures in the finite setting

In practice, many statistical results regarding persistence diagrams are stated for sets of diagrams with uniformly bounded number of points [KHN+15, CCO17], and the specific properties of OT_p in this setting are therefore of interest. Introduce for $m \geq 0$ the subset $\mathcal{M}_{\leq m}^p$ of $\mathcal{M}^p$ defined as $\mathcal{M}_{\leq m}^p := \{\mu \in \mathcal{M}^p,\ \mu(\Omega) \leq m\}$, and the set $\mathcal{M}_f^p$ of finite persistence measures, $\mathcal{M}_f^p := \bigcup_{m \geq 0} \mathcal{M}_{\leq m}^p$. Define similarly the set $\mathcal{D}_{\leq m}$ (resp. $\mathcal{D}_f$). Note that the assumption $\mathrm{Pers}_p(a) < \infty$ is always satisfied for a finite diagram a (which is not true for general Radon measures), so that the exponent p is not needed when defining $\mathcal{D}_{\leq m}$ and $\mathcal{D}_f$.

Proposition 3.10. *$\mathcal{M}_f^p$ (resp. $\mathcal{D}_f$) is dense in $\mathcal{M}^p$ (resp. $\mathcal{D}^p$) for the metric OT_p.*

Proof. This is a straightforward consequence of Lemma 3.2. Indeed, if $\mu \in \mathcal{M}^p$ and $r > 0$, then (3.6) implies that $\mu^{(r)}$ is of finite mass while $\mu^{(r)} \to \mu$ when $r \to 0$. □

Let $\tilde{\Omega} = \Omega \sqcup \{\partial\Omega\}$ be the quotient of $\overline{\Omega}$ by the closed subset $\partial\Omega$—i.e. we encode the diagonal by just one point (still denoted by $\partial\Omega$). The distance d on $\overline{\Omega}^2$ induces naturally a function $\tilde{d}$ on $\tilde{\Omega}^2$, defined for $x, y \in \Omega$ by $\tilde{d}(x,y) = d(x,y)$, $\tilde{d}(x, \partial\Omega) = \tilde{d}(\partial\Omega, x) = d(x, \mathbf{p}_{\partial\Omega}(x))$ and $\tilde{d}(\partial\Omega, \partial\Omega) = 0$. However, $\tilde{d}$ is not a distance since one can have $\tilde{d}(x,y) > \tilde{d}(x, \partial\Omega) + \tilde{d}(y, \partial\Omega)$. Define

$$\rho(x,y) := \min\{\tilde{d}(x,y), \tilde{d}(x, \partial\Omega) + \tilde{d}(y, \partial\Omega)\}. \tag{3.9}$$

It is straightforward to check that ρ is a distance on $\tilde{\Omega}$ and that $(\tilde{\Omega}, \rho)$ is a Polish space. One can then define the Wasserstein distance $W_{p,\rho}$ with respect to ρ for finite measures on $\tilde{\Omega}$ which have the same masses, that is the infimum of $\tilde{C}_p(\tilde{\pi}) := \iint_{\tilde{\Omega}^2} \rho(x,y)^p \mathrm{d}\tilde{\pi}(x,y)$, for $\tilde{\pi}$ a transport plan with corresponding marginals (see Section 2.2.2). The following result states that the problem of computing the OT_p metric between two persistence measures with finite masses can be turn into the one of computing the Wasserstein

distances between two measures supported on $\tilde{\Omega}$ with the same mass. For the sake of simplicity, we assume here that $d(x,y) = \|x-y\|_q$ with $q > 1$ so that for $x \in \Omega$, the set $\arg\min\{y \in \partial\Omega,\ d(x,y)\}$ is given by the orthogonal projection $\mathbf{p}_{\partial\Omega}(x)$ of x onto $\partial\Omega$. The following result could be seamlessly adapted to the case $q = 1$.

Proposition 3.11. *Let $\mu, \nu \in \mathcal{M}_f^p$ and $r \geq \mu(\Omega) + \nu(\Omega)$. Define $\tilde{\mu} = \mu + (r - \mu(\Omega))\delta_{\partial\Omega}$ and $\tilde{\nu} = \nu + (r - \nu(\Omega))\delta_{\partial\Omega}$. Then $\mathrm{OT}_p(\mu,\nu) = W_{p,\rho}(\tilde{\mu},\tilde{\nu})$.*

Before proving Proposition 3.11, we need to introduce two lemmas. The first one essentially states that an element $\pi \in \mathrm{Adm}(\mu,\nu)$ can be turned into an element $\tilde{\pi} \in \Pi(\tilde{\mu},\tilde{\nu})$ such that π and $\tilde{\pi}$ have the same transport cost (for d and $\tilde{d}$ respectively), and vice-versa. The second lemma gives a first inequality to prove Proposition 3.11. The proofs of these lemmas have been delayed to Section 3.5.

Lemma 3.12. *Let $\mu, \nu \in \mathcal{M}_f^p$ and $r \geq \max(\mu(\Omega), \nu(\Omega))$. Let $\tilde{\mu} := \mu + (r - \mu(\Omega))\delta_{\partial\Omega}$, $\tilde{\nu} := \nu + (r - \nu(\Omega))\delta_{\partial\Omega}$ and $\mathbf{p}_{\partial\Omega} : \Omega \to \partial\Omega$ be the orthogonal projection on the diagonal.*

1. *Define $T(\mu,\nu)$ the set of plans $\pi \in \mathrm{Adm}(\mu,\nu)$ satisfying $\pi(\{(x,y) \in \Omega \times \partial\Omega,\ y \neq \mathbf{p}_{\partial\Omega}(x)\}) = \pi(\{(x,y) \in \partial\Omega \times \Omega,\ x \neq \mathbf{p}_{\partial\Omega}(y)\}) = 0$ along with $\pi(\partial\Omega \times \partial\Omega) = 0$. Then, $\mathrm{Opt}_p(\mu,\nu) \subset T(\mu,\nu)$.*

2. *Let $\pi \in T(\mu,\nu)$ be such that $\mu(\Omega) + \pi(\partial\Omega \times \Omega) \leq r$. Define $\iota(\pi) \in \Pi(\tilde{\mu},\tilde{\nu})$ by, for Borel sets $A, B \subset \Omega$,*

$$\begin{cases} \iota(\pi)(A \times B) = \pi(A \times B), \\ \iota(\pi)(A \times \{\partial\Omega\}) = \pi(A \times \partial\Omega), \\ \iota(\pi)(\{\partial\Omega\} \times B) = \pi(\partial\Omega \times B), \\ \iota(\pi)(\{\partial\Omega\} \times \{\partial\Omega\}) = r - \mu(\Omega) - \pi(\partial\Omega \times \Omega) \geq 0. \end{cases} \tag{3.10}$$

   Then, $C_p(\pi) = \iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\iota(\pi)(x,y)$.

3. *Let $\tilde{\pi} \in \Pi(\tilde{\mu},\tilde{\nu})$. Define $\kappa(\tilde{\pi}) \in T(\mu,\nu)$ by (where A, B are Borel sets),*

$$\begin{cases} \kappa(\tilde{\pi})(A \times B) = \tilde{\pi}(A \times B), & A, B \subset \Omega, \\ \kappa(\tilde{\pi})(A \times B) = \tilde{\pi}((A \cap s^{-1}(B)) \times \{\partial\Omega\}), & A \subset \Omega, B \subset \partial\Omega, \\ \kappa(\tilde{\pi})(A \times B) = \tilde{\pi}(\{\partial\Omega\} \times (B \cap s^{-1}(A))), & A \subset \partial\Omega, B \subset \Omega, \\ \kappa(\tilde{\pi})(\partial\Omega, \partial\Omega) = 0. \end{cases}$$

   Then, $\iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}(x,y) = C_p(\kappa(\tilde{\pi}))$.

Lemma 3.13. *Let $\mu, \nu \in \mathcal{M}_f^p$ and $r \geq \max(\mu(\Omega), \nu(\Omega))$. Let $\tilde{\mu} := \mu + (r - \mu(\Omega))\delta_{\partial\Omega}$, $\tilde{\nu} := \nu + (r - \nu(\Omega))\delta_{\partial\Omega}$. Then, $\mathrm{OT}_p(\mu, \nu) \leq W_{p,\rho}(\tilde{\mu}, \tilde{\nu})$.*

Proof of Proposition 3.11. Let $\pi \in T(\mu, \nu)$. As $\mu(\Omega) + \pi(\partial\Omega \times \Omega) \leq \mu(\Omega) + \nu(\Omega) \leq r$, one can define $\tilde{\pi} = \iota(\pi)$. Since $\rho(x, y) \leq d(x, y)$, we have $\tilde{C}_p(\tilde{\pi}) \leq \iint d(x, y)^p \mathrm{d}\tilde{\pi}(x, y) = C_p(\pi)$ (Lemma 3.12). Taking the infimum gives $W_{p,\rho}(\tilde{\mu}, \tilde{\nu}) \leq \mathrm{OT}_p(\mu, \nu)$. The other inequality holds according to Lemma 3.13. □

Remark 3.5. A careful read of the proof of Proposition 3.11 and relative lemmas shows that one could replace the metric ρ by

$$\rho_p : (x, y) \mapsto \min\{d(x, y), (d(x, \partial\Omega)^p + d(\partial\Omega, y)^p)^{\frac{1}{p}}\}.$$

This fact will be used in Section 3.4.

Remark 3.6. The starting idea of this theorem—informally,"adding the mass of one diagram to the other and vice-versa"—is known in TDA as a *bipartite graph matching* [EH10, Ch. VIII.4] and used in practical computations [KMN17]. Here, Proposition 3.11 states that solving this bipartite graph matching problem can be formalized as computing a Wasserstein distance on the metric space $(\tilde{\Omega}, \rho)$ and, as such, makes sense (and remains true) for more general measures.

Remark 3.7. Proposition 3.11 is useful for numerical purpose since it allows in applications, when dealing with a finite set of finite measures (in particular diagrams), to directly use the various tools developed in computational optimal transport [PC17] to compute Wasserstein distances. This alternative to the combinatorial algorithms considered in the literature [KMN17, TMMH14] will be presented in Chapter 5, based on [LCO18]. This result is also helpful to prove the existence of Fréchet means of sets of persistence measures (see Chapter 4).

3.3 The OT_∞ distance

In classical optimal transport, the ∞-Wasserstein distance is known to have a much more erratic behavior than its $p < \infty$ counterparts [San15, Section 5.5.1]. However, in the context of persistence diagrams, the $\mathbf{d}_\infty$ distance defined in Eq. (2.1) appears naturally as an interleaving distance between persistence modules (see Appendix A.2) and satisfies strong stability results: it is thus worthy of interest. It also happens that, when restricted to

diagrams having some specific finiteness properties, most irregular behaviors are suppressed and a convenient characterization of convergence exists.

Definition 3.1. Let $\mathrm{spt}(\mu)$ denote the support of a measure μ and

$$\mathrm{Pers}_\infty(\mu) := \sup\{d(x,\partial\Omega),\ x \in \mathrm{spt}(\mu)\}.$$

Define

$$\mathcal{M}^\infty := \{\mu \in \mathcal{M},\ \mathrm{Pers}_\infty(\mu) < \infty\} \quad \text{and} \quad \mathcal{D}^\infty := \mathcal{D} \cap \mathcal{M}^\infty. \tag{3.11}$$

For $\mu, \nu \in \mathcal{M}^\infty$ and $\pi \in \mathrm{Adm}(\mu,\nu)$, let

$$C_\infty(\pi) := \sup\{d(x,y),\ (x,y) \in \mathrm{spt}(\pi)\}$$

and let

$$\mathrm{OT}_\infty(\mu,\nu) := \inf_{\pi \in \mathrm{Adm}(\mu,\nu)} C_\infty(\pi). \tag{3.12}$$

The set of transport plans minimizing (3.12) is denoted by $\mathrm{Opt}_\infty(\mu,\nu)$.

Proposition 3.14. *Let $\mu, \nu \in \mathcal{M}^\infty$. For the vague topology on E_Ω,*

- *the map $\pi \in \mathrm{Adm}(\mu,\nu) \mapsto C_\infty(\pi)$ is lower semi-continuous.*
- *The set $\mathrm{Opt}_\infty(\mu,\nu)$ is a non-empty sequentially compact set.*
- *OT_∞ is lower semi-continuous.*

Moreover, OT_∞ is a distance on $\mathcal{M}^\infty$.

The proofs of these results are found in Section 3.5.
As in the case $p < \infty$, OT_∞ and $\mathbf{d}_\infty$ coincide on $\mathcal{D}^\infty$.

Proposition 3.15. *For $a, b \in \mathcal{D}^\infty$, $\mathrm{OT}_\infty(a,b) = d_\infty(a,b)$.*

Proof. Consider two diagrams $a, b \in \mathcal{D}^\infty$, written as $a = \sum_{i\in I} \delta_{x_i}$ and $b = \sum_{j\in J} \delta_{y_j}$, where $I, J \subset \mathbb{N}^*$ are (potentially infinite) sets of indices. The marginals constraints imply that a plan $\pi \in \mathrm{Adm}(\mu,\nu)$ is supported on $(\{x_i\}_i \cup \partial\Omega) \times (\{y_j\}_j \cup \partial\Omega)$, and the cost of such a plan can always be reduced if some of the mass $\pi(\{x_i\}, \partial\Omega)$ (resp. $\pi(\partial\Omega, \{y_j\})$) is sent on another point

than the projection of x_i (resp. y_j) on the diagonal $\partial\Omega$. Introduce the matrix C indexed on $(-J \cup I) \times (-I \cup J)$ defined by

$$\begin{cases} C_{i,j} = d(x_i, y_j) & \text{for } i,j > 0, \\ C_{i,j} = d(\partial\Omega, y_j) & \text{for } i < 0, j > 0, \\ C_{i,j} = d(x_i, \partial\Omega) & \text{for } i > 0, j < 0, \\ C_{i,j} = 0 & \text{for } i,j < 0. \end{cases} \tag{3.13}$$

In this context, an element of $\mathrm{Opt}(a,b)$ can be written a matrix P indexed on $(-J \cup I) \times (-I \cup J)$, and marginal constraints state that P must belong to the set of doubly stochastic matrices $\mathcal{S}$. Therefore, $\mathrm{OT}_\infty(a,b) = \inf_{P\in\mathcal{S}} \sup\{C_{i,j},\ (i,j) \in \mathrm{spt}(P)\}$, where $\mathcal{S}$ is the set of doubly stochastic matrices indexed on $(-J \cup I) \times (-I \cup J)$, and $\mathrm{spt}(P)$ denotes the support of P, that is the set $\{(i,j),\ P_{i,j} > 0\}$.

Let $P \in \mathcal{S}$. For any $k \in \mathbb{N}$, and any set of distinct indices $\{i_1, \dots, i_k\} \subset -J \cup I$, we have

$$k = \sum_{k'=1}^{k} \underbrace{\sum_{j\in -I\cup J} P_{i_{k'},j}}_{=1} = \sum_{j\in -I\cup J} \underbrace{\sum_{k'=1}^{k} P_{i_{k'},j}}_{\leq 1}.$$

Thus, the cardinality of $\{j,\ \exists k' \text{ such that } (i_{k'}, j) \in \mathrm{spt}(P)\}$ must be larger than k. Said differently, the marginals constraints impose that any set of k points in a must be matched to *at least* k points in b (points are counted with eventual repetitions here). Under such conditions, the Hall's marriage theorem (see [Hal86, p. 51]) guarantees the existence of a permutation matrix P' with $\mathrm{spt}(P') \subset \mathrm{spt}(P)$. As a consequence,

$$\begin{aligned} \sup\{C_{i,j},\ (i,j) \in \mathrm{spt}(P)\} &\geq \sup\{C_{i,j},\ (i,j) \in \mathrm{spt}(P')\} \\ &\geq \inf_{P'\in\mathcal{S}'} \sup\{C_{i,j},\ (i,j) \in \mathrm{spt}(P')\} = \mathbf{d}_\infty(a,b), \end{aligned}$$

where $\mathcal{S}'$ denotes the set of permutations matrix indexed on $(-J \cup I) \times (-I \cup J)$. Taking the infimum on $P \in \mathcal{S}$ on the left-hand side and using that $\mathcal{S}' \subset \mathcal{S}$ finally gives that $\mathrm{OT}_\infty(a,b) = \mathbf{d}_\infty(a,b)$. □

Proposition 3.16. *The space $(\mathcal{M}^\infty, \mathrm{OT}_\infty)$ is complete.*

Proof. Let $(\mu_n)_n$ be a Cauchy sequence for OT_∞. Fix a compact $K \subset \Omega$, and pick $\varepsilon = d(K, \partial\Omega)/2$. There exists n_0 such that for $n > n_0$, $\mathrm{OT}_\infty(\mu_n, \mu_{n_0}) < \varepsilon$. Let $K_\varepsilon := \{x \in \Omega,\ d(x,K) \leq \varepsilon\}$. By considering $\pi_n \in \mathrm{Opt}_\infty(\mu_n, \mu_{n_0})$, and since $\mathrm{OT}_\infty(\mu_n, \mu_{n_0}) < \varepsilon$, we have that

$$\mu_n(K) = \pi_n(K \times \overline{\Omega}) = \pi_n(K \times K_\varepsilon) \leq \mu_{n_0}(K_\varepsilon). \tag{3.14}$$

Therefore, $(\mu_n(K))_n$ is uniformly bounded, and Proposition B.1 implies that $(\mu_n)_n$ is relatively compact. Finally, the exact same computations as in the proof of the completeness for $p < \infty$ (see Section 3.5) show that $(\mu_n)_n$ converges for the OT_∞ metric. □

Remark 3.8. Unlike the case $p < \infty$, the space $\mathcal{D}^\infty$ (and therefore $\mathcal{M}^\infty$) is not separable. Indeed, for $I \subset \mathbb{N}$, define the diagram $a_I := \sum_{i \in I} \delta_{(i,i+1)} \in \mathcal{D}^\infty$. The family $\{a_I,\ I \subset \mathbb{N}\}$ is uncountable, and for two distinct I, I', $\mathrm{OT}_\infty(a_I, a_{I'}) = \frac{\sqrt{2}}{2}$. This result is similar to [BV18, Theorem 4.20].

We now show that the direct implication in Theorem 3.5 still holds in the case $p = \infty$.

Proposition 3.17. *Let $\mu, \mu_1, \mu_2, \ldots$ be measures in $\mathcal{M}^\infty$. If $\mathrm{OT}_\infty(\mu_n, \mu) \to 0$, then $(\mu_n)_n$ converges vaguely to μ.*

The proof is found in Section 3.5.

Remark 3.9. As for the case $1 \leq p < \infty$, Proposition 3.17 implies that OT_∞ metricizes the vague convergence, and thus using Proposition 3.15, we have that $(\mathcal{D}^\infty, d_\infty)$ is closed in $(\mathcal{M}^\infty, \mathrm{OT}_\infty)$ and is—in particular—complete.

In contrast to the case $p < \infty$, a converse implication for Proposition 3.17 does not hold, even on the subspace of persistence diagrams (see Figure 3.1). To recover a space with a structure more similar to the one of $\mathcal{D}^p$, it is useful to look at a smaller set. Introduce $\mathcal{D}_0^\infty$ the set of persistence diagrams such that for all $r > 0$, there are a finite number of points of the diagram of persistence larger than r and recall that $\mathcal{D}_f$ denotes the set of persistence diagrams with finite number of points.

Proposition 3.18. *The closure of $\mathcal{D}_f$ for the distance OT_∞ is $\mathcal{D}_0^\infty$.*

Proof. Consider $a \in \mathcal{D}_0^\infty$. By definition, for all $n \in \mathbb{N}$, a has a finite number of points with persistence larger than $\frac{1}{n}$, so that the restriction a_n of a to points with persistence larger than $\frac{1}{n}$ belongs to $\mathcal{D}_f$. As $\mathrm{OT}_\infty(a, a_n) \leq \frac{1}{n} \to 0$, $\mathcal{D}_0^\infty$ is contained in the closure of $\mathcal{D}_f$.

Conversely, consider a diagram $a \in \mathcal{D}^\infty \backslash \mathcal{D}_0^\infty$. There is a constant $r > 0$ such that a has infinitely many points with persistence larger than r. For any finite diagram $a' \in \mathcal{D}_f$, we have $\mathrm{OT}_\infty(a', a) \geq r$, thus a is not the limit for the OT_∞ metric of any sequence in $\mathcal{D}_f$. □

Remark 3.10. The space $\mathcal{D}_0^\infty$ is exactly the set introduced in [BGMP14, Theorem 3.5] as the completion of $\mathcal{D}_f$ for the bottleneck metric d_∞. Here,

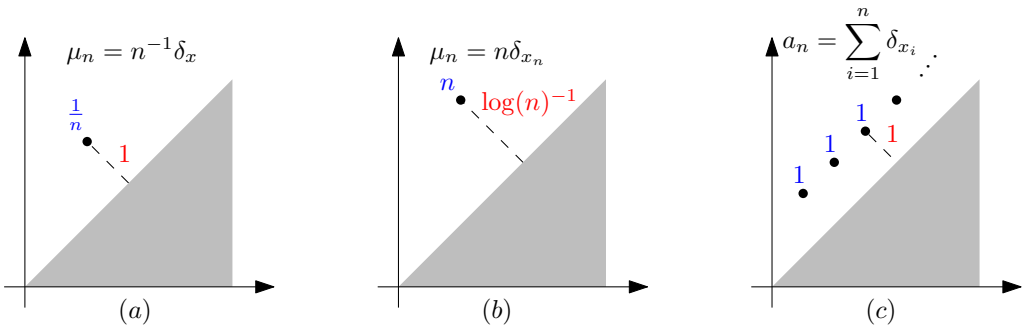


Figure 3.1: Illustration of differences between OT_p, OT_∞, and vague convergences. Blue color represents the mass on a point while red color designates distances. (a) A case where $\mathrm{OT}_p(\mu_n, 0) \to 0$ for any $p < \infty$ while $\mathrm{OT}_\infty(\mu_n, 0) = 1$. (b) A case where $\mathrm{OT}_\infty(\mu_n, 0) \to 0$ while for all $p < \infty$, $\mathrm{OT}_p(\mu_n, \mu) \to \infty$. (c) A sequence of persistence diagrams $a_n \in \mathcal{D}^\infty$, where $(a_n)_n$ converges vaguely to $a = \sum_i \delta_{x_i}$ and $\mathrm{Pers}_\infty(a_n) = \mathrm{Pers}_\infty(a)$, but (a_n) does not converge to a for OT_∞.

we recover that $\mathcal{D}_0^\infty$ is complete as a closed subset of the complete space $\mathcal{D}^\infty$.

Define for $r > 0$ and $a \in \mathcal{D}$, $a^{(r)}$ the persistence diagram restricted to $\{x \in \Omega,\ d(x, \partial\Omega) > r\}$ (as in Lemma 3.2). The following characterization of convergence holds in $\mathcal{D}_0^\infty$.

Proposition 3.19. *Let $a, a_1, a_2, \ldots$ be persistence diagrams in $\mathcal{D}_0^\infty$. Then,*

$$\mathrm{OT}_\infty(a_n, a) \to 0 \Leftrightarrow \begin{cases} a_n \xrightarrow{v} a, \\ (a_n^{(r)})_n \text{ is tight for all positive } r. \end{cases}$$

Proof of Proposition 3.19. Let us prove first the direct implication. Proposition 3.17 states that the convergence with respect to OT_∞ implies the vague convergence. Fix $r > 0$. By definition, $a^{(r)}$ is made of a finite number of points, all included in some open bounded set $U \subset \Omega$. As $a_n^{(r)}(U^c)$ is a sequence of integers, the bottleneck convergence implies that for n large enough, $a_n^{(r)}(U^c)$ is equal to 0. Thus, $(a_n^{(r)})_n$ is tight.

Let us prove the converse. Consider $a \in \mathcal{D}_0^\infty$ and a sequence $(a_n)_n$ that converges vaguely to a, with $(a_n^{(r)})$ tight for all $r > 0$. Fix $r > 0$ and let $x_1, \ldots, x_K$ be an enumeration of the points in $a^{(r)}$, the point x_k being present with multiplicity $m_k \in \mathbb{N}$. Denote by $B(x, \varepsilon)$ (resp. $\overline{B}(x, \varepsilon)$) the open (resp. closed) ball of radius ε centered at x. By Portmanteau theorem,

for ε small enough,

$$\begin{cases} \liminf_{n\to\infty} a_n(B(x_k,\varepsilon)) \geq a(B(x_k,\varepsilon)) = m_k \\ \limsup_{n\to\infty} a_n(\overline{B}(x_k,\varepsilon)) \leq a(\overline{B}(x_k,\varepsilon)) = m_k, \end{cases}$$

so that, for n large enough, there are exactly m_k points of a_n in $B(x_k,\varepsilon)$ (since $(a_n(B(x_k,\varepsilon)))_n$ is a converging sequence of integers). The tightness of $(a_n^{(r)})_n$ implies the existence of some compact $K \subset \Omega$ such that for n large enough, $a_n^{(r)}(K^c) = 0$ (as the measures take their values in $\mathbb{N}$). Applying Portmanteau's theorem to the closed set $K' := K \setminus \bigcup_{i=1}^{K} B(x_i,\varepsilon)$ gives

$$\limsup_{n\to\infty} a_n^{(r)}(K') \leq a^{(r)}(K') = 0.$$

This implies that for n large enough, there are no other point in a_n with persistence larger than r and thus $\mathrm{OT}_\infty(a^{(r)}, a_n)$ is smaller than $r+\varepsilon$. Finally,

$$\limsup_{n\to\infty} \mathrm{OT}_\infty(a_n, a) \leq \limsup_{n\to\infty} \mathrm{OT}_\infty(a_n, a^{(r)}) + r \leq 2r + \varepsilon.$$

Letting $\varepsilon \to 0$ then $r \to 0$, the bottleneck convergence holds. □

As for the case of OT_p metrics, this allows us to characterize relatively compact subsets of $\mathcal{D}_0^\infty$. A similar result can be found in [PMK19, Theorem 11]. Here again,

Corollary 3.20. *A set $F \subset \mathcal{D}_0^\infty$ is relatively compact for the metric OT_∞ if and only if for all compact K, $\sup\{a(K),\ a \in F\}$ is finite and for all $r > 0$, there exists a compact K such that $a^r(\Omega\setminus K) = 0$ for all $a \in F$.*

3.4 Duality

We end this chapter by providing a duality result for the OT_p metrics we introduced. Our approach essentially mimics the standard proof done in standard optimal transport [Vil08, Ch. 5] and also builds on some elements introduced in [FG10].

Let $c(x,y) := d(x,y)^p$ in the following. Define also

$$\tilde{c}(x,y) := \min\{c(x,y), c(x,\partial\Omega) + c(\partial\Omega,y)\} = \rho_p^p,$$

where ρ_p was introduced in Remark 3.5. Introduce

$$\mathcal{A} := \{(x,y) \in \overline{\Omega},\ \tilde{c}(x,y) = c(x,y)\}.$$

The following definitions are standard in convex optimization and are routinely used in optimal transport theory when it comes to study duality.

Definition 3.2. For $\varphi : \overline{\Omega} \to \mathbb{R}$, define its *c-transform* as

$$\varphi^c(y) = \inf_x c(x, y) - \varphi(x).$$

We say that φ is *c-concave* if it is the c-transform of some function ψ, that is $\varphi = \psi^c$.

Define then the *c-superdifferential* of a c-concave function φ as the set

$$\partial^c \varphi := \{(x, y) \in \overline{\Omega} \times \overline{\Omega},\ \varphi(x) + \varphi^c(y) = c(x, y)\}.$$

Finally, say that a subset X of $\overline{\Omega} \times \overline{\Omega}$ is *c-cyclically monotone* if for any $K \in \mathbb{N}$ and any finite sample $(x_1, y_1) \dots (x_K, y_K) \in X$, and any permutation $\sigma \in \mathfrak{S}_K$, one has

$$\sum_{i=1}^{K} c(x_i, y_i) \leq \sum_{i=1}^{K} c(x_i, y_{\sigma(i)}). \tag{3.15}$$

Notions of $\tilde{c}$-concavity, $\tilde{c}$-transform, $\tilde{c}$-superdifferential, $\tilde{c}$-cyclical monotonicity are defined similarly.

The following proposition in central to obtain a duality result.

Proposition 3.21. *Let $\mu, \nu \in \mathcal{M}_c$ and $\pi \in \mathrm{Adm}(\mu, \nu)$. We have equivalence between*

(i) $\pi \in \mathrm{Opt}(\mu, \nu)$.

(ii) $\mathrm{spt}(\pi) \subset \mathcal{A}$ *and* $\mathrm{spt}(\pi)$ *is $\tilde{c}$-cyclically monotone.*

(iii) There exists a c-concave function φ such that φ and φ^c both vanish on $\partial\Omega$ and $\mathrm{spt}(\pi) \subset \partial^c \varphi$.

As a corollary, if $\mu_n \xrightarrow{v} \mu$ and $\nu_n \xrightarrow{v} \nu$, and $\pi_n \in \mathrm{Opt}(\mu_n, \nu_n)$, then $\pi_n \xrightarrow{v} \pi \in \mathrm{Opt}(\mu, \nu)$.

The proof essentially follows [FG10] and can be found in Section 3.5.

Theorem 3.22 (Duality)**.** *One has*

$$L_c(\mu, \nu) := \inf_{\pi \in \mathrm{Adm}(\mu,\nu)} \langle c, \pi \rangle = \sup_{f,g \in \Psi_c} \langle f, \mu \rangle + \langle g, \nu \rangle \tag{3.16}$$

where Ψ_c denotes the set of continuous functions $f, g : \overline{\Omega} \to \mathbb{R}$ satisfying

- $\forall x, y,\ f(x) + g(y) \leq c(x, y)$.
- $\forall x, f(x) \leq c(x, \partial\Omega)$ *and* $\forall y, g(y) \leq c(\partial\Omega, y)$.

Proof. The proof relies on the equivalence between (i) and (iii) in Proposition 3.21. Let φ be a continuous function supported on $\overline{\Omega}$ that vanishes on the diagonal $\partial\Omega$. By definition of the c-transform, for all $x, y \in \overline{\Omega}$, one has $\varphi(x) + \varphi^c(y) \leq c(x, y)$. Therefore, for any $\pi \in \mathrm{Adm}(\mu, \nu)$, we have

$$\int_\Omega \varphi(x)\mathrm{d}\mu(x) + \int_\Omega \varphi^c(y)\mathrm{d}\nu(y) \leq \iint_{\overline{\Omega}\times\overline{\Omega}} c(x, y)\mathrm{d}\pi(x, y)$$

thanks to the marginal constraints (and that φ, φ^c both vanishes on the diagonal). Thus,

$$\sup_\varphi \langle \varphi, \mu\rangle + \langle \varphi^c, \nu\rangle \leq \inf_{\pi\in\mathrm{Adm}(\mu,\nu)} \langle c, \pi\rangle\,, \tag{3.17}$$

where the supremum in the left hand side is taken over the functions φ continuous that vanishes on the diagonal. This type of inequality are known as a *duality gap*.

Now, consider $\pi \in \mathrm{Opt}(\mu, \nu)$. We know thanks to (iii) that one can pick φ such that $\mathrm{spt}(\pi) \subset \partial^c\varphi$, which immediate gives that $\langle \varphi, \mu\rangle + \langle \varphi^c, \nu\rangle = \langle c, \pi\rangle = L_c(\mu, \nu)$. Therefore, it follows that the duality gap in (3.17) is actually 0.

Finally, to prove the claimed result, it suffices to observe that by definition of the c-transform, one has $g(y) \leq \varphi^c(y)$ for any function g that satisfies $\varphi(x) + g(y) \leq c(x, y)$, and thus (changing φ for f), we obtain the inequality. □

These results have some nice consequences in the particular case $p = 1$, that is $c(x, y) = d(x, y)$. One can observe (see [San15, Prop. 3.1]) that in this case, a function $\varphi : \overline{\Omega} \to \mathbb{R}$ is c-concave if and only if it is Lipschitz continuous with constant less than 1 (with respect to d), and for such a function φ, one has $\varphi^c = -\varphi$. Thus, we obtain the following Kantorovich-Rubinstein duality result:

Proposition 3.23. *Let $p = 1$, that is $c(x, y) = d(x, y) = \|x - y\|$, and $\mu, \nu \in \mathcal{M}^1$. We have*

$$\mathrm{OT}_1(\mu, \nu) = \max_{f\in\mathcal{L}} (\mu - \nu)(f), \tag{3.18}$$

where $\mathcal{L}$ denotes the set of Lipschitz continuous functions $f : \overline{\Omega} \to \mathbb{R}$ with Lipschitz constant less than 1 and that satisfy $f(\partial\Omega) = 0$.

Note that a similar duality result for finite persistence diagram metrics was proved in [BE19].

3.5 Proofs

Proofs of Proposition 3.1 and Proposition 3.14.

- For $\pi \in \mathrm{Adm}(\mu,\nu)$ supported on E_Ω, and for any compact sets $K,\ K' \subset \Omega$, one has $\pi((K\times\overline{\Omega})\cup(\overline{\Omega}\times K')) \leq \mu(K)+\nu(K') < \infty$. As any compact subset of E_Ω is included in a set of the form $(K\times\overline{\Omega})\cup(\overline{\Omega}\times K')$, Proposition B.1 implies that $\mathrm{Adm}(\mu,\nu)$ is relatively compact for the vague convergence on E_Ω. Also, if a sequence $(\pi_n)_n$ in $\mathrm{Adm}(\mu,\nu)$ converges vaguely to some $\pi \in \mathcal{M}(E_\Omega)$, then the marginals of π are still μ and ν, implying that $\mathrm{Adm}(\mu,\nu)$ is closed in $\mathcal{M}(E_\Omega)$: it is therefore sequentially compact.

- To prove the second point of Proposition 3.1, consider $\pi, \pi_1, \pi_2, \dots$ such that $\pi_n \xrightarrow{v} \pi$, and introduce $\pi'_n : A \mapsto \iint_A d(x,y)^p \mathrm{d}\pi_n$. The sequence $(\pi'_n)_n$ still converges vaguely to $\pi' : A \mapsto \iint_A d(x,y)^p \mathrm{d}\pi$. Portmanteau theorem (Proposition B.4) applied with the open set E_Ω to the measures $\pi'_n \xrightarrow{v} \pi'$ implies that $C_p(\pi) = \pi'(E_\Omega) \leq \liminf_n \pi'_n(E_\Omega) = \liminf_n C_p(\pi_n)$, i.e. C_p is lower semi-continuous.

- We now prove the lower semi-continuity of C_∞. Let $(\pi_n)_n$ be a sequence converging vaguely to π on E_Ω and let $r > \liminf_{n\to\infty} C_\infty(\pi_n)$. The set $U_r = \{(x,y) \in E_\Omega,\ d(x,y) > r\}$ is open. By Portmanteau theorem (Proposition B.4), we have $0 = \liminf_{n\to\infty} \pi_n(U_r) \geq \pi(U_r)$. Therefore, $\mathrm{spt}(\pi) \subset U_r^c$ and $C_\infty(\pi) \leq r$. As this holds for any $r > \liminf_{n\to\infty} C_\infty(\pi_n)$, we have $\liminf_{n\to\infty} C_\infty(\pi_n) \geq C_\infty(\pi)$.

- We show that for any $1 \leq p \leq \infty$, the lower semi-continuity of C_p and the sequential compactness of $\mathrm{Adm}(\mu,\nu)$ imply that 1. $\mathrm{Opt}_p(\mu,\nu)$ is a non-empty compact set for the vague topology on E_Ω and that 2. OT_p is lower semi-continuous.

  1. Let $(\pi_n)_n$ be a minimizing sequence of (2.22) in $\mathrm{Adm}(\mu,\nu)$. As $\mathrm{Adm}(\mu,\nu)$ is sequentially compact, it has an adherence value π, and the lower semi-continuity implies that

$$C_p(\pi) \leq \liminf_{n\to\infty} C_p(\pi_n) = \mathrm{OT}_p^p(\mu,\nu),$$

so that $\mathrm{Opt}_p(\mu,\nu)$ is non-empty. Using once again the lower semi-continuity of C_p, if a sequence in $\mathrm{Opt}_p(\mu,\nu)$ converges to some limit, then the cost of the limit is smaller than (and thus equal to) $\mathrm{OT}_p^p(\mu,\nu)$, i.e. the limit is in $\mathrm{Opt}_p(\mu,\nu)$. The set $\mathrm{Opt}_p(\mu,\nu)$ being closed in the sequentially compact set $\mathrm{Adm}(\mu,\nu)$, it is also sequentially compact.

2. Let $\mu_n \xrightarrow{v} \mu$ and $\nu_n \xrightarrow{v} \nu$. Up to taking a subsequence, on may assume that $\mathrm{OT}_p(\mu_n,\nu_n)$ converges (eventually to infinity). Consider $\pi_n \in \mathrm{Opt}_p(\mu_n,\nu_n)$. For any compact sets K, $K' \subset \Omega$, one has $\pi_n((K\times\overline{\Omega})\cup(\overline{\Omega}\times K')) \leq \sup_n \mu_n(K) + \sup_n \nu_n(K') < \infty$. Therefore, by Proposition B.1, there exists a subsequence $(\pi_{n_k})_k$ which converges vaguely to some $\pi \in \mathrm{Adm}(\mu,\nu)$. As the marginals of π_{n_k} converges to μ and ν, $\pi \in \mathrm{Adm}(\mu,\nu)$. Therefore,

$$\mathrm{OT}_p^p(\mu,\nu) \leq C_p(\pi) \leq \liminf_{n\to\infty} C_p(\pi_n) = \liminf_{n\to\infty} \mathrm{OT}_p^p(\mu_n,\nu_n).$$

- Finally, we prove that OT_p is a metric on $\mathcal{M}^p$. Let $\mu,\nu,\lambda \in \mathcal{M}^p$. The symmetry of OT_p is clear. If $\mathrm{OT}_p(\mu,\nu)=0$, then there exists $\pi \in \mathrm{Adm}(\mu,\nu)$ supported on $\{(x,x),\ x\in\Omega\}$. Therefore, for a Borel set $A\subset\Omega$, $\mu(A) = \pi(A\times\overline{\Omega}) = \pi(A\times A) = \pi(\overline{\Omega}\times A) = \nu(A)$, and $\mu=\nu$. To prove the triangle inequality, we need a variant on the gluing lemma, stated in [FG10, Lemma 2.1]: for $\pi_{12}\in\mathrm{Opt}(\mu,\nu)$ and $\pi_{23}\in\mathrm{Opt}(\nu,\lambda)$ there exists a measure $\gamma\in\mathcal{M}(\overline{\Omega}^3)$ such that the marginal corresponding to the first two entries (resp. two last entries), when restricted to E_Ω, is equal to π_{12} (resp. π_{23}), and induces a zero cost on $\partial\Omega\times\partial\Omega$. Therefore,

$$\begin{aligned}
\mathrm{OT}_p(\mu,\lambda) &\leq \left(\int_{\overline{\Omega}^2} d(x,z)^p \mathrm{d}\gamma(x,y,z)\right)^{1/p} \\
&\leq \left(\int_{\overline{\Omega}^2} d(x,y)^p \mathrm{d}\gamma(x,y,z)\right)^{1/p} + \left(\int_{\overline{\Omega}^2} d(y,z)^p \mathrm{d}\gamma(x,y,z)\right)^{1/p} \\
&= \left(\int_{\overline{\Omega}^2} d(x,y)^p \mathrm{d}\pi_{12}(x,y)\right)^{1/p} + \left(\int_{\overline{\Omega}^2} d(y,z)^p \mathrm{d}\pi_{23}(y,z)\right)^{1/p} \\
&= \mathrm{OT}_p(\mu,\nu) + \mathrm{OT}_p(\nu,\lambda).
\end{aligned}$$

The proof is similar for $p=\infty$.

□

Proof of Proposition 3.4. We first show the separability. Consider for $k>0$ a partition of Ω into squares (C_i^k) of side length 2^{-k}, centered at points

x_i^k. Let F be the set of all measures of the form $\sum_{i\in I} q_i \delta_{x_i^k}$ for q_i positive rationals, $k > 0$ and I a finite subset of $\mathbb{N}$. Our goal is to show that the countable set F is dense in $\mathcal{M}^p$. Fix $\varepsilon > 0$, and $\mu \in \mathcal{M}^p$. The proof is in three steps.

1. Since $\mathrm{Pers}_p(\mu) < \infty$, there exists a compact $K \subset \Omega$ such that $\mathrm{Pers}_p(\mu) - \mathrm{Pers}_p(\mu_0) < \varepsilon^p$, where μ_0 is the restriction of μ to K. By considering the transport plan between μ and μ_0 induced by the identity map on K and the projection onto the diagonal on $\overline{\Omega}\backslash K$, it follows that $\mathrm{OT}_p^p(\mu, \mu_0) \leq \mathrm{Pers}_p(\mu) - \mathrm{Pers}_p(\mu_0) \leq \varepsilon^p$.

2. Consider k such that $2^{-k} \leq \varepsilon/(\sqrt{2}\mu(K)^{1/p})$ and denote by I the indices corresponding to squares C_i^k intersecting K. Let us introduce $\mu_1 = \sum_{i\in I}^{\infty} \mu_0(C_i^k)\delta_{x_i^k}$. One can create a transport map between μ_0 and μ_1 by mapping each square C_i^k to its center x_i^k, so that
$$\mathrm{OT}_p(\mu_0, \mu_1) \leq \left(\sum_i \mu_0(C_i^k)(\sqrt{2} \cdot 2^{-k})^p \right)^{1/p} \leq \mu(K)^{1/p}\sqrt{2} \cdot 2^{-k} \leq \varepsilon.$$

3. Consider, for $i \in I$, q_i a rational number satisfying $q_i \leq \mu_0(C_i^k)$ and $|\mu_0(C_i^k) - q_i| \leq \varepsilon^p / \left(\sum_{i\in I} d(x_i^k, \partial\Omega)^p\right)$. Let $\mu_2 = \sum_{i\in I} q_i \delta_{x_i^k}$. Consider the transport plan between μ_2 and μ_1 that fully transports μ_2 onto μ_1, and transport the remaining mass in μ_1 onto the diagonal. Then,
$$\mathrm{OT}_p(\mu_1, \mu_2) \leq \left(\sum_{i\in I} |\mu_0(C_i^k) - q_i| d(x_i^k, \partial\Omega)^p \right)^{1/p} \leq \varepsilon.$$

As $\mu_2 \in F$ and $\mathrm{OT}_p(\mu, \mu_2) \leq 3\varepsilon$, the separability is proven.

To prove that the space is complete, consider a Cauchy sequence $(\mu_n)_n$. As $(\mathrm{Pers}_p(\mu_n))_n = (\mathrm{OT}_p^p(\mu_n, 0))_n$ is a Cauchy sequence, it is bounded. Therefore, for $K \subset \Omega$ a compact set, (3.6) implies that $\sup_n \mu_n(K) < \infty$. Proposition B.1 implies that $(\mu_n)_n$ is relatively compact for the vague convergence on Ω: consider $(\mu_{n_k})_k$ a subsequence converging vaguely on Ω to some measure μ. By the lower semi-continuity of OT_p,
$$\mathrm{Pers}_p(\mu) = \mathrm{OT}_p^p(\mu, 0) \leq \liminf_{k\to\infty} \mathrm{OT}_p^p(\mu_{n_k}, 0) < \infty,$$
so that $\mu \in \mathcal{M}^p$. Using once again the lower semi-continuity,
$$\mathrm{OT}_p(\mu_n, \mu) \leq \liminf_{k\to\infty} \mathrm{OT}_p(\mu_n, \mu_{n_k})$$
$$\lim_{n\to\infty} \mathrm{OT}_p(\mu_n, \mu) \leq \lim_{n\to\infty} \liminf_{k\to\infty} \mathrm{OT}_p(\mu_n, \mu_{n_k}) = 0,$$
ensuring that $\mathrm{OT}_p(\mu_n, \mu) \to 0$, that is the space is complete. □

Proof of the direct implication of Theorem 3.5. Let $\mu, \mu_1, \mu_2, \dots$ be elements of $\mathcal{M}^p$ and assume that the sequence $(\mathrm{OT}_p(\mu_n, \mu))_n$ converges to 0. The triangle inequality implies that $\mathrm{Pers}_p(\mu_n) = \mathrm{OT}_p^p(\mu_n, 0)$ converges to $\mathrm{Pers}_p(\mu) = \mathrm{OT}_p^p(\mu, 0)$. Let $f \in C_c(\Omega)$, whose support is included in some compact set K. For any $\varepsilon > 0$, there exists a Lipschitz function f_ε, with Lipschitz constant L and whose support is included in K, with the ∞-norm $\|f - f_\varepsilon\|_\infty$ smaller than ε. The convergence of $\mathrm{Pers}_p(\mu_n)$ and (3.6) imply that $\sup_k \mu_k(K) < \infty$. Let $\pi_n \in \mathrm{Opt}_p(\mu_n, \mu)$, we have

$$\begin{aligned}|\mu_n(f) - \mu(f)| &\leq |\mu_n(f - f_\varepsilon)| + |\mu(f - f_\varepsilon)| + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| \\ &\leq (\mu_n(K) + \mu(K))\varepsilon + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| \\ &\leq (\sup_k \mu_k(K) + \mu(K))\varepsilon + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)|.\end{aligned}$$

Also,

$$\begin{aligned}|\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| &\leq \iint_{\overline{\Omega}^2} |f_\varepsilon(x) - f_\varepsilon(y)| \mathrm{d}\pi_n(x, y) \qquad \text{where } \pi_n \in \mathrm{Opt}(\mu_n, \mu) \\ &\leq L \iint_{(K \times \overline{\Omega}) \cup (\overline{\Omega} \times K)} d(x, y) \mathrm{d}\pi_n(x, y) \\ &\leq L \pi_n((K \times \overline{\Omega}) \cup (\overline{\Omega} \times K))^{1 - \frac{1}{p}} \left(\iint_{(K \times \overline{\Omega}) \cup (\overline{\Omega} \times K)} d(x, y)^p \mathrm{d}\pi_n(x, y) \right)^{\frac{1}{p}} \\ &\qquad \text{by Hölder's inequality.} \\ &\leq L \left(\sup_k \mu_k(K) + \mu(K) \right)^{1 - \frac{1}{p}} \mathrm{OT}_p(\mu_n, \mu) \to 0.\end{aligned}$$

Therefore, taking the limsup in n and then letting $\varepsilon \to 0$, we obtain that $\mu_n(f) \to \mu(f)$. □

Proof of the converse implication of Theorem 3.5. Let $\mu, \mu_1, \mu_2 \dots$ be elements of $\mathcal{M}^p$ and assume that $\mu_n \xrightarrow{v} \mu$ and $\mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu)$. Since

$$\mathrm{OT}_p(\mu_n, \mu) \leq \mathrm{OT}_p(\mu_n, 0) + \mathrm{OT}_p(\mu, 0) = \mathrm{Pers}_p(\mu_n)^{1/p} + \mathrm{Pers}_p(\mu)^{1/p},$$

the sequence $(\mathrm{OT}_p(\mu_n, \mu))_n$ is bounded. Thus, if we show that $(\mathrm{OT}_p(\mu_n, \mu))_n$ admits 0 as unique accumulation point, then the convergence holds. Up to extracting a subsequence, we may assume that $(\mathrm{OT}_p(\mu_n, \mu))_n$ converges to some limit. Let $(\pi_n)_n \in \mathrm{Opt}(\mu_n, \mu)^{\mathbb{N}}$ be the corresponding optimal transport plans. Let K be a compact subset of Ω. Recall (Proposition B.1 in

Appendix B) that relative compactness for the vague convergence of a sequence $(\mu_n)_n$ is equivalent to $\sup_n\{\mu_n(K)\} < \infty$ for every compact $K \subset \Omega$. Therefore, for any compact $K \subset \Omega$, and $n \in \mathbb{N}$,

$$\pi_n((K \times \Omega) \cup (\Omega \times K)) \leq \mu_n(K) + \mu(K) \leq \sup_k \mu_k(K) + \mu(K) < \infty.$$

As any compact of E_Ω is included is some set of the form $(K \cup \Omega) \times (\Omega \cup K)$, for $K \subset \Omega$ compact, using Proposition B.1 again, it follows that $(\pi_n)_n$ is also relatively compact for the vague convergence.

Let thus π be the limit of any converging subsequence of $(\pi_n)_n$, which indexes are still denoted by n. As $\mu_n \xrightarrow{v} \mu$, π is necessarily in $\mathrm{Opt}_p(\mu, \mu)$ (see [FG10, Prop. 2.3]), i.e. π is supported on $\{(x, x), x \in \Omega\}$. The vague convergence of $(\mu_n)_n$ and the convergence of $(\mathrm{Pers}_p(\mu_n))_n$ to $\mathrm{Pers}_p(\mu)$ imply that for a given compact set $K \subset \Omega$, whose complementary set in Ω is denoted by K^c, its interior set is denoted by $\mathring{K}$, and its boundary by ∂K we have

$$\begin{aligned}
&\limsup_{n\to\infty} \int_{K^c} d(x, \partial\Omega)^p \mathrm{d}\mu_n(x) \\
&= \limsup_{n\to\infty} \left(\mathrm{Pers}_p(\mu_n) - \int_K d(x, \partial\Omega)^p \mathrm{d}\mu_n(x) \right) \\
&= \lim_n \mathrm{Pers}_p(\mu) - \liminf_n \int_{\mathring{K}} d(x, \partial\Omega)^p \mathrm{d}\mu_n(x) - \liminf_n \int_{\partial K} d(x, \partial\Omega)^p \mathrm{d}\mu_n(x) \\
&\leq \mathrm{Pers}_p(\mu) - \int_{\mathring{K}} d(x, \partial\Omega)^p \mathrm{d}\mu(x) \text{ by Portmanteau theorem (see Appendix B)} \\
&= \int_{\overline{K^c}} d(x, \partial\Omega)^p \mathrm{d}\mu(x).
\end{aligned}$$

Recall that $\mathrm{Pers}_p(\mu) < \infty$. Therefore, for $\varepsilon > 0$, there exists some compact set $K \subset \Omega$, with

$$\limsup_n \int_{K^c} d(x, \partial\Omega)^p \mathrm{d}\mu_n(x) < \varepsilon \quad \text{and} \quad \int_{K^c} d(x, \partial\Omega)^p \mathrm{d}\mu(x) < \varepsilon. \quad (3.19)$$

Now, consider the following transport plan $\tilde{\pi}_n$ (consider informally that what went from K to K^c and from K^c to K is now transported onto the

diagonal, while everything else is unchanged):

$$\begin{cases} \tilde{\pi}_n = \pi_n & \text{on } K^2 \sqcup (K^c)^2, \\ \tilde{\pi}_n = 0 & \text{on } K \times K^c \sqcup K^c \times K, \\ \tilde{\pi}_n(A \times B) = \pi_n(A \times (K^c \sqcup B)) & \text{for } A \subset K,\ B \subset \partial\Omega, \\ \tilde{\pi}_n(A \times B) = \pi_n(A \times (K \sqcup B)) & \text{for } A \subset K^c,\ B \subset \partial\Omega, \\ \tilde{\pi}_n(A \times B) = \pi_n((K^c \sqcup A) \times B) & \text{for } A \subset \partial\Omega,\ B \subset K, \\ \tilde{\pi}_n(A \times B) = \pi_n((K \sqcup A) \times B) & \text{for } A \subset \partial\Omega,\ B \subset K^c. \end{cases} \tag{3.20}$$

Note that $\tilde{\pi}_n \in \mathrm{Adm}(\mu_n, \mu)$: for instance, for $A \subset K$ a Borel set, $\tilde{\pi}_n(A \times \overline{\Omega}) = \pi_n(A \times K) + \pi_n(A \times (K^c \sqcup \partial\Omega)) = \pi_n(A \times \overline{\Omega}) = \mu_n(A)$, and it is shown likewise that the other constraints are satisfied. As $\tilde{\pi}_n$ is suboptimal, $\mathrm{OT}_p^p(\mu_n, \mu) \leq \int_{\overline{\Omega}^2} d(x,y)^p \mathrm{d}\tilde{\pi}_n(x,y)$. The latter integral is equal to a sum of different terms, and we show that each of them converges to 0. Assume without loss of generality that the compact set K belongs to an increasing sequence of compact sets whose union is Ω, with $\pi(\partial(K \times K)) = 0$ for all compacts of the sequence.

- We have $\iint_{K^2} d(x,y)^p \mathrm{d}\tilde{\pi}_n(x,y) = \iint_{K^2} d(x,y)^p \mathrm{d}\pi_n(x,y)$. The lim sup of the integral is smaller than $\iint_{K^2} d(x,y)^p \mathrm{d}\pi(x,y)$ by Portmanteau theorem (applied to the sequence $(d(x,y)^p \mathrm{d}\pi_n(x,y))_n$), and, recalling that π is supported on the diagonal of E_Ω, this integral is equal to 0.

- For optimality reasons, any optimal transport plan must be supported on $\{d(x,y)^p \leq d(x,\partial\Omega)^p + d(y,\partial\Omega)^p\}$ (this fact is detailed in [FG10, Prop. 2.3]). It follows that

$$\begin{aligned} \iint_{(K^c)^2} d(x,y)^p \mathrm{d}\tilde{\pi}_n(x,y) &= \iint_{(K^c)^2} d(x,y)^p \mathrm{d}\pi_n(x,y) \\ &\leq \int_{K^c} d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) + \int_{K^c} d(y,\partial\Omega)^p \mathrm{d}\mu(y). \end{aligned}$$

  Taking the lim sup in n, and then letting K go to Ω, this quantity converges to 0 by (3.19).

- We have

$$\iint_{K \times \partial\Omega} d(x,\partial\Omega)^p \mathrm{d}\tilde{\pi}_n(x,y) = \int_K d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) - \iint_{K^2} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y).$$

  By Portmanteau theorem applied to the sequence $(d(x,\partial\Omega)^p \mathrm{d}\mu_n(x))_n$, we obtain that

$$\limsup_n \int_K d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) \leq \int_K d(x,\partial\Omega)^p \mathrm{d}\mu(x).$$

Applying once again Portmanteau theorem on the second term to the sequence $(d(x,y)^p \mathrm{d}\pi_n(x,y))_n$, and using that π is supported on the diagonal of E_Ω, we have

$$\begin{aligned}\limsup_n -\iint_{K^2} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) &\leq -\iint_{K^2} d(x,\partial\Omega)^p \mathrm{d}\pi(x,y)\\ &= -\int_K d(x,\partial\Omega)^p \mathrm{d}\mu(x)\end{aligned}$$

as we recall recall that $\pi(\partial(K\times K))=0$. Therefore, the lim sup of this (non-negative) integral is equal to 0.

- The three remaining terms (corresponding to the three last lines of the definition (3.20)) are treated likewise.

Finally, we have proven that $(\mathrm{OT}_p(\mu_n,\mu))_n$ is bounded and that for any converging subsequence $(\mu_{n_k})_k$, $\mathrm{OT}_p(\mu_{n_k},\mu)$ converges to 0. It follows that $\mathrm{OT}_p(\mu_n,\mu)\to 0$. □

Proof of Lemma 3.12.

1. Consider $\pi\in\mathrm{Adm}(\mu,\nu)$, and define π' that coincides with π on $\Omega\times\Omega$, and is such that we enforce that the mass located at some $x\in\Omega$ transported on the diagonal to be transported on its orthogonal projection $\mathbf{p}_{\partial\Omega}(x)$: more precisely, for all Borel set $A\subset\Omega$, $B\subset\partial\Omega$, $\pi'(A\times B)=\pi((\mathbf{p}_{\partial\Omega}^{-1}(B)\cap A)\times B)$ and $\pi'(B\times A)=\pi(B\times(\mathbf{p}_{\partial\Omega}^{-1}(B)\cap A))$. Note that $\pi'\in T(\mu,\nu)$. Since $\mathbf{p}_{\partial\Omega}(x)$ is the unique minimizer of $y\mapsto d(x,y)^p$, it follows that $C_p(\pi')\leq C_p(\pi)$, with equality if and only if $\pi\in T(\mu,\nu)$, and thus $\mathrm{Opt}_p(\mu,\nu)\subset T(\mu,\nu)$.

2. Write $\tilde\pi=\iota(\pi)$. The mass $\tilde\pi(\{\partial\Omega\}\times\{\partial\Omega\})$ is non-negative by definition. One has for all Borel sets $A\subset\Omega$,

$$\begin{aligned}\tilde\pi(A\times\tilde\Omega) &= \tilde\pi(A\times\Omega)+\tilde\pi(A\times\{\partial\Omega\})\\ &= \pi(A\times\Omega)+\pi(A\times\partial\Omega)=\pi(A\times\overline{\Omega})=\mu(A)=\tilde\mu(A).\end{aligned}$$

Similarly, $\tilde\pi(\tilde\Omega\times B)=\tilde\nu(B)$ for all $B\subset\Omega$. Observe also that

$$\tilde\pi(\{\partial\Omega\}\times\tilde\Omega)=\tilde\pi(\{\partial\Omega\}\times\{\partial\Omega\})+\tilde\pi(\{\partial\Omega\}\times\Omega)=r-\mu(\Omega)=\tilde\mu(\{\partial\Omega\}).$$

Similarly, $\tilde{\pi}(\tilde{\Omega} \times \{\partial\Omega\}) = \tilde{\nu}(\{\partial\Omega\})$. It gives that $\iota(\pi) \in \Pi(\tilde{\mu}, \tilde{\nu})$, so that ι is well defined. Observe that

$$\begin{aligned}\iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}(x,y) = &\iint_{\Omega\times\Omega} d(x,y)^p \mathrm{d}\pi(x,y) \\ &+ \int_{\Omega} d(x,\partial\Omega)^p \mathrm{d}\pi(x,\partial\Omega) \\ &+ \int_{\Omega} d(\partial\Omega,y)^p \mathrm{d}\pi(\partial\Omega,y) + 0 \\ &= C_p(\pi) \text{ as } \pi \in T(\mu,\nu).\end{aligned}$$

3. Write $\pi = \kappa(\tilde{\pi})$. For $A \subset \Omega$ a Borel set,

$$\begin{aligned}\pi(A \times \overline{\Omega}) &= \pi(A \times \Omega) + \pi(A \times \partial\Omega) \\ &= \tilde{\pi}(A \times \Omega) + \tilde{\pi}(A \times \{\partial\Omega\}) = \tilde{\pi}(A \times \tilde{\Omega}) = \mu(A).\end{aligned}$$

Similarly, $\pi(\overline{\Omega} \times B) = \nu(B)$ for all $B \subset \Omega$. Therefore, $\pi \in \mathrm{Adm}(\mu,\nu)$, and by construction, if a point $x \in \Omega$ is transported on $\partial\Omega$, it is transported on $\mathbf{p}_{\partial\Omega}(x)$, so that $\pi \in T(\mu,\nu)$. Observe that $\mu(\Omega) + \pi(\partial\Omega\times\Omega) \leq \tilde{\pi}(\tilde{\Omega}\times\tilde{\Omega}) = r$, so that $\iota(\pi)$ is well defined. Also, $\iota(\pi) = \tilde{\pi}$, so that, according to point 2, $C_p(\pi) = \iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}(x,y)$.

□

Proof of Lemma 3.13. Let $\tilde{\pi} \in \Pi(\tilde{\mu}, \tilde{\nu})$. Define the set $H := \{(x,y) \in \tilde{\Omega}^2,\ \rho(x,y) = d(x,y)\}$, and let H^c be its complementary set in $\tilde{\Omega}^2$, i.e. the set where $\rho(x,y) = d(x,\partial\Omega) + d(\partial\Omega,y)$. Define $\tilde{\pi}' \in \mathcal{M}(\tilde{\Omega}^2)$ by, for Borel sets $A, B \subset \Omega$:

$$\begin{cases}\tilde{\pi}'(A \times B) = \tilde{\pi}((A \times B) \cap H) \\ \tilde{\pi}'(A \times \{\partial\Omega\}) = \tilde{\pi}((A \times \tilde{\Omega}) \cap H^c) + \tilde{\pi}(A \times \{\partial\Omega\}) \\ \tilde{\pi}'(\{\partial\Omega\} \times B) = \tilde{\pi}((\tilde{\Omega} \times B) \cap H^c) + \tilde{\pi}(\{\partial\Omega\} \times B).\end{cases}$$

We easily check that $\tilde{\pi}' \in \Pi(\tilde{\mu}, \tilde{\nu})$. Also, using $(a+b)^p \geq a^p + b^p$ for positive

a, b, we have

$$\begin{aligned}\iint_{\tilde{\Omega}\times\tilde{\Omega}} \rho(x,y)^p \mathrm{d}\tilde{\pi}(x,y) &= \iint_H d(x,y)^p \mathrm{d}\tilde{\pi}(x,y) \\ &\quad + \iint_{H^c} (d(x,\partial\Omega) + d(\partial\Omega,y))^p \mathrm{d}\tilde{\pi}(x,y) \\ &\geq \iint_H d(x,y)^p \mathrm{d}\tilde{\pi}'(x,y) \\ &\quad + \iint_{H^c} \left(d(x,\partial\Omega)^p + d(y,\partial\Omega)^p\right) \mathrm{d}\tilde{\pi}(x,y) \\ &= \iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}'(x,y) \\ &\geq \inf_{\tilde{\pi}'\in\Pi(\tilde{\mu},\tilde{\nu})} \iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}'(x,y).\end{aligned}$$

We conclude by taking the infimum on $\tilde{\pi}$ that

$$W_{p,\rho}(\tilde{\mu},\tilde{\nu}) \geq \inf_{\tilde{\pi}'\in\Pi(\tilde{\mu},\tilde{\nu})} \iint_{\tilde{\Omega}\times\tilde{\Omega}} d(x,y)^p \mathrm{d}\tilde{\pi}'(x,y).$$

Since $\rho(x,y) \leq d(x,y)$, it follows that

$$W_{p,\rho}^p(\tilde{\mu},\tilde{\nu}) = \inf_{\tilde{\pi}\in\Pi(\tilde{\mu},\tilde{\nu})} \iint_{\tilde{\Omega}^2} d(x,y)^p \mathrm{d}\tilde{\pi}(x,y). \tag{3.21}$$

Since d is continuous, the infimum in the right hand side of (3.21) is reached [Vil08, Theorem 4.1]. Consider thus $\tilde{\pi} \in \Pi(\tilde{\mu},\tilde{\nu})$ which realizes the infimum. We can write, using Lemma 3.12,

$$\begin{aligned}W_{p,\rho}^p(\tilde{\mu},\tilde{\nu}) &= \iint_{\tilde{\Omega}^2} d(x,y)^p \mathrm{d}\tilde{\pi}(x,y) = \iint_{\overline{\Omega}\times\overline{\Omega}} d(x,y)^p \mathrm{d}\kappa(\tilde{\pi})(x,y) \\ &\geq \inf_{\pi\in T(\mu,\nu)} \iint_{\overline{\Omega}\times\overline{\Omega}} d(x,y)^p \mathrm{d}\pi(x,y) = \mathrm{OT}_p^p(\mu,\nu),\end{aligned}$$

which concludes the proof. □

Proof of Proposition 3.17. Let $f \in C_c(\Omega)$, whose support is included in some compact set K. For any $\varepsilon > 0$, there exists a L-Lipschitz function f_ε, whose support is included in K, with $\|f - f_\varepsilon\|_\infty \leq \varepsilon$. Observe that $\sup_k \mu_k(K) < \infty$ using the same arguments than for (3.14). Let $\pi_n \in \mathrm{Opt}_\infty(\mu_n,\mu)$. We have

$$\begin{aligned}|\mu_n(f) - \mu(f)| &\leq |\mu_n(f - f_\varepsilon)| + |\mu(f - f_\varepsilon)| + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| \\ &\leq (\mu_n(K) + \mu(K))\varepsilon + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| \\ &\leq (\sup_k \mu_k(K) + \mu(K))\varepsilon + |\mu_n(f_\varepsilon) - \mu(f_\varepsilon)|.\end{aligned}$$

Also,

$$\begin{aligned}|\mu_n(f_\varepsilon) - \mu(f_\varepsilon)| &\leq \iint_{\overline{\Omega}^2} |f_\varepsilon(x) - f_\varepsilon(y)| \mathrm{d}\pi_n(x,y) \\ &\leq L \iint_{(K\times\overline{\Omega})\cup(\overline{\Omega}\times K)} d(x,y)\mathrm{d}\pi_n(x,y) \\ &\leq LC_\infty(\pi_n)(\pi_n(K\times\overline{\Omega}) + \pi_n(\overline{\Omega}\times K)) \\ &\leq L\mathrm{OT}_\infty(\mu_n,\mu)\left(\sup_k \mu_k(K) + \mu(K)\right) \to 0.\end{aligned}$$

Therefore, taking the lim sup in n and then letting ε go to 0, we obtain that $\mu_n(f) \to \mu(f)$. □

Proof of Proposition 3.21. (i) ⇒ (ii). Let $\pi \in \mathrm{Opt}(\mu,\nu)$. Define π' as in [FG10, Eq. (7)], that is

$$\pi' := \pi|_{\mathcal{A}} + (p_1, \mathbf{p}_{\partial\Omega}\circ p_1)_\# \pi|_{\overline{\Omega}\times\overline{\Omega}\setminus\mathcal{A}} + (\mathbf{p}_{\partial\Omega}\circ p_2, p_2)_\#\pi|_{\overline{\Omega}\times\overline{\Omega}\setminus\mathcal{A}},$$

where $\mathbf{p}_{\partial\Omega}$ is the orthogonal projection onto the diagonal, and p_1, p_2 are the projections of a measure onto its first and second marginal respectively. One can observe that $\pi \in \mathrm{Adm}(\mu,\nu)$, that $\mathrm{spt}(\pi) \subset \mathcal{A}$, and that $C_p(\pi') \leq C_p(\pi)$ with a strict inequality if $\mathrm{spt}(\pi) \not\subset \mathcal{A}$. As such, any optimal transport plan must be supported on $\mathcal{A}$.

In order to prove $\tilde{c}$-cyclical monotonicity, we first introduce μ_n, ν_n two measures with finite mass. In this context, we know from Proposition 3.11 and Remark 3.5 that $\mathrm{OT}_p(\mu_n,\nu_n) = W_{p,\tilde{c}}(\tilde{\mu}_n,\tilde{\nu}_n)$ and that an optimal transport plan $\tilde{\pi}_n$ between $\tilde{\mu}_n$ and $\tilde{\nu}_n$ for the cost $\tilde{c}$ induces an optimal transport plan π_n between μ_n and ν_n for the cost c. Using [Vil08, Thm. 5.10, part (ii)], it follows that $\mathrm{spt}(\tilde{\pi}_n)$ is $\tilde{c}$-cyclically monotone, and thus that $\mathrm{spt}(\pi_n)$ is also $\tilde{c}$-cyclically monotone. Thus, consider $\mu,\nu \in \mathcal{M}^p$ and two sequences $\mu_n \xrightarrow{v} \mu$, $\nu_n \xrightarrow{v} \nu$. Consider $\pi_n \in \mathrm{Opt}(\mu_n,\nu_n)$ and let π be the limit of a subsequence for the vague convergence. Having that $\mathrm{spt}(\pi_n)$ is $\tilde{c}$-cyclically monotone means that for all $K \in \mathbb{N}$, the measure $\pi_n^{\otimes K}$ is supported on set $\{(x_i,y_i)_{i=1}^K$ satisfying (3.15) $\}$, which is closed as $\tilde{c}$ is continuous. Therefore, it follows that $\pi^{\otimes K}$ is also supported on this set and $\mathrm{spt}(\pi)$ is $\tilde{c}$-cyclically monotone also.

In order to prove (ii) ⇒ (iii), as remarked in [FG10], it suffices to observe that these statements do not deal with the mass of π and its marginals, but only about its support. As such, one can just reproduce the Step 3 of the proof of [Vil08, Thm. 5.10, part (i)] along with the computations

made in the proof of [FG10, Prop. 2.3]. More precisely, it suffices to fix $(x_0, y_0) \in \mathrm{spt}(\pi)$ and to define

$$\varphi(x) := \sup_{K\in\mathbb{N}} \sup\{[\tilde{c}(x_1, y_0) - \tilde{c}(x_0, y_0)] + [\tilde{c}(x_2, y_1) - \tilde{c}(x_1, y_1)] \\ + \cdots + \tilde{c}(x, y_K) - \tilde{c}(x_K, y_K)]\},$$

where the second supremum is taken over $(x_1, y_1) \dots (x_K, y_K) \in \mathrm{spt}(\pi)$. Now, φ is $\tilde{c}$-concave and that $\mathrm{spt}(\pi) \subset \partial^{\tilde{c}}\varphi$. Furthermore, for $x, y \in \partial\Omega$, one has $\varphi(x) + \varphi^{\tilde{c}}(y) = \tilde{c}(x, y) = 0$. As such, up to adding a constant, we have that φ and $\varphi^{\tilde{c}}$ both vanish on $\partial\Omega$. To prove that these observations remain true when replacing $\tilde{c}$ by c, following [FG10], we observe that $\mathrm{spt}(\pi) \subset \partial^{\tilde{c}}\varphi \cap \mathcal{A} \subset \partial^c\varphi$. Then, recalling that $\varphi(x) = 0$ for $x \in \partial\Omega$, it follows $\varphi^c(x) = \varphi(x) + \varphi^c(x) = c(x, x) = 0$, proving that φ^c also vanishes on $\partial\Omega$.

Finally, to prove (iii) $\Rightarrow$ (i), consider $\pi, \pi' \in \mathrm{Adm}(\mu, \nu)$ (supported on E_Ω) and assume that $\pi \subset \partial^c\varphi$ for some c-concave function φ with φ, φ^c that vanish on $\partial\Omega$. Observe that for all x, y,

$$\varphi(x) \le c(x, \partial\Omega), \quad , \varphi^c(y) \le c(\partial\Omega, y),$$

so that $\mu(\varphi), \mu(\varphi^c) < \infty$ for any $\mu \in \mathcal{M}^p$. Finally, observe that for any $x, y \in \overline{\Omega}$, $\varphi(x) + \varphi^c(y) \le c(x, y)$ (while equality holds on $\partial^c\varphi$). It allows us to write

$$\begin{aligned} C_p(\pi) &= \iint_{\partial^c\varphi} c(x, y)\mathrm{d}\pi(x, y) \\ &= \iint_{\Omega\times\overline{\Omega}} \varphi(x)\mathrm{d}\pi(x, y) + \iint_{\overline{\Omega}\times\Omega} \varphi^c(y)\mathrm{d}\pi(x, y) \\ &= \int \varphi(x)\mathrm{d}\mu(x) + \int \varphi^c(y)\mathrm{d}\nu(y) \\ &= \iint_{\Omega\times\overline{\Omega}} \varphi(x)\mathrm{d}\pi'(x, y) + \iint_{\overline{\Omega}\times\Omega} \varphi^c(y)\mathrm{d}\pi'(x, y) \\ &= \iint_{\overline{\Omega}\times\overline{\Omega}} (\varphi(x) + \varphi^c(y))\mathrm{d}\pi'(x, y) \\ &\le \iint c(x, y)\mathrm{d}\pi'(x, y) = C_p(\pi'). \end{aligned}$$

Therefore, $\pi \in \mathrm{Opt}(\mu, \nu)$.

The last assertion is known as *stability of optimal transport* in standard optimal transport community [Vil08, Thm. 5.20], and mentioned in [FG10]. It is an immediate corollary of the equivalence between (i) and (ii): let

$\pi_n \in \mathrm{Opt}(\mu_n, \nu_n)$. As such, for all n, the support of π_n is a subset of $\mathcal{A}$ and is $\tilde{c}$-cyclically monotone. Up to extracting a subsequence, $(\pi_n)_n$ converges vaguely to some π, which also has its support included in $\mathcal{A}$ and $\tilde{c}$-cyclically monotone, as $\tilde{c}$ is continuous. Therefore, $\pi \in \mathrm{Opt}(\mu, \nu)$. □

Chapter 4

Fréchet means in the space of persistence measures

Abstract

In this chapter, we state the existence of Fréchet means for probability distributions supported on $\mathcal{M}^p$. We start with the finite case (i.e. averaging finitely many persistence measures) and then extend the result to any probability distribution with finite p-th moment. We then study the specific case of distribution supported on $\mathcal{D}^p$ (i.e. averaging persistence diagrams) and show that in the finite setting, the set of Fréchet means is a convex set whose extreme points are in $\mathcal{D}^d$ (i.e. are actual persistence diagrams). As in the previous chapter, some technical proofs have been deferred to a specific section (§4.4). Note that computational aspects are not addressed here: they will be covered in the next part of the manuscript, which is dedicated to applications.

Remark 4.1. In this chapter, we will assume that $1 < p < \infty$ and $1 < q < \infty$ (recall $d(\cdot, \cdot) = \| \cdot - \cdot \|_q$). These assumptions will ensure that (i) the projection of $x \in \Omega$ onto $\partial\Omega$ is uniquely defined and (ii) the Fréchet mean of k points $x_1 \dots x_k$ in $\overline{\Omega}$, i.e. minimizer of $x \mapsto \sum_{i=1}^k \|x - x_i\|_q^p$, is also uniquely defined; two facts used in our proofs.

Recall that $(\mathcal{M}^p, \mathrm{OT}_p)$ is a Polish space, and let W_p denote the Wasserstein distance (see Section 2.2.2) between probability measures supported on $(\mathcal{M}^p, \mathrm{OT}_p)$. We denote by $\mathcal{W}^p(\mathcal{M}^p)$ the space of probability measures P supported on $\mathcal{M}^p$, equipped with the W_p metric, which are at a finite

distance from $\delta_{\mathbf{0}}$—the Dirac mass supported on the empty diagram—i.e.

$$W_p^p(P, \delta_{\mathbf{0}}) = \int_{\nu \in \mathcal{M}^p} \mathrm{OT}_p^p(\nu, \mathbf{0}) \mathrm{d}P(\nu) = \int_{\nu \in \mathcal{M}^p} \mathrm{Pers}_p(\nu) \mathrm{d}P(\nu) < \infty.$$

Definition 4.1. Consider $P \in \mathcal{W}^p(\mathcal{M}^p)$. A measure $\mu^* \in \mathcal{M}^p$ is a *Fréchet mean* of P if it minimizes the *energy functional*

$$\mathcal{E} : \mu \in \mathcal{M}^p \mapsto \int_{\nu \in \mathcal{M}^p} \mathrm{OT}_p^p(\mu, \nu) \mathrm{d}P(\mu).$$

4.1 Fréchet means in the finite case

Let P be of the form $\sum_{i=1}^N \lambda_i \delta_{\mu_i}$ with $N \in \mathbb{N}$, μ_i a persistence measure of finite mass m_i, and $(\lambda_i)_i$ non-negative weights that sum to 1. Let $\mathcal{E}$ be the corresponding energy functional. Define $m_{\mathrm{tot}} := \sum_{i=1}^N m_i$. To prove the existence of Fréchet means for such a P, we show that, in this case, Fréchet means correspond to Fréchet means for the Wasserstein distance of some distribution on $\mathcal{M}^p_{m_{\mathrm{tot}}}(\tilde{\Omega})$, the set of measures on $\tilde{\Omega}$ that all have the same mass m_{tot} (see Section 3.2), a problem well studied in the literature [AC11, CE10, COO15a]. Recall from Section 3.2 that $\tilde{\Omega}$ denotes the closed half-plane $\overline{\Omega}$ where the points of the boundary $\partial\Omega$ have been identified. This space is naturally equipped with a metric ρ deduced from d by enforcing the triangle inequality, that is

$$\rho(x, y) = \min(d(x, y), d(x, \mathbf{p}_{\partial\Omega}(x)) + d(\mathbf{p}_{\partial\Omega}(y), y)).$$

We start with a lemma which claims that if a measure μ has too much mass (larger than m_{tot}), then it cannot be a Fréchet mean of $\mu_1 \dots \mu_N$. Its proof can be find in Section 4.4. The idea of the proof is to show that if a measure μ has some mass that is mapped to the diagonal in each transport plan between μ and μ_i, then we can build a measure μ' by "removing" this mass, and then observe that such a measure μ' has a smaller energy $\mathcal{E}(\mu')$.

Lemma 4.1. *We have* $\inf\{\mathcal{E}(\mu),\ \mu \in \mathcal{M}^p\} = \inf\{\mathcal{E}(\mu), \mu \in \mathcal{M}^p_{\leq m_{\mathrm{tot}}}\}$.

Let $W_{p,\rho}$ denote the Wasserstein distance between measures with same mass supported on the metric space $(\tilde{\Omega}, \rho)$ (see Sections 2.2.2 and 3.2).

Proposition 4.2. *Let*

$$\begin{aligned} \Psi : \mathcal{M}^p_{\leq m_{\mathrm{tot}}} &\to \mathcal{M}^p_{m_{\mathrm{tot}}}(\tilde{\Omega}) \\ \mu &\mapsto \mu + (m_{\mathrm{tot}} - \mu(\Omega))\delta_{\partial\Omega}. \end{aligned}$$

The functionals

$$\mathcal{E} : \mu \in \mathcal{M}^p_{\leq m_{\text{tot}}} \mapsto \sum_{i=1}^{N} \lambda_i \text{OT}^p_p(\mu, \mu_i)$$

and

$$\mathcal{F} : \tilde{\mu} \in \mathcal{M}^p_{m_{\text{tot}}}(\tilde{\Omega}) \mapsto \sum_{i=1}^{N} \lambda_i W^p_{p,\rho}\left(\tilde{\mu}, \Psi(\mu_i)\right),$$

have the same infimum values and $\arg\min \mathcal{E} = \Psi^{-1}(\arg\min \mathcal{F})$.

The existence of minimizers $\tilde{\mu}$ of $\mathcal{F}$, that is "Wasserstein barycenter" (i.e. Fréchet means for the Wasserstein distance) of $\tilde{P} := \sum_{i=1}^{N} \lambda_i \delta_{\tilde{\mu}_i}$, is well-known (see [AC11, Theorem 8]). Proposition 4.2 asserts that $\Psi^{-1}(\tilde{\mu})$ is a minimizer of $\mathcal{E}$ on $\mathcal{M}^p_{\leq m_{\text{tot}}}$, and thus a Fréchet mean of P according to Lemma 4.1. We therefore have the following corollary, namely the existence of Fréchet means in the finite case.

Corollary 4.3. *Let $\mu_1 \dots \mu_N$ be N persistence measures in $\mathcal{M}^p$, and let $\lambda_1 \dots \lambda_N$ be non-negative weights that sum to one. The distribution $P = \sum_{i=1}^{N} \lambda_i \delta_{\mu_i} \in \mathcal{W}^p(\mathcal{M}^p)$ admits a Fréchet mean.*

4.2 Existence and consistency of Fréchet means

We now extend the results of the previous section to the Fréchet means of general probability distributions supported on $\mathcal{M}^p$. First, we show a *consistency* result, in the vein of [LGL16, Theorem 3].

Proposition 4.4. *Let $(P_n)_n, P$ be probability measures in $\mathcal{W}^p(\mathcal{M}^p)$. Assume that each P_n has a Fréchet mean μ_n and that $W_p(P_n, P) \to 0$. Then, the sequence $(\mu_n)_n$ is relatively compact in $(\mathcal{M}^p, \text{OT}_p)$, and any limit of a converging subsequence is a Fréchet mean of P.*

To prove this result, we introduce the following lemma, whose proof can be found in Section 4.4.

Lemma 4.5. *Let $\mu, \mu_1, \mu_2, \dots \in \mathcal{M}^p$. Then, $\text{OT}_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \xrightarrow{v} \mu$ and there exists a persistence measure $\nu \in \mathcal{M}^p$ such that $\text{OT}_p(\mu_n, \nu) \to \text{OT}_p(\mu, \nu)$.*

Proof of Proposition 4.4. We first prove the relative compactness of $(\mu_n)_n$. Following Proposition B.1, we can show that for every compact $K \subset \Omega$, $\sup\{\mu_n(K)\} < \infty$. Consider a compact set $K \subset \Omega$. We have, because of (3.6),

$$\begin{aligned}\mu_n(K)^{\frac{1}{p}} &\leq \frac{1}{d(K,\partial\Omega)}\mathrm{OT}_p(\mu_n, \mathbf{0}) = \frac{1}{d(K,\partial\Omega)}W_p(\delta_{\mu_n}, \delta_{\mathbf{0}}) \\ &\leq \frac{1}{d(K,\partial\Omega)}\left(W_p(\delta_{\mu_n}, P_n) + W_p(P_n, \delta_{\mathbf{0}})\right)\end{aligned}$$

Since μ_n is a Fréchet mean of P_n, it minimizes $\{W_p(\delta_\nu, P_n),\ \nu \in \mathcal{M}^p\}$, and in particular $W_p(\delta_{\mu_n}, P_n) \leq W_p(\delta_{\mathbf{0}}, P_n)$. Furthermore, since we assume that $W_p(P_n, P) \to 0$, we have in particular that $\sup_n W_p(P_n,\ \delta_{\mathbf{0}}) < \infty$. As a consequence $\sup_n \mu_n(K) < \infty$, and it follows (Proposition B.1) that the sequence $(\mu_n)_n$ is relatively compact for the vague convergence.

Without relabeling, assume thus that $(\mu_n)_n \xrightarrow{v} \mu$. One must proof that μ is a Fréchet mean of P. At this stage, we can follow the same computations as [LGL16, Proof of Theorem 3]. Let ν be a persistence measure. We write

$$\begin{aligned}\int \mathrm{OT}_p^p(\nu, \xi)\mathrm{d}P(\xi) &= W_p^p(\delta_\nu, P) \\ &= \lim_{n\to\infty} W_p^p(\delta_\nu, P_n) \\ &\geq \lim_{n\to\infty} W_p^p(\delta_{\mu_n}, P_n) \\ &= \lim_{n\to\infty} \int \mathrm{OT}_p^p(\mu_n, \xi)\mathrm{d}P_n(\xi) \\ &\geq \int \liminf_{n\to\infty} \mathrm{OT}_p^p(\mu_n, \xi)\mathrm{d}P_n(\xi) \\ &\geq \int \mathrm{OT}_p^p(\mu, \xi)\mathrm{d}P(\xi) \qquad \text{by lower semi-continuity.}\end{aligned}$$

This proves that μ must be a barycenter of P. Furthermore, taking $\nu = \mu$ in the above series of inequalities yields

$$W_p(\delta_{\mu_n}, P_n) \to W_p(\delta_\mu, P),$$

and then by triangle inequality, and using that $W_p(P_n, P) \to 0$,

$$W_p(\delta_{\mu_n}, P) \to W_p(\delta_\mu, P).$$

Therefore, we can write

$$\begin{aligned}\int \mathrm{OT}_p^p(\mu,\xi)\mathrm{d}P(\xi) &= W_p^p(\delta_\mu, P)\\ &= \lim_{n\to\infty} W_p^p(\delta_{\mu_n}, P)\\ &= \lim_{n\to\infty} \int \mathrm{OT}_p^p(\mu_n,\xi)\mathrm{d}P(\xi)\\ &\geq \int \liminf_{n\to\infty} \mathrm{OT}_p^p(\mu_n,\xi)\mathrm{d}P(\xi)\\ &\geq \int \mathrm{OT}_p^p(\mu,\xi)\mathrm{d}P(\xi),\end{aligned}$$

forcing in particular that

$$\int \underbrace{\liminf_{n\to\infty} \mathrm{OT}_p^p(\mu_n,\xi) - \mathrm{OT}_p^p(\mu,\xi)}_{\geq 0}\,\mathrm{d}P(\xi) = 0,$$

and thus, P-almost surely, we have

$$\liminf_{n\to\infty} \mathrm{OT}_p^p(\mu_n,\xi) = \mathrm{OT}_p^p(\mu,\xi).$$

We eventually obtain the existence of some $\nu \in \mathcal{M}^p$ such that, up to a subsequence,

$$\mathrm{OT}_p(\mu_n,\nu) \to \mathrm{OT}_p(\mu,\nu).$$

Lemma 4.5 gives the desired conclusion. □

Proposition 4.4 allows us to generalize existence of barycenters in the finite case (Section 4.1) to general probability distributions supported on $\mathcal{M}^p$.

Theorem 4.6. *For any probability distribution P supported on $\mathcal{M}^p$ with finite p-th moment, the set of Fréchet means of P is a non-empty compact convex set of $\mathcal{M}^p$.*

Note that here, *convexity* must be understood in the sense of linear combination of measures, that is a convex set $C \subset \mathcal{M}^p$ is such that for any two measures $\mu, \nu \in C$, and any $t \in (0,1)$, one has $(1-t)\mu + t\nu \in C$.

Proof. We first prove the non-emptiness. Let $P = \sum_{i=1}^N \lambda_i \mu_i$ be a probability measure on $\mathcal{M}^p$ with finite support $\mu_1, \ldots, \mu_N$ (but with no assumption on finiteness of masses). According to Proposition 3.10, there exists sequences $(\mu_i^{(n)})_n$ in $\mathcal{M}_f^p$ with $\mathrm{OT}_p(\mu_i^{(n)}, \mu_i) \to 0$. As a consequence of the

result of Section 4.1, the probability measures $P^{(n)} := \sum_i \lambda_i \delta_{\mu_i^{(n)}}$ admit Fréchet means. Furthermore, $W_p^p(P^{(n)}, P) \leq \sum_i \lambda_i \mathrm{OT}_p^p(\mu_i^{(n)}, \mu_i)$ so that this quantity converges to 0 as $n \to \infty$. It follows from Proposition 4.4 that P admits a Fréchet mean.

If P has infinite support, following [LGL16], it can be approximated (in W_p) by a empirical probability measure $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{\mu_i}$ where the μ_i are i.i.d. from P. We know that P_n admits a Fréchet mean since its support is finite, and thus, applying Proposition 4.4 once again, we obtain that P admits a Fréchet mean.

Finally, the compactness of the set of Fréchet means follows from Proposition 4.4 and the convexity of the set from the convexity of OT_p^p (Lemma 3.9). □

4.3 Fréchet means in $\mathcal{D}^p$

We now prove the existence of Fréchet means for distributions of persistence diagram (i.e. probability distributions supported on $\mathcal{D}^p$), extending the results of [MMH11], in which authors prove their existence for specific probability distributions (namely distributions with compact support or specific rates of decay). Theorem 4.8 below asserts two different things: that $\arg\min\{\mathcal{E}(a),\ a \in \mathcal{D}^p\}$ is non empty, and that

$$\min\{\mathcal{E}(a),\ a \in \mathcal{D}^p\} = \min\{\mathcal{E}(\mu),\ \mu \in \mathcal{M}^p\},$$

that is an optimal persistence measure cannot perform strictly better than an optimal persistence diagram when averaging persistence diagrams. As for Fréchet means in $\mathcal{M}^p$, we start with the finite case. The following lemma actually gives a geometric description of the set of Fréchet means obtained when averaging a finite number of finite diagrams.

Lemma 4.7. *Consider $a_1, \ldots, a_N \in \mathcal{D}_f$, weights $(\lambda_i)_i$ that sum to 1, and let $P := \sum_{i=1}^N \lambda_i \delta_{a_i}$. Then, the set of minimizers of $\mu \mapsto \sum_{i=1}^N \lambda_i \mathrm{OT}_p^p(\mu, a_i)$ is a non empty convex subset of $\mathcal{M}_f$ whose extreme points belong to $\mathcal{D}_f$. In particular, P admits a Fréchet mean in $\mathcal{D}_f$.*

The proof of this lemma relies on *integer programming* results and is found in Section 4.4. Note that, as a straightforward consequence, if P has a unique minimizer in $\mathcal{D}_f$ (which is generically true [Tur13]), then so it does in $\mathcal{M}_f$ (and this unique minimizer is a persistence diagram).

Theorem 4.8. *For any probability distribution P supported on $\mathcal{D}^p$ with finite p-th moment, the set of Fréchet means of P contains an element of $\mathcal{D}^p$. Furthermore, if P is supported on a finite set of finite persistence diagrams, then the set of the Fréchet means of P is a convex set whose extreme points are in $\mathcal{D}^p$.*

Proof. The second assertion of the theorem is stated in Lemma 4.7. To prove the existence of a Fréchet mean which is a persistence diagram, we argue as in the proof of Theorem 4.6, using additionally the fact that $\mathcal{D}^p$ is closed in $\mathcal{M}^p$ (Corollary 3.6). □

4.4 Proofs

Proof of Lemma 4.1. Let $\mu \in \mathcal{M}^p$. Let $\pi_i \in \mathrm{Opt}_p(\mu_i, \mu)$ for $i = 1, \ldots, N$. The measure $A \subset \Omega \mapsto \pi_i(\partial\Omega \times A)$ is absolutely continuous with respect to μ. Therefore, it has a density f_i with respect to μ. Define for $A \subset \Omega$ a Borel set,

$$\mu'(A) := \mu(A) - \int_A \min_j f_j(x) \mathrm{d}\mu(x),$$

and, for $i = 1, \ldots, N$, a measure π_i', equal to π_i on $\Omega \times \overline{\Omega}$ and which satisfies for $A \subset \Omega$ a Borel set,

$$\pi_i'(\partial\Omega \times A) = \pi_i'(\mathbf{p}_{\partial\Omega}(A) \times A) := \pi_i(\partial\Omega \times A) - \int_A \min_j f_j(x) \mathrm{d}\mu(x),$$

where s is the orthogonal projection on $\partial\Omega$. As $\pi_i(\partial\Omega \times A) = \int_A f_i(x) \mathrm{d}\mu(x)$, $\pi_i'(A)$ is non-negative, and as $\pi_i(\partial\Omega \times A) \leq \mu(A)$, it follows that $\mu'(A)$ is non-negative. To prove that $\pi_i' \in \mathrm{Adm}(\mu_i, \mu')$, it is enough to check that for $A \subset \Omega$, $\pi_i'(\overline{\Omega} \times A) = \mu'(A)$:

$$\begin{aligned}\pi_i'(\overline{\Omega} \times A) &= \pi_i(\Omega \times A) + \pi_i(\partial\Omega \times A) - \int_A \min_j f_j(x) \mathrm{d}\mu(x) \\ &= \mu(A) - \int_A \min_j f_j(x) \mathrm{d}\mu(x) = \mu'(A).\end{aligned}$$

Also,

$$\begin{aligned}\mu'(\Omega) &= \int_\Omega (1 - \min_j f_j) \mathrm{d}\mu(x) \leq \sum_{j=1}^N \int_\Omega (1 - f_j) \mathrm{d}\mu(x) \\ &= \sum_{j=1}^N (\mu(\Omega) - \pi_j(\partial\Omega \times \Omega)) = \sum_{j=1}^N (\pi_j(\overline{\Omega} \times \Omega) - \pi_j(\partial\Omega \times \Omega)) \\ &= \sum_{j=1}^N \pi_j(\Omega \times \Omega) \leq \sum_{j=1}^N \pi_j(\Omega \times \overline{\Omega}) = \sum_{j=1}^N m_j = m_{\mathrm{tot}}.\end{aligned}$$

and thus $\mu'(\Omega) \leq m_{\mathrm{tot}}$. To conclude, observe that

$$\begin{aligned}\mathcal{E}(\mu') &\leq \sum_{i=1}^N \lambda_i C_p(\pi_i') = \sum_{i=1}^N \lambda_i \left(\iint_{\Omega \times \overline{\Omega}} d(x,y)^p \mathrm{d}\pi_i(x,y) \right. \\ &\qquad \left. + \iint_{\partial\Omega \times \Omega} d(x,y)^p \mathrm{d}\pi_i(x,y) - \int_\Omega d(x, \partial\Omega)^p \min_j f_j(x) \mathrm{d}\mu(x) \right) \\ &\leq \sum_{i=1}^N \lambda_i C_p(\pi) = \mathcal{E}(\mu).\end{aligned}$$

□

Proof of Proposition 4.2. Let G be the set of $\mu \in \mathcal{M}^p$ such that, for all i, there exists $\pi_i \in \mathrm{Opt}_p(\mu_i, \mu)$ with $\pi_i(\Omega, \partial\Omega) = 0$. By point 2 of Lemma 3.12, for $\mu \in G$ and $\pi_i \in \mathrm{Opt}_p(\mu_i, \mu)$ with $\pi_i(\Omega, \partial\Omega) = 0$, $\iota(\pi_i)$ is well defined and satisfies

$$\mathrm{OT}_p^p(\mu_i, \mu) = C_p(\pi_i) = \iint_{\tilde{\Omega} \times \tilde{\Omega}} d(x,y)^p \mathrm{d}\iota(\pi_i)(x,y) \geq \tilde{C}_p(\iota(\pi_i)) \geq W_{p,\rho}^p(\tilde{\mu}_i, \tilde{\mu}),$$

so that $\mathcal{F}(\Psi(\mu)) \leq \mathcal{E}(\mu)$. As, by Lemma 3.13, $\mathcal{E} \leq \mathcal{F} \circ \Psi$, we therefore have $\mathcal{E}(\mu) = \mathcal{F}(\Psi(\mu))$ for $\mu \in G$.

We now show that if $\mu \notin G$, then there exists $\mu' \in \mathcal{M}^p$ with $\mathcal{E}(\mu') < \mathcal{E}(\mu)$. Let $\mu \notin G$ and $\pi_i \in \mathrm{Opt}_p(\mu_i, \mu)$. Assume that for some i, we have $\pi_i(\Omega, \partial\Omega) > 0$, and introduce $\nu \in \mathcal{M}^p$ defined as $\nu(A) = \pi_i(A, \partial\Omega)$ for $A \subset \Omega$. Define

$$T : \Omega \ni x \mapsto \arg\min_{y \in \Omega} \left\{ \lambda_i d(x,y)^p + \sum_{j \neq i} \lambda_j d(y, \partial\Omega)^p \right\} \in \Omega.$$

Note that since $p > 1$, this map is well defined (the minimizer is unique due to strict convexity) and continuous thus measurable. Consider the measure

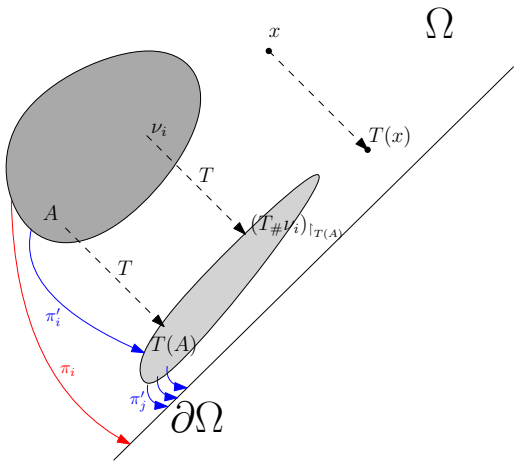


Figure 4.1: Global picture of the proof. The main idea is to observe that the cost induced by π_i (red) is strictly larger than the sum of costs induces by the π_i's (blue), which leads to a strictly better energy.

$\mu' = \mu + (T_\#\nu)$, where $T_\#\nu$ is the push-forward of ν by the application T. Consider the transport plan π_i' deduced from π_i where ν is transported onto $T_\#\nu$ instead of being transported to $\partial\Omega$ (see Figure 4.1). More precisely, π_i' is the measure on $\overline{\Omega} \times \overline{\Omega}$ defined by, for Borel sets $A, B \subset \Omega$:

$$\pi_i'(A \times B) = \pi_i(A \times B) + \nu(A \cap T^{-1}(B)),$$
$$\pi_i'(A \times \partial\Omega) = 0, \quad \pi_i'(\partial\Omega \times B) = \pi_i(\partial\Omega \times B).$$

We have $\pi_i' \in \mathrm{Adm}(\mu_i, \mu')$. Indeed, for Borel sets $A, B \subset \Omega$:

$$\pi_i'(A \times \overline{\Omega}) = \pi_i'(A \times \Omega) = \pi_i(A \times \Omega) + \nu(A) = \pi_i(A \times \overline{\Omega}) = \mu_i(A),$$

and

$$\begin{aligned}\pi_i'(\overline{\Omega} \times B) &= \pi_i'(\Omega \times B) + \pi_i'(\partial\Omega \times B) \\ &= \pi_i(\Omega \times B) + \nu(T^{-1}(B)) + \pi_i(\partial\Omega \times B) \\ &= \mu(B) + T_\#\nu(B) = \mu'(B).\end{aligned}$$

Using π_i' instead of π_i changes the transport cost by the quantity

$$\int_\Omega [d(x, T(x))^p - d(x, \partial\Omega)^p] \mathrm{d}\nu(x) < 0.$$

In a similar way, we define for $j \neq i$ the plan $\pi_j' \in \mathrm{Adm}(\mu_j, \mu')$ by transporting the mass induced by the newly added $(T_\#\nu)$ to the diagonal

$\partial\Omega$. Using these modified transport plans increases the total cost by

$$\sum_{j\neq i} \lambda_j \int_\Omega d(T(x), \partial\Omega)^p \mathrm{d}\nu(x).$$

One can observe that

$$\int_\Omega \left[\lambda_i \left(d(x, T(x))^p - d(x, \partial\Omega)^p \right) + \sum_{j\neq i} \lambda_j d(T(x), \partial\Omega)^p \right] \mathrm{d}\nu(x) < 0$$

due to the definition of T and $\nu(\Omega) > 0$.

Therefore, the total transport cost induced by the $(\pi'_i)_{i=1\dots N}$ is strictly smaller than $\mathcal{E}(\mu)$, and thus $\mathcal{E}(\mu') < \mathcal{E}(\mu)$. Finally, we have

$$\inf_{\mu\in\mathcal{M}^p_{\leq m_{tot}}} \mathcal{E}(\mu) = \inf_{\mu\in G} \mathcal{E}(\mu) = \inf_{\mu\in G} \mathcal{F}(\Psi(\mu)) \geq \inf_{\mu\in\mathcal{M}^p_{\leq m_{tot}}} \mathcal{F}(\Psi(\mu)) \geq \inf_{\mu\in\mathcal{M}^p_{\leq m_{tot}}} \mathcal{E}(\mu),$$

where the last inequality comes from $\mathcal{F}\circ\Psi \geq \mathcal{E}$ (Lemma 3.13). Therefore, $\inf\mathcal{E} = \inf\mathcal{F}\circ\Psi$, which is equal to $\inf\mathcal{F}$, as Ψ is a bijection. Also, if μ is a minimizer of $\mathcal{E}$ (should it exist), then $\mu \in G$ and $\mathcal{E}(\mu) = \mathcal{F}(\Psi(\mu))$. Therefore, as the infimum are equal, $\Psi(\mu)$ is a minimizer of $\mathcal{F}$. Reciprocally, if $\tilde{\mu}$ is a minimizer of $\mathcal{F}$, then, by Lemma 3.13, $\mathcal{F}(\tilde{\mu}) \geq \mathcal{E}(\Psi^{-1}(\tilde{\mu}))$, and, as the infimum are equal, $\Psi^{-1}(\tilde{\mu})$ is a minimizer of $\mathcal{E}$. □

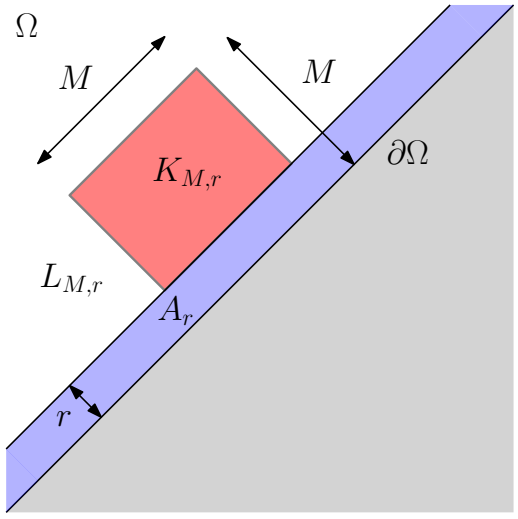


Figure 4.2: Partition of Ω used in the proof of Lemma 4.5.

Proof of Lemma 4.5. For the direct implication, take $\nu = \mathbf{0}$ and apply Theorem 3.5.

Let us prove the converse implication. Assume that $\mu_n \xrightarrow{v} \mu$ and $\mathrm{OT}_p(\mu_n, \nu) \to \mathrm{OT}_p(\mu, \nu)$ for some $\nu \in \mathcal{D}^p$. The vague convergence of $(\mu_n)_n$ implies that μ^p is the only possible accumulation point for weak convergence of the sequence $(\mu_n^p)_n$. Therefore, it is sufficient to show that the sequence $(\mu_n^p)_n$ is relatively compact for weak convergence (i.e. tight and bounded in total variation, see Proposition B.2). Indeed, this would mean that (μ_n^p) converges weakly to μ^p, or equivalently by Proposition B.3 that $\mu_n \xrightarrow{v} \mu$ and $\mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu)$. The conclusion is then obtained thanks to Theorem 3.5.

Thus, let $(\mu_n)_n$ be any subsequence and $(\pi_n)_n$ be corresponding optimal transport plans between μ_n and ν. The vague convergence of $(\mu_n)_n$ implies that $(\pi_n)_n$ is relatively compact with respect to the vague convergence on E_Ω. Let π be a limit of any converging subsequence of $(\pi_n)_n$, which indexes are still denoted by n. One can prove that $\pi \in \mathrm{Opt}(\mu, \nu)$ (see [FG10, Prop. 2.3]. For $r > 0$, define $A_r := \{x \in \Omega,\ d(x, \partial\Omega) \leq r\}$ and write $\overline{A}_r$ for $A_r \cup \partial\Omega$. Consider $\eta > 1$. We can write

$$
\begin{aligned}
&\int_{A_r} d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) = \iint_{A_r \times \overline{\Omega}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \\
&= \iint_{A_r \times (\Omega \setminus A_{\eta r})} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) + \iint_{\overline{A}_r \times \overline{A}_{\eta r}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \\
&\overset{(*)}{\leq} \frac{1}{(\eta-1)^p} \iint_{A_r \times (\Omega \setminus A_{\eta r})} d(x,y)^p \mathrm{d}\pi_n(x,y) + \iint_{\overline{A}_r \times \overline{A}_{\eta r}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \\
&\leq \frac{1}{(\eta-1)^p} \mathrm{OT}_p^p(\mu_n,\nu) \\
&\quad + 2^{p-1} \left(\iint_{\overline{A}_r \times \overline{A}_{\eta r}} d(x,y)^p \mathrm{d}\pi_n(x,y) + \iint_{\overline{A}_r \times \overline{A}_{\eta r}} d(y,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \right) \\
&\leq \frac{1}{(\eta-1)^p} \mathrm{OT}_p^p(\mu_n,\nu) \\
&\quad + 2^{p-1} \left(\mathrm{OT}_p^p(\mu_n,\nu) - \iint_{E_\Omega \setminus (\overline{A}_r \times \overline{A}_{\eta r})} d(x,y)^p \mathrm{d}\pi_n(x,y) + \int_{A_{\eta r}} d(y,\partial\Omega)^p \mathrm{d}\nu(y) \right)
\end{aligned}
$$

where $(*)$ holds because $d(x,y) \geq (\eta-1)r \geq (\eta-1)d(x,\partial\Omega)$ for $(x,y) \in$

$A_r \times A_{\eta r}^c$. Therefore,

$$\limsup_{n\to\infty} \int_{A_r} d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) \leq \frac{1}{(\eta-1)^p} \mathrm{OT}_p^p(\mu,\nu) + 2^{p-1}\Bigg(\mathrm{OT}_p^p(\mu,\nu) - \iint\limits_{E_\Omega\backslash(\overline{A}_r\times\overline{A}_{\eta r})} d(x,y)^p \mathrm{d}\pi(x,y) + \int_{A_{\eta r}} d(y,\partial\Omega)^p \mathrm{d}\nu(y)\Bigg)$$

Note that at the last line, we used Portmanteau theorem (see Proposition B.4) on the sequence of measures $(d(x,y)^p \mathrm{d}\pi_n(x,y))_n$ for the open set $E_\Omega\backslash(\overline{A}_r \times \overline{A}_{\eta r})$. Letting $r \to 0$, then η goes to infinity, one obtains

$$\lim_{r\to 0} \limsup_{n\to\infty} \int_{A_r} d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) = 0.$$

The second part consists in showing that there can not be mass escaping "at infinity" in the subsequence $(\mu_n^p)_n$. Fix $r, M > 0$. For $x \in \Omega$, denote $\mathbf{p}_{\partial\Omega}(x)$ the projection of x on $\partial\Omega$. Pose

$$K_{M,r} := \{x \in \Omega\backslash A_r,\ d(x,\partial\Omega) < M, d(\mathbf{p}_{\partial\Omega}(x),0) < M\}$$

and $L_{M,r}$ the closure of $\Omega\backslash(A_r \cup K_{M,r})$ (see Figure 4.2). For $r' > 0$,

$$\begin{aligned}
\int_{L_{M,r}} d(x,\partial\Omega)^p \mathrm{d}\mu_n(x) &= \iint\limits_{L_{M,r}\times\overline{\Omega}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \\
&= \iint\limits_{L_{M,r}\times(L_{M/2,r'}\cup\overline{A}_{r'})} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) + \iint\limits_{L_{M,r}\times K_{M/2,r'}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y) \\
&\leq 2^{p-1} \iint\limits_{L_{M,r}\times(L_{M/2,r'}\cup\overline{A}_{r'})} d(x,y)^p \mathrm{d}\pi_n(x,y) \\
&\quad + 2^{p-1} \iint\limits_{L_{M,r}\times(L_{M/2,r'}\cup\overline{A}_{r'})} d(\partial\Omega,y)^p \mathrm{d}\pi_n(x,y) \\
&\quad + \iint\limits_{L_{M,r}\times K_{M/2,r'}} d(x,\partial\Omega)^p \mathrm{d}\pi_n(x,y).
\end{aligned}$$

We treat the three parts of the sum separately. As before, taking the lim sup in n and letting $M \to \infty$, the first part of the sum converges to 0 (apply

Portmanteau theorem on the open set $E_\Omega \backslash (L_{M,r} \times (L_{M/2,r'} \cup \overline{A}_{r'}))$. The second part is smaller than

$$2^{p-1} \int_{L_{M/2,r'} \cup A_{r'}} d(y, \partial\Omega)^p \mathrm{d}\nu(y),$$

which converges to 0 as $M \to \infty$ and $r' \to 0$. For the third part, notice that if $(x, y) \in L_{M,r} \times K_{M/2,r'}$, then

$$d(x, \partial\Omega) \leq d(x, \mathbf{p}_{\partial\Omega}(y)) \leq d(x, y) + d(y, \mathbf{p}_{\partial\Omega}(y)) \leq d(x, y) + \frac{M}{2} \leq 2d(x, y).$$

Therefore,

$$\begin{aligned} \iint\limits_{L_{M,r} \times K_{M/2,r'}} d(x, \partial\Omega)^p \mathrm{d}\pi_n(x, y) &\leq 2^p \iint\limits_{L_{M,r} \times K_{M/2,r'}} d(x, y)^p \mathrm{d}\pi_n(x, y) \\ &\leq 2^p \iint\limits_{L_{M,r} \times \overline{\Omega}} d(x, y)^p \mathrm{d}\pi_n(x, y). \end{aligned}$$

As before, it is shown that $\limsup_n \iint_{L_{M,r} \times \overline{\Omega}} d(x, y)^p \mathrm{d}\pi_n(x, y)$ converges to 0 when M goes to infinity by applying Portmanteau theorem on the open set $E_\Omega \backslash (L_{M,r} \times \overline{\Omega})$.

Finally, we have shown, that by taking r small enough and M large enough, one can find a compact set $\overline{K_{M,r}}$ such that $\int_{\Omega \backslash \overline{K_{M,r}}} d(x, \partial\Omega)^p \mathrm{d}\mu_n = \mu_n^p(\Omega \backslash \overline{K_{M,r}})$ is uniformly small: $(\mu_n^p)_n$ is tight. As we have

$$\mu_n^p(\Omega) = \mathrm{Pers}_p(\mu_n) = \mathrm{OT}_p^p(\mu_n, 0) \leq (\mathrm{OT}_p(\mu_n, \nu) + \mathrm{OT}_p(\nu, 0))^p \to (\mathrm{OT}_p(\mu, \nu) + \mathrm{OT}_p(\nu, 0))^p,$$

it is also bounded in total variation. Hence, $(\mu_n^p)_n$ is relatively compact for the weak convergence: this concludes the proof. □

Proof of Lemma 4.7. Let $P = \sum_{i=1}^N \lambda_i \delta_{a_i}$ be a probability distribution with $a_i \in \mathcal{D}_f$ of mass $m_i \in \mathbb{N}$, and define $m_{\mathrm{tot}} = \sum_{i=1}^N m_i$. By Proposition 4.2, every Fréchet mean a of P is in correspondence with a Fréchet mean for the Wasserstein distance $\tilde{a}$ of $\tilde{P} = \sum_{i=1}^N \lambda_i \delta_{\tilde{a}_i}$, where $\tilde{a}_i = a_i + (m_{\mathrm{tot}} - m_i)\delta_{\partial\Omega}$, with a being the restriction of $\tilde{a}$ to Ω.

Fix $m \in \mathbb{N}$, and let $\tilde{a}_1, \ldots, \tilde{a}_N$ be point measures of mass m in $\tilde{\Omega}$. Write $\tilde{a}_i = \sum_{j=1}^m \delta_{x_{i,j}}$, so that $x_{i,j} \in \tilde{\Omega}$ for $1 \leq i \leq N,\ 1 \leq j \leq m$, with the $x_{i,j}$s non-necessarily distinct. Define

$$T : (x_1, \ldots, x_N) \in \tilde{\Omega}^N \mapsto \arg\min \left\{ \sum_{i=1}^N \lambda_i \rho(x_i, y)^p,\ y \in \tilde{\Omega} \right\} \in \tilde{\Omega}. \tag{4.1}$$

Since we assume $p > 1$, T is well-defined and is continuous (the minimizer is unique by strict convexity). Using the localization property stated in [COO15a, Section 2.2], we know that the support of a Fréchet mean of $\tilde{P}$ is included in the finite set

$$S := \{T(x_{1,j_1}, \dots, x_{N,j_N}),\ 1 \leq j_1, \dots, j_N \leq m\}.$$

Let $K = m^N$ and let $z_1, \dots, z_K$ be an enumeration of the points of S (with potential repetitions). Denote by $\mathrm{Gr}(z_k)$ the N elements $x_1, \dots, x_N$, with $x_i \in \mathrm{spt}(\tilde{a}_i)$, such that $z_k = T(x_1, \dots, x_N)$. It is explained in [COO15a, §2.3], that finding a Fréchet mean of $\tilde{P}$ is equivalent to finding a minimizer of the problem

$$\inf_{(\gamma_1,\dots,\gamma_N)\in\Pi} \sum_{i=1}^{N} \lambda_i \iint_{\tilde{\Omega}^2} \rho(x_i, y)^p \mathrm{d}\gamma_i(x_i, y), \tag{4.2}$$

where Π is the set of plans $(\gamma_i)_{i=1,\dots,N}$, with γ_i having for first marginal $\tilde{a}_i$, and such that all γ_is share the same (non-fixed) second marginal. Furthermore, we can assume without loss of generality that $(\gamma_1 \dots \gamma_N)$ is supported on $(\mathrm{Gr}(z_k), z_k)_k$, i.e. a point z_k in the Fréchet mean is necessary transported to its corresponding grouping $\mathrm{Gr}(z_k)$ by (optimal) $\gamma_1, \dots \gamma_N$ [COO15a, §2.3]. For such a minimizer, the common second marginal is a Fréchet mean of $\tilde{P}$.

A potential minimizer of (4.2) is described by a vector $\gamma = (\gamma_{i,j,k}) \in \mathbb{R}_+^{NmK}$ such that:

$$\begin{cases} \text{for } 1 \leq i \leq N,\ 1 \leq j \leq m, & \sum_{k=1}^{K} \gamma_{i,j,k} = 1 \text{ and} \\ \text{for } 2 \leq i \leq N,\ 1 \leq k \leq K, & \sum_{j=1}^{m} \gamma_{1,j,k} = \sum_{j=1}^{m} \gamma_{i,j,k}. \end{cases} \tag{4.3}$$

Let $c \in \mathbb{R}^{NmK}$ be the vector defined by $c_{i,j,k} = \mathbf{1}\{x_{i,j} \in \mathrm{Gr}(z_k)\}\lambda_i \rho(x_{i,j}, z_k)^p$. Then, the problem (4.2) is equivalent to

$$\underset{\gamma\in R_+^{NmK}}{\text{minimize}}\, \gamma^T c \quad \text{under the constraints (4.3).} \tag{4.4}$$

The set of Fréchet means of P are in bijection with the set of minimizers of this Linear Programming problem (see [Sch03, §5.15]), which is given by a face of the polyhedron described by the equations (4.3). Hence, if we show that this polyhedron is integer (i.e. its vertices have integer values), then it would imply that the extreme points of the set of Fréchet means of P are point measures, concluding the proof. The constraints (4.3) are described by a matrix A of size $(Nm + (N-1)K) \times NmK$ and a vector $b = [\mathbf{1}_{Nm}, \mathbf{0}_{(N-1)K}]$, such that $\gamma \in \mathbb{R}^{NmK}$ satisfies (4.3) if and only if $A\gamma = b$.

A sufficient condition for the polyhedron $\{Ax \leq b\}$ to be integer is to satisfy the following property (see [Sch03, Section 5.17]): for all $u \in \mathbb{Z}^{NmK}$, the dual problem

$$\max\{y^T b,\ y \geq 0 \text{ and } y^T A = u\} \tag{4.5}$$

has either no solution (i.e. there is no $y \geq 0$ satisfying $y^T A = u$), or it has an integer optimal solution y.

For y satisfying $y^T A = u$, write $y = [y^0, y^1]$ with $y^0 \in \mathbb{R}^{Nm}$ and $y^1 \in \mathbb{R}^{(N-1)K}$, so that y^0 is indexed on $1 \leq i \leq N,\ 1 \leq j \leq m$ and y^1 is indexed on $2 \leq i \leq N,\ 1 \leq k \leq K$. One can check that, for $2 \leq i \leq N,\ 1 \leq j \leq m,\ 1 \leq k \leq K$:

$$u_{1,j,k} = y^0_{1,j} + \sum_{i'=2}^{N} y^1_{i',k} \quad \text{and} \quad u_{i,j,k} = y^0_{i,j} - y^1_{i,k}, \tag{4.6}$$

so that,

$$\begin{aligned} y^T b &= \sum_{i=1}^{N}\sum_{j=1}^{m} y^0_{i,j} = \sum_{j=1}^{m} y^0_{1,j} + \sum_{i=2}^{N}\sum_{j=1}^{m} y^0_{i,j} \\ &= \sum_{j=1}^{m}(u_{1,j,k} - \sum_{i=2}^{N} y^1_{i,k}) + \sum_{i=2}^{N}\sum_{j=1}^{m}(u_{i,j,k} + y^1_{i,k}) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{m} u_{i,j,k}. \end{aligned}$$

Therefore, the function $y^T b$ is constant on the set $P := \{y \geq 0,\ y^T A = u\}$, and any point of the set is an argmax. We need to check that if the set P is non-empty, then it contains a vector with integer coordinates: this would conclude the proof. A solution of the homogeneous equation $y^T A = 0$ satisfies $y^0_{i,j} = y^1_{i,k} = \lambda_i$ for $i \geq 2$ and $y^0_{1,j} = -\sum_{i=2}^{N} y^1_{i,k} = -\sum_{i=2}^{N} \lambda_i$ and reciprocally, any choice of $\lambda_i \in \mathbb{R}$ gives rise to a solution of the homogeneous equation. For a given u, one can verify that the set of solutions of $y^T A = u$ is given, for $\lambda_i \in \mathbb{R}$, by

$$\begin{cases} y^0_{1,j} = \sum_{i=1}^{N} u_{i,j,k} - \sum_{i=2}^{N} \lambda_i \\ y^0_{i,j} = \lambda_i \text{ for } i \geq 2, \\ y^1_{i,k} = -u_{i,j,k} + \lambda_i \text{ for } i \geq 2. \end{cases}$$

Such a solution exists if and only if for all j, $U_j := \sum_{i=1}^{N} u_{i,j,k}$ does not depend on k and for $i \geq 2$, $U_{i,k} := u_{i,j,k}$ does not depend on j. For such a

vector u, P corresponds to the $\lambda_i \geq 0$ with $\lambda_i \geq \max_k U_{i,k}$ and $U_j \geq \sum_{i=1}^{N} \lambda_i$. If this set is non empty, it contains as least the point corresponding to $\lambda_i = \max\{0, \max_k U_{i,k}\}$, which is an integer: this point is integer valued, concluding the proof. □

Part II

Applications in statistics and learning

Chapter 5

Fast estimation of Fréchet means for persistence diagrams

Abstract

In this chapter, we address the problem of estimating Fréchet means (a.k.a. barycenters) in the space of persistence diagrams from a computational perspective. A theoretical study of Fréchet means of persistence diagrams and measures can be found in Chapter 4, where existence results are provided. Here, we propose an approach to approximate Fréchet means of persistence diagrams in a convex, parallelizable, and GPU-friendly way. This work is essentially taken from [LCO18], although being significantly more detailed and revisited under the light of Chapter 3.

Implementation resources.

- Implementation in Gudhi of Wasserstein distances and Lagrangian barycenters for persistence diagrams.
- (To appear in `Gudhi`) Eulerian barycenters using entropic regularization.
- A tutorial in Gudhi for barycenters of persistence diagrams.

5.1 Preliminary remarks and problem formulation

Let Ω denote the open half-plane

$$\Omega := \{(b, d) \in \mathbb{R}^2,\ b > d\}, \tag{5.1}$$

and let $\partial\Omega$ denote its boundary, namely the diagonal

$$\partial\Omega := \{(t, t),\ t \in \mathbb{R}\}. \tag{5.2}$$

Eventually, let $\overline{\Omega}$ denote $\Omega \sqcup \partial\Omega$.

For now, we fix a parameter $p \in [1, +\infty)$. Recall from Definition 2.4 in Section 2.1 that persistence diagrams are defined as locally finite point measures supported on Ω, that is of the form $\sum_{x\in X} n_x\delta_x$, where $X \subset \Omega$ is locally finite, and $n_x \in \mathbb{N}$ for all $x \in X$. Recall that persistence diagrams can be equipped with metrics $\mathbf{d}_p$, defining a metric space $(\mathcal{D}^p, \mathbf{d}_p)$, and which depends on a ground metric d which is set to be $d(x, y) := \|x - y\|_p$ in this chapter. The space of persistence diagrams is denoted by $\mathcal{D}$. It is a subset of the space of non-negative Radon measures supported on Ω, denoted by $\mathcal{M}(\Omega)$ or $\mathcal{M}$ if there is no ambiguity.

Here, as we target numerical applications, we will only consider finite sets of finite persistence diagrams. Therefore, we can assume that our diagrams are supported on $[0, 1]^2 \cap \Omega$.

Problem formulation In this context, we consider the following problem. Given a set of (finite) persistence diagrams $\nu_1, \dots \nu_N \in \mathcal{D}^p$, and non-negative weights $(\lambda_i)_{1\leq i\leq N}$ that sum to one, we are looking for minimizers of

$$\mathcal{E} : \mu \mapsto \sum_{i=1}^{N} \lambda_i \mathbf{d}_p(\mu, \nu_i)^p. \tag{5.3}$$

This problem was initially introduced in [TMMH14], we refer the reader to Section 2.1.3 for more details. This problem is known to have (non-unique, in general) solutions in both $\mathcal{D}$ and $\mathcal{M}$ (and the minimum values are the same). However, these exact solutions are generally intractable. The goal is thus to provide a way to estimate such minimizers in practice, ideally in a large-scale manner.

Diagram metrics as an optimal transport problems (reminder). In this context, following the results of Chapter 3 (namely, Proposition 3.3

and Proposition 3.11), the metric $\mathbf{d}_p$ between two persistence diagrams $\mu = \sum_{i=1}^{n_1} \delta_{x_i}$ and $\nu = \sum_{j=1}^{n_2} \delta_{y_j}$ can be reformulated as a *balanced* optimal transport problem (see Section 2.2) on the space $\tilde{\Omega}$ obtained from $\overline{\Omega}$ by identifying all the points in $\partial\Omega$ to a single point (still denoted by $\partial\Omega$). First, note that d naturally induces a ground cost on $\tilde{\Omega} \times \tilde{\Omega}$ (still denoted by d) by setting $d(x, \{\partial\Omega\}) := d(x, \mathbf{p}_{\partial\Omega}(x))$, where $\mathbf{p}_{\partial\Omega}(x)$ is the orthogonal projection of $x \in \Omega$ onto the diagonal $\partial\Omega$.

Introduce the operator

$$\begin{aligned} \mathbf{R} : \mathcal{M}(\Omega) &\to \mathcal{M}(\{\partial\Omega\}) \\ \mu &\mapsto \mu(\Omega)\delta_{\partial\Omega}, \end{aligned} \tag{5.4}$$

and let $\tilde{\mu} = \mu + \mathbf{R}\nu$ and $\tilde{\nu} = \nu + \mathbf{R}\mu$. Let $a = (\mathbf{1}_{n_1}, n_2) \in \mathbb{R}^{n_1+1}$ and $b = (\mathbf{1}_{n_2}, n_1) \in \mathbb{R}^{n_2+1}$ be the weight vectors of μ and ν respectively, where $\mathbf{1}_k = (1 \dots 1) \in \mathbb{R}^k$. Let C be the cost matrix with block structure

$$C = \begin{pmatrix} \hat{C} & u \\ v^T & 0 \end{pmatrix} \in \mathbb{R}^{(n_1+1)\times(n_2+1)}, \tag{5.5}$$

with $u_i = d(x_i, \partial\Omega)^p$, $v_j = d(\partial\Omega, y_j)^p$ and $\hat{C} = (d(x_i, y_j)^p)_{ij}$.

It reads

$$\mathbf{d}_p(\mu, \nu)^p = \mathbf{L}_C(a, b) \tag{5.6}$$

where $\mathbf{L}_C$ is defined by the following optimization problems, dual of each other:

$$\mathbf{L}_C(a, b) = \min_{P \in \Pi(a,b)} \langle C, P \rangle = \min_{P \in \Pi(a,b)} \sum_{ij} C_{ij} P_{ij}, \tag{5.7}$$

$$= \max_{f,g \in \Psi_C} \langle a, f \rangle + \langle b, g \rangle = \max_{f,g \in \Psi_C} \sum_i a_i f_i + \sum_j b_j g_j, \tag{5.8}$$

where the constraint sets are respectively defined as

$$\Pi(a, b) := \{P \in \mathbb{R}^{(n_1+1)\times(n_2+1)},\ P\mathbf{1}_{n_2+1} = a,\ P^T\mathbf{1}_{n_1+1} = b\}, \tag{5.9}$$

$$\Psi_C := \{(f, g) \in \mathbb{R}^{n_1+1} \times \mathbb{R}^{n_2+1},\ \forall i, j, f_i + g_j \leq C_{ij}\}. \tag{5.10}$$

Remark 5.1. Note that the cost d defined on $\tilde{\Omega} \times \tilde{\Omega}$ is not a metric as one could have $d(x, y) \geq d(x, \partial\Omega) + d(\partial\Omega, y)$. To recover a true metric space, we can define

$$\rho(x, y) := \min\{d(x, y), d(x, \partial\Omega) + d(\partial\Omega, y)\}.$$

It is proved (Proposition 3.11) that using either ρ or d leads to the same optimal transport plans and optimal transport costs, so actually these two costs are interchangeable. Formally, it yields with previous notations,

$$\mathbf{L}_C(a, b) = W^p_{p,\rho}(\hat{\mu}, \hat{\nu}).$$

From a computational perspective, using $d(x, y)^p = \|x - y\|_p^p$, which is a *separable cost*, is highly beneficial when implementing the *entropic regularization* of diagram distances.

5.2 A Lagrangian approach

The first algorithm proposed to address the problem (5.3), assuming $p = 2$, was also introduced in [TMMH14]. It relies on a Lagrangian approach and is summarized in Algorithm 1. As it performs barycenter estimation relying on the Munkres (Hungarian) algorithm, it is referred to as the *B-Munkres algorithm* in the following. This algorithm is proved to converged to a *local* minimum of $\mathcal{E}$ in $\mathcal{D}$, see Figure 5.1 for an illustrative example. In this section, we propose to revisit this algorithm by adopting an optimal transport perspective.

For $x_1 \dots x_m \in \tilde{\Omega}$, assume that $x_1 \dots x_k \in \Omega$ (i.e. are off-diagonal points) and $x_{k+1} \dots x_m = \partial\Omega$, let $\overline{x}$ denote the arithmetic mean of $x_1 \dots x_k$, and define

$$\text{mean}(x_1 \dots x_m) := \arg\min_x \sum_{i=1}^{n} d(x, x_i)^2 = \frac{k}{m}\overline{x} + \frac{m-k}{m}\mathbf{p}_{\partial\Omega}(\overline{x}).$$

Interestingly, under the light of (5.6), Algorithm 1 can formally be retrieved as a slight variation of algorithms introduced independently in the optimal transport literature at a similar time in [CD14]. To clarify this, we introduce the following proposition.

Proposition 5.1. *Let $\nu_1 \dots \nu_N$ be N persistence diagrams of respective masses (number of points) $n_1 \dots n_N$. Let $m_{\text{tot}} := \sum_i n_i$. Define $w_{n_i} := (\mathbf{1}_{n_i}, m_{\text{tot}} - n_i) \in \mathbb{R}^{n_i+1}$. Minimizing the energy*

$$\mathcal{E} : \mu \mapsto \sum_{i=1}^{N} \mathbf{d}_p^p(\mu, \nu_i)$$

is equivalent to minimizing

$$\tilde{\mathcal{E}} : (\tilde{\Omega})^{m_{\text{tot}}} \to \mathbb{R}$$
$$(y_1 \dots y_{m_{\text{tot}}}) \mapsto \mathbf{L}_{C_i}(w_k, w_{n_i})$$

Algorithm 1 Estimation of Fréchet means (B-Munkres) [TMMH14]

Input: Set of diagrams $\nu_1 \dots \nu_N$, initial guess $\{y_1 \dots y_k\}$.
Output: A persistence diagram $\mu = \sum_j \delta_{y_j}$ that is a local minima of $\mathcal{E}$ on $\mathcal{D}$ with $p = 2$.
Let `converged = False`.
while `not converged` **do**
  for $i = 1 \dots N$ **do**
    Compute γ_i optimal partial matching between μ and ν_i.
  end for
  for each $y_j \in \mu$ **do**
    Let $x_j^i \in \nu_i$ be such that $(y_j, x_j^i) \in \gamma_i$
    $y_j \leftarrow \text{mean}(x_j^i, 1 \leq i \leq N)$
  end for
  if None of the $(y_j)_j$ has changed **then**
    `converged` $\leftarrow$ `True`.
    return μ.
  end if
end while

where $k := |\{i,\ y_i \in \Omega\}|$ *(number of off-diagonal points) and* C_i *is the pairwise distance matrix defined in* (5.5) *(between the off-diagonal points of* $\mu := \sum_{i,\ y_i \in \Omega} \delta_{y_i}$ *and* ν_i*). Furthermore, if* $(y_1 \dots y_{m_{\text{tot}}})$ *is a minimizer of* $\tilde{\mathcal{E}}$*, then* μ *is a minimizer of* $\mathcal{E}$*, and vice-versa.*

Proof. In Chapter 4, we proved (Proposition 4.2 and Lemma 4.7) that minimizing $\mathcal{E}$ on $\mathcal{D}$ was equivalent to minimize

$$\begin{aligned} \mathcal{F} : \mathcal{D}_{m_{\text{tot}}}(\tilde{\Omega}) &\to \mathbb{R} \\ \tilde{\mu} &\mapsto \sum_{i=1}^{N} W_{p,\rho}^p(\tilde{\mu}, \nu_i + (m_{\text{tot}} - n_i)\delta_{\partial\Omega}), \end{aligned}$$

where $\mathcal{D}_{m_{\text{tot}}}(\tilde{\Omega})$ denotes the measures of the form $\tilde{\mu} = \sum_{j=1}^{m_{\text{tot}}} \delta_{y_j}$, where $y_j \in \tilde{\Omega}$, in sense that if $\tilde{\mu}$ minimizes $\mathcal{F}$, then $\mu := \sum_{i,\ y_i \in \Omega} \delta_{y_j}$ minimizes $\mathcal{E}$, and vice-versa.

Furthermore, in this context (see Remark 5.1),

$$W_{p,\rho}^p(\tilde{\mu}, \nu_i + (m_{\text{tot}} - n_i)\delta_{\partial\Omega}) = \mathbf{L}_{C_i}(w_k, w_{n_i}),$$

proving the claim. □

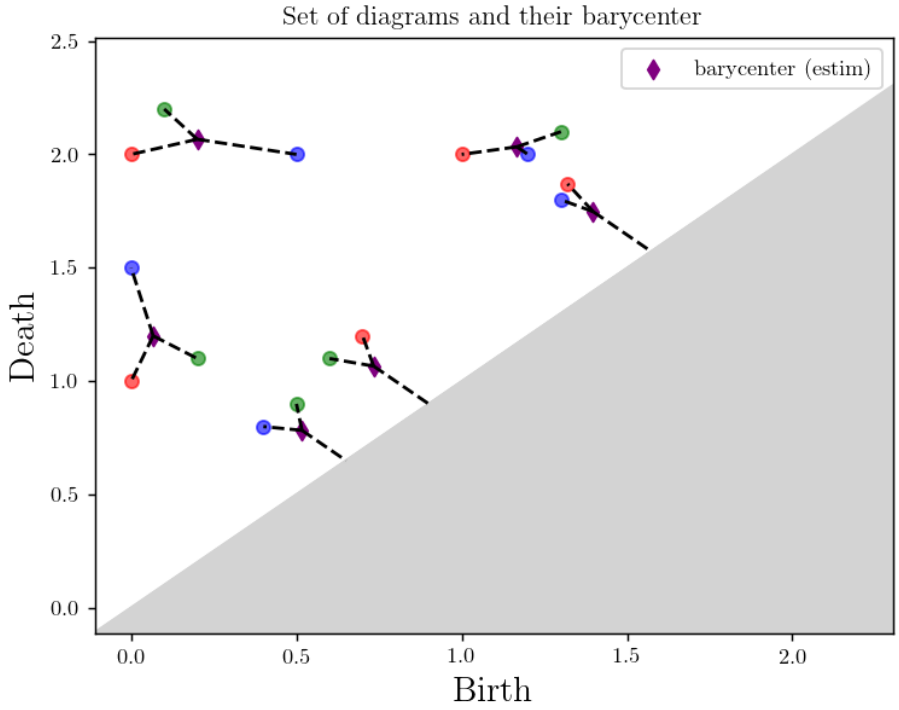


Figure 5.1: Example of output of Algorithm 1.

This problem is *exactly* a standard optimal-transport barycenter problem. A first remark that can be done from this crucial observation is that we immediately obtain that minimizers of $\tilde{\mathcal{E}}$ (thus of $\mathcal{E}$ once we restrict to off-diagonal points) must have their support included in

$$\{\text{mean}(x_1 \dots x_N),\ x_1 \in (a_1 \cup \{\partial\Omega\}) \dots x_N \in (a_N \cup \{\partial\Omega\})\},$$

a property known as the localization property [COO15a, Section 2.2].

In [CD14, §4.3], Cuturi and Doucet propose to consider the problem of minimizing $\tilde{\mathcal{E}}$ as a minimization problem over the positions $\{y_1 \dots y_{m_{\text{tot}}}\}$ describing (the current estimation of) the barycenter. They derive a Newton update rule that reads (in the context of persistence diagrams)

$$\{y_1 \dots y_{m_{\text{tot}}}\} = Y \leftarrow \text{mean}(X_i P_i^{*,T}),$$

where X_i denotes the positions of the points of a_i (completed with a mass $m_{\text{tot}} - a_i(\Omega)$ on $\partial\Omega$), and $P_i^{*,T}$ is an optimal transport plan between Y and X_i for the cost matrix $C(X, Y)$ defined in (5.5). Enforcing P_i to be

a permutation matrix (which is always possible thanks to the Birkhoff-von Neumann theorem, see Section 2.2) exactly yields the update rule of Algorithm 1. To sum up, Algorithm 1 can be understood as a particular instance[1] of [CD14, Algorithm 2].

Limitations. In particular, Algorithm 1 only provides a local minimizer of $\mathcal{E}$ in $\mathcal{D}$. Here, a local minimizer must be understood as an element $\mu \in \mathcal{D}$ for which there is some $\varepsilon > 0$, such that

$$\forall \nu \in \mathcal{D}, d_2(\mu, \nu) \leq \varepsilon \Rightarrow \mathcal{E}(\mu) \leq \mathcal{E}(\nu).$$

One can naturally wonder if such local minimizers do exist, and how "bad" they are. Unfortunately, we prove that $\mathcal{E}$ admits arbitrary bad local minimizers in $\mathcal{D}$.

Lemma 5.2. *For any $\kappa \geq 1$, there exists a set of diagrams such that $\mathcal{E}$ admits a local minimizer μ_{loc} in $\mathcal{D}$ satisfying:*

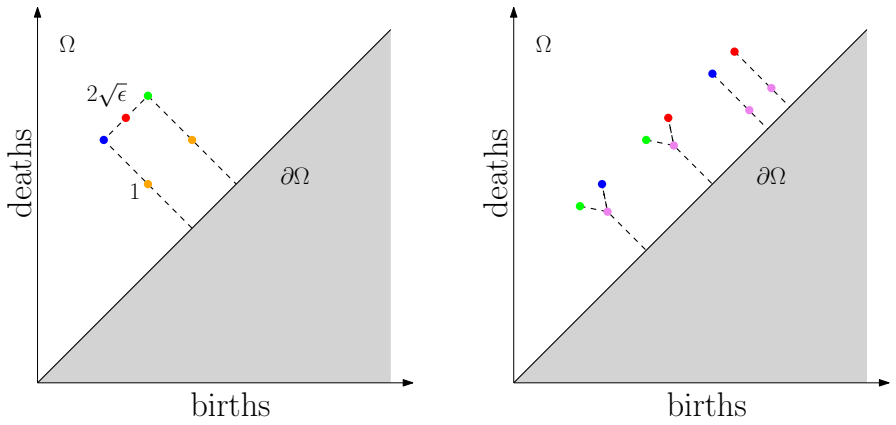
$$\mathcal{E}(\mu_{\text{loc}}) \geq \kappa \mathcal{E}(\mu_{\text{opt}})$$

where μ_{opt} is a global minimizer of $\mathcal{E}$. Furthermore, there exist sets of diagrams so that the B-Munkres algorithm always converges to such a local minimum when initialized with one of the input diagram (as suggested in [TMMH14]).

Proof. Consider the configuration of Figure 5.2, (a): two diagrams with 1 point (blue and green diagram) and their correct barycenter (red diagram) along with the orange diagram (2 points). The orange diagram is a local minimizer of $\mathcal{E}$ in $\mathcal{D}$ (in which the algorithm could get stuck if initialized poorly). It achieves an objective value of $\frac{1}{2}((\frac{1}{2}+\frac{1}{2})^2 + (\frac{1}{2}+\frac{1}{2})^2) = 1$ while the red diagram achieves a value of $\frac{1}{2}(\sqrt{\varepsilon}^2 + \sqrt{\varepsilon}^2) = \varepsilon$. This example proves that there exist configurations of diagrams so that $\widehat{\mathcal{E}}$ has arbitrary bad local minima.

One could hope that when initialized to one of the input diagram (as suggested in [TMMH14]), the algorithm will not get stuck to the orange diagram. Figure 5.2 (b) provides a configuration involving three diagrams with two points each where the algorithm will always get stuck in a bad local minimum when initialized with any of the three diagrams. The analysis is similar to previous statement. □

[1] More precisely, a slight variation, as Cuturi and Doucet only considered the Euclidean cost. On the quotient space $\tilde{\Omega}$, d is not the Euclidean cost, but computations adapt faithfully thanks to the block structure of C mentioned in (5.5).



(a) Example of arbitrary bad local minima of $\widehat{\mathcal{E}}$.

(b) Failing configuration for B-Munkres algorithm.

Figure 5.2: Example of simple configurations in which the B-Munkres algorithm will converge to arbitrarily bad local minima

Aside from its non-convexity, Algorithm 1 suffers in practice from a high computational cost. Indeed, solving partial matching problems between two diagrams of size n scales (just as any optimal transport problem) essentially in $\mathcal{O}(n^3)$. As Algorithm 1 requires, at each iteration, to compute N optimal partial matchings, it is untractable when the number of diagram N and the typical size of the diagram n increase, which is confirmed numerically (see Figure 5.8 below).

Improving on these two issues. Algorithm 1 produces an estimation of a Fréchet mean by optimizing on positions: it is a *Lagrangian* approach (see Section 2.2). Lagrangian approaches are known to lead to non-convex optimization problems (see for instance [CD14, §4.4]). In contrast, we adopt below a *Eulerian* approach which—at the cost of discretizing the ground space—allows us to recover a convex optimization problem.

In terms of computational efficiency, we rely on an *entropic regularization* of optimal transport, see Section 2.2.4 for an introduction. A careful adaptation of computational techniques developed in computational optimal transport [PC17] yields to a scalable problem than can be run efficiently on GPUs, drastically improving running times on large sets of large diagrams.

Remark 5.2. We end this section by mentioning a concurrent approach to improve on Algorithm 1 introduced in [VBT19, KVT19]. Briefly speaking, it consists in using the Auction algorithm—as adapted in [KMN17]—to

estimate optimal partial matchings between persistence diagrams. This approach significantly reduces the running time in Algorithm 1 and has the benefits of providing an output that is an actual persistence diagram (while the approach we develop below returns a measure supported on the discretized groundspace). However, the resulting problem is still not convex.

5.3 Entropic approximation for distances and means of persistence diagrams

Following the correspondance established in (5.6), entropic regularization of optimal transport metrics (we refer the reader to Section 2.2.4 for more details) can be used to approximate the diagram distance $\mathbf{d}_p(\cdot,\cdot)$. Given two persistence diagrams μ, ν with respective mass n_1, n_2, let $n := n_1 + n_2$, $a = (\mathbf{1}_{n_1}, n_2) \in \mathbb{R}^{n_1+1}$, $b = (\mathbf{1}_{n_2}, n_1) \in \mathbb{R}^{n_2+1}$, and let C be the cost matrix as defined in (5.5). In this context, given a regularization parameter $\gamma > 0$, one can approximate the distance between μ and ν using one of the following optimization problems, dual of each other:

$$\begin{aligned}
\mathbf{d}_p(\mu,\nu)^p \simeq \mathbf{d}_p^\gamma(\mu,\nu)^p &:= \inf_{P\in\Pi(a,b)} \langle P, C\rangle + \gamma h(P), && (5.11)\\
&= \sup_{(f,g)\in\mathbb{R}^{n_1+1}\times\mathbb{R}^{n_2+1}} \langle a, f\rangle + \langle b, g\rangle && (5.12)\\
&\quad - \varepsilon \sum_{ij} \exp\left(\frac{f_i + g_j - C_{ij}}{\varepsilon}\right), && (5.13)\\
&=: \mathbf{L}_C^\gamma(a,b) && (5.14)
\end{aligned}$$

where $\Pi(a,b)$ is defined in (5.9), $h(P) := \sum_{ij} P_{ij}(\log(P_{ij}) - 1)$ is the *negentropy*. For an optimal P^γ, we define the *Sinkhorn cost* as $\mathbf{S}_C^\gamma(a,b) := \langle P^\gamma, C\rangle$. Note that one has $\mathbf{S}_C^\gamma(a,b) \geq \mathbf{L}_C^\gamma(a,b)$.

The seminal strength of the minimization problem involved in (5.11) is that it is strictly convex and thus admits a unique solution P^γ. This solution can be approximated by iterating the *Sinkhorn map*

$$S : (u,v) \in \mathbb{R}^{n_1+1} \times \mathbb{R}^{n_2+1} \mapsto \left(\frac{a}{Kv}, \frac{b}{K^T u}\right), \tag{5.15}$$

where $K = \exp(-C/\varepsilon) \in \mathbb{R}^{(n_1+1)\times(n_2+1)}$ (where exponentiation must be understood term-wise). Indeed, if (u^γ, v^γ) is a fixed point of (5.15), then it turns out that $u^\gamma K v^\gamma = P^\gamma$. Furthermore, for any initialization (u_0, v_0),

the sequence defined by

$$u_{t+1} = \frac{a}{Kv_t} \tag{5.16}$$

$$v_{t+1} = \frac{b}{K^T u_{t+1}}, \tag{5.17}$$

converges to an optimal pair (u^γ, v^γ). Note also the relation $(u^\gamma, v^\gamma) = (\exp(f^\gamma/\gamma), \exp(g^\gamma/\gamma))$, where (f^γ, g^γ) is optimal in (5.14).

5.3.1 Error control

The above definition gives us a way to approximate $\mathbf{d}_p$ by $\mathbf{d}_p^\gamma$ by simply iterating the Sinkhorn map S (note that in practice, additional care is required to obtain much faster computations). Of course, one might ask for some control on the error made by using $\mathbf{d}_p^\gamma$ instead of $\mathbf{d}_p$. There are two sources of error:

- The regularization error induced by the parameter γ
- The fact that, in practice, we do not exactly reach a fixed point of S as this would required to perform "infinitely many" iterations.

Therefore, in practice, we obtain a transport plan $P_t^\gamma = \mathrm{diag}(u_t^\gamma) K \mathrm{diag}(v_t^\gamma)$ where (u_t^γ, v_t^γ) is the output after t iterations of the Sinkhorn map (5.15). The error made by approximating $\mathbf{d}_p(\mu,\nu)^p$ by $\langle P_t^\gamma, C\rangle$ has been studied in [AWR17]. Adapted to our context, their bounds read

$$|\mathbf{d}_p(\mu,\nu)^p - \langle P_t^\gamma, C\rangle| \leq 2\gamma n \log(n) + \mathrm{dist}(P_t^\gamma, \Pi(a,b))\|C\|_\infty \tag{5.18}$$

where

$$\mathrm{dist}(P, \Pi(a,b)) := \|P\mathbf{1} - a\|_1 + \|P^T\mathbf{1} - b\|_1,$$

(that is, error on marginals). It remains to understand at which rate $\mathrm{dist}(P, \Pi(a,b)) \to 0$ as the number of iterations $t \to \infty$. In [DGK18], authors prove that iterating the Sinkhorn map (5.15) gives a plan P_t^γ satisfying $\mathrm{dist}(P_t^\gamma, \Pi(a,b)) < \varepsilon$ in $\mathcal{O}\left(\frac{\|C\|_\infty^2}{\gamma\varepsilon} + \ln(n)\right)$ iterations. Given (5.18), a natural choice is thus to take $\gamma = \frac{\varepsilon}{n\ln(n)}$ for a desired precision ε, which lead to a total of $\mathcal{O}\left(\frac{n\ln(n)\|C\|_\infty^2}{\varepsilon^2}\right)$ iterations in the Sinkhorn loop. These results can be used to pre-tune parameters t and γ to control the approximation error due to smoothing. However, these are worst-case bounds, controlled by max-norms, and are often too pessimistic in practice. To overcome this

phenomenon, we introduce on-the-fly error control, using approximate solutions to the smoothed primal (5.11) and dual (5.14) optimal transport problems, which provide upper and lower bounds on the optimal transport cost.

Upper and Lower Bounds. The Sinkhorn algorithm, after at least one iteration ($t \geq 1$), produces feasible variables for the (non-regularized) dual problem (5.8):

$$(f_t^\gamma, f_t^\gamma) = (\gamma \log(u_t^\gamma), \gamma \log(v_t^\gamma)) \in \Psi_C,$$

providing a natural lower bound of the target quantity

$$\langle f_t^\gamma, a \rangle + \langle g_t^\gamma, b \rangle \leq \sup_{f,g \in \Psi_C} \langle a, f \rangle + \langle b, g \rangle = \mathbf{d}_p(\mu, \nu)^p.$$

However, it turns out that in practice, this quantity poorly as a lower bound of the true optimal transport cost (see Figure 5.3 and Section 5.3.4 below) in most of our experiments. To improve on this, we computed the so called *C-transform* $(f_t^\gamma)^c$ of f_t^γ [San15, §1.6], defined as:

$$\forall j, (f_t^\gamma)_j^c = \max_i \{C_{ij} - f_i\}, \ j \leq n_2 + 1.$$

Applying a C^T-transform on $(f_t^\gamma)^c$, we recover two vectors $(f_t^\gamma)^{c\bar{c}} \in \mathbb{R}^{n_1+1}, (f_t^\gamma)^c \in \mathbb{R}^{n_2+1}$. For any feasible f, g, we have that [PC17, Prop 3.1]

$$\langle f, a \rangle + \langle g, b \rangle \leq \langle f^{c\bar{c}}, a \rangle + \langle f^c, b \rangle .$$

When C's top-left block is the squared Euclidean metric, this problem can be cast as that of computing the Moreau envelope [Mor65] of f. In a Eulerian setting and when our measures are described as histograms supported on a finite regular grid (which we consider below), the C-transform in linear time with respect to the grid resolution d by using either a linear-time Legendre transform or the Parabolic Envelope algorithm [Luc10, §2.2.1,§2.2.2].

Unlike dual iterations, the *primal* iterate P_t^γ does *not* belong to the transport polytope $\Pi(a, b)$ after a finite number t of iterations. We use the `rounding_to_feasible` algorithm introduced in [AWR17] to compute efficiently a feasible approximation R_t^γ of P_t^γ that does belong to $\Pi(a, b)$. Putting these two elements together, we obtain

$$\underbrace{\langle (f_t^\gamma)^{c\bar{c}}, a \rangle + \langle (f_t^\gamma)^c, b \rangle}_{m_t^\gamma} \leq \mathbf{L}_C(a, b) \leq \underbrace{\langle R_t^\gamma, C \rangle}_{M_t^\gamma}. \tag{5.19}$$

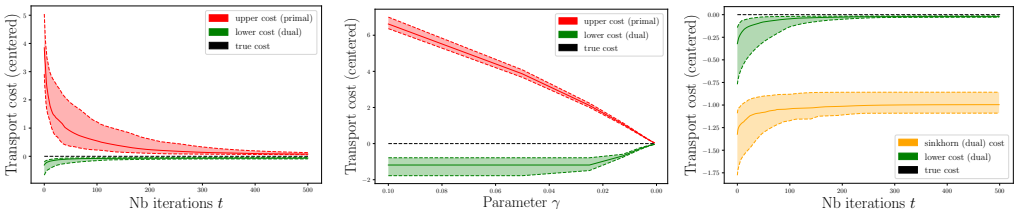


Figure 5.3: *(Left)* $M_t^\gamma := \langle R_t^\gamma, C\rangle$ (red) and $m_t^\gamma := \langle \alpha_t^{c\bar{c}}, a\rangle + \langle \alpha_t^c, b\rangle$ (green) as a function of t, the number of iterations of the Sinkhorn map (t ranges from 1 to 500, with fixed $\gamma = 10^{-3}$). *(Middle)* Final M^γ (red) and m^γ (green) provided by Alg.1, computed for decreasing γs, ranging from 10^{-1} to 5.10^{-4}. For each value of γ, Sinkhorn loop is run until $d(P_t^\gamma, \Pi(a,b)) < 10^{-3}$. Note that the γ-axis is flipped. *(Right)* Influence of $c\bar{c}$-transform for the Sinkhorn dual cost. *(orange)* The dual cost $\langle \alpha_t^\gamma, a\rangle + \langle \beta_t^\gamma, b\rangle$, where $(\alpha_t^\gamma, \beta_t^\gamma)$ are Sinkhorn dual variables (before the C-transform). *(green)* Dual cost after C-transform, i.e. with $((\alpha_t^\gamma)^{c\bar{c}}, (\alpha_t^\gamma)^c)$. Experiment run with $\gamma = 10^{-3}$ and 500 iterations.

Therefore, after iterating the Sinkhorn map (5.15) t times, we have that if $M_t^\gamma - m_t^\gamma$ is below a certain criterion ε, then we can guarantee that $\langle R_t^\gamma, C\rangle$ is *a fortiori* an ε-approximation of $\mathbf{L}_C(a,b)$. Observe that one can also have a relative error control: if one has $M_t^\gamma - m_t^\gamma \leq \varepsilon M_t^\gamma$, then $(1-\varepsilon)M_t^\gamma \leq L_C(a,b) \leq M_t^\gamma$. Note that m_t^γ might be negative but can always be replaced by $\max(m_t^\gamma, 0)$ since C has non-negative entries (and therefore $\mathbf{L}_C(a,b) \geq 0$), while M_t^γ is always non-negative.

5.3.2 Computational aspects

Discretization. As mentioned in the beginning of the section, we assume in the remaining that our diagrams have their support in $[0,1]^2 \cap \Omega$. An explained previously, updating positions (Lagrangian approaches) generally leads to non-convex problem. To recover a convex problem, we propose instead to adopt an Eulerian approach: just as a persistence diagram μ can be describe as a set of positions $\{x_1 \dots x_n\}$, it can also be seen as a map $x \in \Omega \mapsto 1$ if x is a point of the diagram, 0 otherwise. That is, we will optimize the mass put on each x in Ω. Of course, this theoretically leads to an infinite-dimensional optimization problem. In practice, we have to discretize the ground space to recover a tractable optimization problem, that has the huge advantage of being convex.

From a numerical perspective, encoding persistence diagrams as histograms on the square offers numerous advantages. Given a uniform grid of size $d \times d$ on $[0,1]^2$, we associate to a given diagram μ a *matrix-shaped*

histogram $\mathbf{a} \in \mathbb{R}^{d\times d}$ such that $\mathbf{a}_{ij}$ is the number of points in μ belonging to the cell located at position (i,j) in the grid (we use bold-faced small letters to insist on the fact that these histograms must be stored as square matrices). Note that, in order to account for the mass located on $\partial\Omega$, we must add an additional coordinate that will be denoted by Δ. We extend the operator $\mathbf{R}$ to histograms, associating to a histogram $\mathbf{a} \in \mathbb{R}^{d\times d}$ its total mass on the Δ coordinate (the (d^2+1)-th one).

Convolutions. In the Eulerian setting, where diagrams are matrix-shaped histograms of size $d \times d = d^2$, the cost matrix would have size $d^2 \times d^2$. Since one may want to use large values of d to have low discretization error (typically $d = 100$), instantiating the cost matrix naively is usually intractable. However, [SDGP+15] showed that for regular grids endowed with a separable cost, each Sinkhorn iteration (as well as other key operations such as evaluating Sinkhorn's cost $\mathbf{S}^\gamma_C = \langle P^\gamma, C\rangle$) can be performed using Gaussian convolutions, which amounts to performing matrix multiplications of size $d\times d$, without having to manipulate $d^2\times d^2$ matrices at any step. The framework we consider is slightly different due to the extra dimension $\{\partial\Omega\}$ (and the corresponding row and column in (5.5)), but equivalent computational properties hold. This observation is crucial from a numerical perspective. The idea is sketch in Figure 5.4.

Implementation details. Let $(\mathbf{u}, u_\Delta)$ be a pair where $\mathbf{u} \in \mathbb{R}^{d\times d}$ is a matrix-shaped histogram and $u_\Delta \in \mathbb{R}_+$ is a real number accounting for the mass located the Δ coordinate (i.e. on $\{\partial\Omega\}$). We denote by $\overrightarrow{\mathbf{u}}$ the $d^2 \times 1$ column vector obtained when reshaping $\mathbf{u}$. The $(d^2+1)\times(d^2+1)$ cost matrix C and corresponding kernel K are given by

$$C = \begin{pmatrix} \widehat{C} & \overrightarrow{\mathbf{c}_\Delta} \\ \overrightarrow{\mathbf{c}_\Delta}^T & 0 \end{pmatrix}, \quad K = \begin{pmatrix} \widehat{K} := e^{-\frac{\widehat{C}}{\gamma}} & \overrightarrow{\mathbf{k}_\Delta} := e^{-\frac{\overrightarrow{\mathbf{c}_\Delta}}{\gamma}} \\ \overrightarrow{\mathbf{k}_\Delta}^T & 1 \end{pmatrix},$$

where $\widehat{C} = (\|(i,i') - (j,j')\|_p^p)_{ii',jj'}$, $\mathbf{c}_\Delta = (\|(i,i') - \pi_\Delta((i,i'))\|_p^p)_{ii'}$. The crucial point is that C and K as defined above will never be instantiated, because we can rely instead on $\mathbf{c} \in \mathbb{R}^{d\times d}$ defined as $\mathbf{c}_{ij} = |i-j|^p$ and $\mathbf{k} = e^{-\frac{\mathbf{c}}{\gamma}}$.

Proposition 5.3 (Iteration of Sinkhorn map)**.** *The application of K to $(\mathbf{u}, u_\Delta)$ can be performed as:*

$$(\mathbf{u}, u_\Delta) \mapsto \left(\mathbf{k}(\mathbf{k}\mathbf{u}^T)^T + u_\Delta \mathbf{k}_\Delta, \langle \mathbf{u}, \mathbf{k}_\Delta\rangle + u_\Delta\right) \tag{5.20}$$

where $\langle\cdot,\cdot\rangle$ denotes the Froebenius dot-product in $\mathbb{R}^{d\times d}$.

We now introduce $\mathbf{m} := \mathbf{k} \odot \mathbf{c}$ and $\mathbf{m}_\Delta := \mathbf{k}_\Delta \odot \mathbf{c}_\Delta$ ($\odot$ denotes term-wise multiplication).

Proposition 5.4 (Computation of $\mathbf{S}_C^\gamma$)**.** *Let $(\mathbf{u}, u_\Delta), (\mathbf{v}, v_\Delta) \in \mathbb{R}^{d \times d+1}$. The transport cost of $P := \mathrm{diag}(\overrightarrow{\mathbf{u}}, u_\Delta) K \mathrm{diag}(\overrightarrow{\mathbf{v}}, v_\Delta)$ can be computed as:*

$$\begin{aligned}\langle \mathrm{diag}(\overrightarrow{\mathbf{u}}, u_\Delta) K \mathrm{diag}(\overrightarrow{\mathbf{v}}, v_\Delta), C \rangle = &\langle \mathrm{diag}(\overrightarrow{\mathbf{u}}) \widehat{K} \mathrm{diag}(\overrightarrow{\mathbf{v}}), \widehat{C} \rangle \\ &+ u_\Delta \langle \mathbf{v}, \mathbf{m}_\Delta \rangle + v_\Delta \langle \mathbf{u}, \mathbf{m}_\Delta \rangle ,\end{aligned} \tag{5.21}$$

where the first term can be computed as:

$$\langle \mathrm{diag}(\overrightarrow{\mathbf{u}}) \widehat{K} \mathrm{diag}(\overrightarrow{\mathbf{v}}), \widehat{C} \rangle = \| \mathbf{u} \odot \left(\mathbf{m}(\mathbf{k}\mathbf{v}^T)^T + \mathbf{k}(\mathbf{m}\mathbf{v}^T)^T \right) \|_1. \tag{5.22}$$

The last proposition addresses the computation of the upper bound involved in (5.19). First, we recall the `rounding_to_feasible` algorithm introduced by Altschuler et al. in [AWR17]. Here, $r(P)$ and $c(P)$ denote respectively the first and second marginal of a matrix P.

Algorithm 2 Rounding algorithm of [AWR17]

1: **Input:** $P \in \mathbb{R}^{d \times d}$, desired marginals r, c.
2: **Output:** $F(P) \in \Pi(r, c)$ close to P.
3: $X = \min\left(\frac{r}{r(P)}, 1\right) \in \mathbb{R}^d$
4: $P' = \mathrm{diag}(X) P$
5: $Y = \min\left(\frac{c}{c(P')}, 1\right) \in \mathbb{R}^d$
6: $P'' = P' \mathrm{diag}(Y)$
7: $e_r = r - r(P''), e_c = c - c(P'')$
8: **return** $F(P) := P'' + e_r e_c^T / \|e_c\|_1$

Finally, consider two histograms $(\mathbf{a}, a_\Delta), (\mathbf{b}, b_\Delta) \in \mathbb{R}^{d \times d} \times \mathbb{R}$, let $R \in \Pi((\mathbf{a}, a_\Delta), (\mathbf{b}, b_\Delta))$ be the rounded matrix of P as introduced in Algorithm 2. Let $r(P), c(P) \in \mathbb{R}^{d \times d} \times \mathbb{R}$ denote the first and second marginal of P respectively. We introduce (minimum and divisions must be understood term-wise):

$$X = \min\left(\frac{(\mathbf{a}, a_\Delta)}{r(P)}, \mathbf{1}\right), \quad Y = \min\left(\frac{(\mathbf{b}, b_\Delta)}{c(\mathrm{diag}(X)P)}, \mathbf{1}\right),$$

along with $P' = \mathrm{diag}(X) P \mathrm{diag}(Y)$ and the marginal errors:

$$(e_r, (e_r)_\Delta) = (\mathbf{a}, a_\Delta) - r(P'), \quad (e_c, (e_c)_\Delta) = (\mathbf{b}, b_\Delta) - c(P'),$$

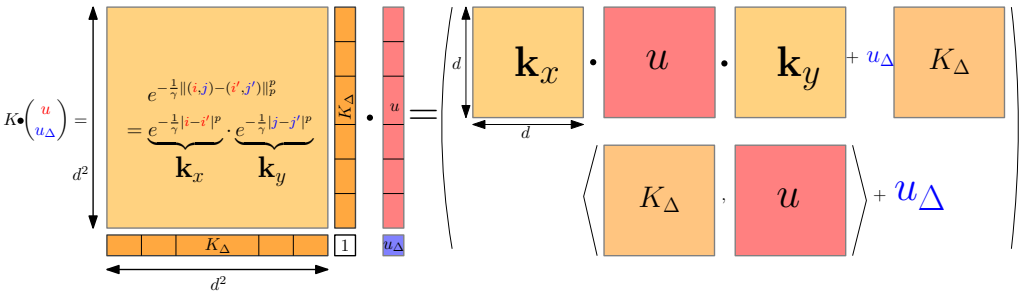


Figure 5.4: Sketch of the convolution technique used to efficiently iterate the Sinkhorn map (5.15) in a Eulerian setting.

Proposition 5.5 (Computation of the upper bound $\langle R, C \rangle$)**.** *The transport cost induced by R can be computed as:*

$$\begin{aligned} \langle R, C \rangle = &\langle \mathrm{diag}(X \odot (\mathbf{u}, u_\Delta)) K \mathrm{diag}(Y \odot (\mathbf{v}, v_\Delta)), C \rangle \\ &+ \frac{1}{\|e_c\|_1 + (e_c)_\Delta} \big(\|e_r^T \mathbf{c} e_c\|_1 + \|e_r \mathbf{c} e_c^T\|_1 \\ &+ (e_c)_\Delta \langle e_r, \mathbf{c}_\Delta \rangle + (e_r)_\Delta \langle e_c, \mathbf{c}_\Delta \rangle \big). \end{aligned} \tag{5.23}$$

Note that the first term can be computed using (5.21)

The proofs of these propositions are purely technical and can be found in Section 5.3.5.

Parallelization and GPU. Using a Eulerian representation is particularly beneficial when applying Sinkhorn's algorithm, as shown by [Cut13]. Indeed, the Sinkhorn map (5.15) only involves matrix-vector operations. Furthermore, the cost matrix C is constant: it does not depend on the pair of histograms involved. When dealing with a large number of histograms, concatenating these histograms and running Sinkhorn's iterations in parallel as matrix-matrix product results in significant speedup that can exploit GPGPU to compare a large number of pairs simultaneously. This makes this approach especially well-suited for large sets of persistence diagrams. This idea is sketched in Figure 5.5.

Algorithm 3 allows us to estimate distances between persistence diagrams in parallel by performing only $(d \times d)$-sized matrix multiplications, leading to a computational scaling in d^3 where d is the grid resolution parameter. Note that a standard stopping threshold in Sinkhorn iteration process is to check the error to marginals $\mathrm{dist}(P, \Pi(\mathbf{a}, \mathbf{b}))$, as motivated by (5.18).

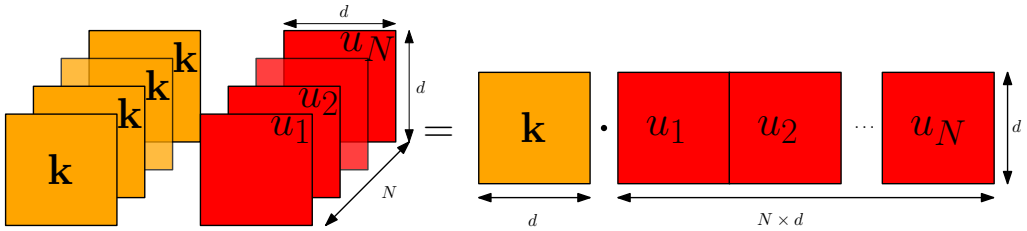


Figure 5.5: Sketch of parallel computations of S when dealing with multiple pairs of diagrams.

Algorithm 3 Sinkhorn cost for persistence diagrams

Input: Pairs of PDs $(D_i, D'_i)_i$, smoothing parameter $\gamma > 0$, grid step $d \in \mathbb{N}$, stopping criterion, initial $(\mathbf{u}, \mathbf{v})$.
Output: Approximation of all $(\mathbf{d}_p(D_i, D'_i)^p)_i$, upper and lower bounds if wanted.
init Cast D_i, D'_i as histograms $\mathbf{a}_i$, $\mathbf{b}_i$ on a $d \times d$ grid
while stopping criterion not reached **do**
  Iterate in parallel (5.15) $(\mathbf{u}, \mathbf{v}) \mapsto S(\mathbf{u}, \mathbf{v})$ using (5.20)
end while
Compute all $\mathbf{S}_C^\gamma(\mathbf{a}_i + \mathbf{R}\mathbf{b}_i, \mathbf{b}_i + \mathbf{R}\mathbf{a}_i)$ using (5.21)
if Want a upper bound **then**
  Compute $\langle R_i, C \rangle$ in parallel using (5.23)
end if
if Want a lower bound **then**
  Compute $\langle (\alpha_t^\gamma)^{c\bar{c}}, \mathbf{a}_i \rangle + \langle (\alpha_t^\gamma)^c, \mathbf{b}_i \rangle$ using [Luc10]
end if

5.3.3 Entropic regularization for Fréchet means in $\mathcal{D}$

In addition to numerical efficiency, a major advantage of the regularized optimal transport is that

$$a \mapsto \mathbf{L}_C^\gamma(a, b)$$

is differentiable. In the Eulerian setting, where C is fixed, its gradient is given by centering the vector $\gamma \log(u^\gamma)$ where u^γ is a fixed point of the Sinkhorn map (5.15) (see [CD14]). This result can be adapted to our framework, namely:

Proposition 5.6. *Let $\nu_1 \dots \nu_N$ be PDs, and $(\mathbf{a}_i)_i$ the corresponding histograms on a $d \times d$ grid. Let C_d be the corresponding $(d^2+1) \times (d^2+1)$ cost*

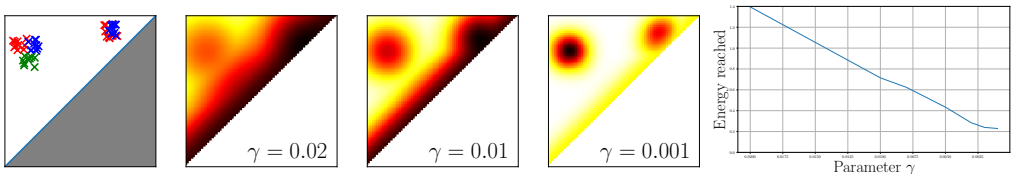


Figure 5.6: Barycenter estimation for different γs with a simple set of 3 PDs (red, blue and green). The smaller the γ, the better the estimation ($\mathcal{E}$ decreases, note the γ-axis is flipped on the right plot), at the cost of more iterations in Alg. 4. The mass appearing along the diagonal is a consequence of entropic smoothing: it does not cost much to delete while it increases the entropy of transport plans.

matrix. Note that C_d does not depend on $\mathbf{x}$. The gradient of the functional

$$\mathcal{E}^{\gamma,d} : \mathbf{x} \mapsto \sum_{i=1}^{N} \mathbf{L}^{\gamma}_{C_d}(\mathbf{x} + \mathbf{R}\mathbf{a}_i, \mathbf{a}_i + \mathbf{R}\mathbf{x})$$

is given by

$$\nabla_{\mathbf{x}} \mathcal{E}^{\gamma,d} = \gamma \left(\sum_{i=1}^{N} \log(u_i^{\gamma}) + \mathbf{R}^T \log(v_i^{\gamma}) \right) \tag{5.24}$$

where $\mathbf{R}^T$ denotes the adjoint operator $\mathbf{R}$ and $(u_i^{\gamma}, v_i^{\gamma})$ is a fixed point of the Sinkhorn map obtained while transporting $\mathbf{x}+\mathbf{R}\mathbf{a}_i$ onto $\mathbf{a}_i+\mathbf{R}\mathbf{x}$ for the cost C.

This result follows from the envelope theorem, taking into account the fact that $\mathbf{x}$ appears in both terms depending on u and v. This formula can be exploited to compute barycenters via gradient descent, yielding Algorithm 4. Following [CD14, §4.2], we use a multiplicative update. This is a particular case of mirror descent [BT03] and is equivalent to a (Bregman) projected gradient descent on the positive orthant, retaining positive coefficients (masses) throughout iterations.

Remark 5.3 (A comparison with linear representations). When doing statistical analysis with persistence diagrams, a standard approach is to transform a diagram into a finite-dimensional vector—in a stable way—and then perform statistical analysis and learning with a Euclidean structure. This approach does not preserve the Wasserstein-like geometry of the diagram space and thus loses the algebraic interpretability of persistence diagrams. Figure 5.7 gives a qualitative illustration of the difference between Wasserstein barycenters (Fréchet mean) of persistence diagrmas and Euclidean barycenters (linear means) of persistence images [AEK+17], a commonly used vectorization for PDs [MKPY16, ZZJS16, OHK18].

Algorithm 4 Estimation of Fréchet means of persistence diagrams

Input: Persistence diagrams $\nu_1, \dots, \nu_N$, learning rate λ, smoothing parameter $\gamma > 0$, grid step $d \in \mathbb{N}$.
Output: Estimated barycenter $\mathbf{x}$
Init: $\mathbf{x}$ uniform measure above the diagonal.
Cast each ν_i as a histogram $\mathbf{a}_i$ on a $d \times d$ grid
while $\mathbf{x}$ changes **do**
  Iterate S defined in (5.15) in parallel between all the pairs $(\mathbf{x} + \mathbf{R}\mathbf{a}_i)_i$ and $(\mathbf{a}_i + \mathbf{R}\mathbf{x})_i$, using (5.20).
  $\nabla := \gamma(\sum_i \log(u_i^\gamma) + \mathbf{R}^T \log(v_i^\gamma))$
  $\mathbf{x} := \mathbf{x} \odot \exp(-\lambda \nabla)$
end while
if Want energy **then**
  Compute $\frac{1}{N} \sum_i \mathbf{S}_C^\gamma(\mathbf{x} + \mathbf{R}\mathbf{a}_i, \mathbf{a}_i + \mathbf{R}\mathbf{x})$ using (5.21)
end if
Return $\mathbf{x}$

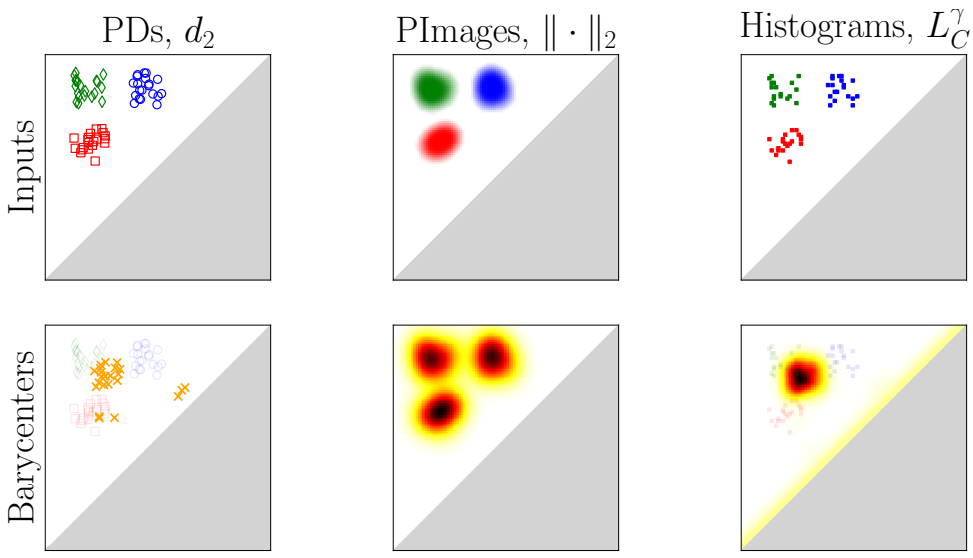


Figure 5.7: Illustration of differences between Fréchet means with Wasserstein and Euclidean geometry. The top row represents input data, namely persistence diagrams (left), vectorization of diagrams as persistence images in $\mathbb{R}^{100\times100}$ (middle, [AEK+17]), and discretization of diagrams as histograms (right). The bottom row represents the estimated barycenters (orange scale) with input data (shaded), using the approach of Algorithm 1 [TMMH14] (left), the arithmetic mean of persistence images (middle) and our optimal transport based approach (right).

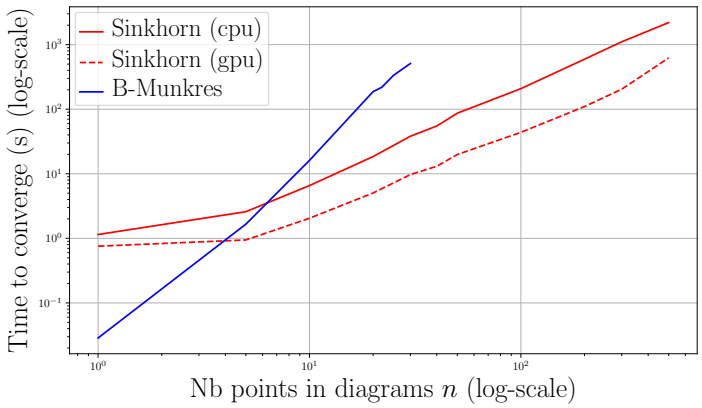


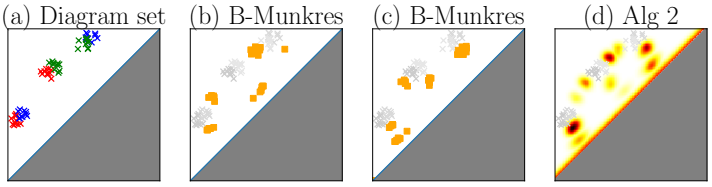
Figure 5.8: Average running times for B-Munkres *(blue)* and Sinkhorn *(red)* algorithms (log-log scale) to average 10 PDs.

Remark 5.4. Aside from its computational strengths (convexity, parallelism, GPU-capable), this Eulerian approach has another major advantage over Lagrangian ones: it straightforwardly extends to any measure supported on $[0,1]^2 \cap \Omega$ and is not restricted to persistence diagram (nor even to measures with finite support). Therefore, this approach would allow us to compute (more precisely, to estimate) distance and barycenters between more complicated measures supported on Ω, such as *expected persistence diagrams* (see Section 7.1).

5.3.4 Experiments

Let us now showcase Algorithm 3 and Algorithm 4 in numerical applications. In the following, experiments are run with $p = 2$, but would work with any finite $p \geq 1$. The choice of $p = 2$ is consistent with the work of [TMMH14] for Fréchet mean estimation.

Fast barycenters Let us compare Algorithm 4 (referred to as *Sinkhorn*) with the combinatorial algorithm of the B-Munkres algorithm Algorithm 1. Figure 5.8 records running times of both algorithms on a set of 10 diagrams having from n to $2n$ points, n ranging from 1 to 500, on Intel Xeon 2.3 GHz (CPU) and P100 (GPU, Sinkhorn only). When running Algorithm 4, the gradient descent is performed until $|\mathcal{E}(\mathbf{x}_{t+1})/\mathcal{E}(\mathbf{x}_t) - 1| < 0.01$, with $\gamma = 10^{-1}/n$ and $d = 50$. Our experiment shows that Algorithm 4 drastically outperforms *B-Munkres* as the number of points n increases. We interrupt *B-Munkres* at $n = 30$, after which computational time becomes an issue.



	Final energy
B-Munkres (*b*)	0.589
B-Munkres (*c*)	0.555
Alg. 2 (*d*)	0.542

Figure 5.9: Qualitative comparison of B-Munkres and Algorithm 4. (*a*) Input set of $N = 3$ diagrams with $n = 20$ points each. (*b*) Output of B-Munkres when initialized on the blue diagram (orange squares) and input data (grey scale). (*c*) Output of B-Munkres initialized on the green diagram. (*d*) Output of Algorithm 4 on a 100×100 grid, $\gamma = 5.10^{-4}$, learning-rate $\lambda = 5$, Sinkhorn stopping criterion (distance to marginals): 0.001, gradient descent performed until $|\mathcal{E}(\mathbf{z}_{t+1})/\mathcal{E}(\mathbf{z}_t) - 1| < 0.01$.—As one can see, localization of masses is similar. Initialization of B-Munkres is made on one of the input diagram as indicated in [TMMH14, Alg. 1], and leads to convergence to different local minima. Our convex approach (Algorithm 4) performs better (lower energy). As a baseline, the energy of the naive arithmetic mean of the three diagrams is 0.72.

Aside the computational efficiency, we highlight the benefits of operating with a convex formulation in Figure 5.9. Due to non-convexity, the B-Munkres algorithm is only guaranteed to converge to a local minima, and its output depends on initialization. We illustrate on a toy set of $N = 3$ diagrams how our algorithm avoids local minima thanks to the Eulerian approach we take.

k-means on large PD sets. We now merge Algorithm 3 and Algorithm 4 in order to perform unsupervised clustering via k-means on PDs. We work with the 3D-shape database provided by [SP04] and generate diagrams in the same way as in [COO15b], working in practice with 5000 diagrams with 50 to 100 points each. The database contains 6 classes: `camel`, `cat`, `elephant`, `horse`, `head` and `face`. In practice, this unsupervised clustering algorithm detects two main clusters: faces and heads on one hand, camels and horses on the other hand are systematically grouped together. Figure 5.10 illustrates the convergence of our algorithm and the computed centroids for the aforementioned clusters.

5.3.5 Proofs of Propositions 5.3, 5.4, 5.5

Proof of Proposition 5.3. Given a histogram $\mathbf{u} \in \mathbb{R}^{d \times d}$ and a mass $u_\Delta \in \mathbb{R}_+$, one can observe that:

$$\widehat{K}\mathbf{u} = \mathbf{k}(\mathbf{k}\mathbf{u}^T)^T. \tag{5.25}$$

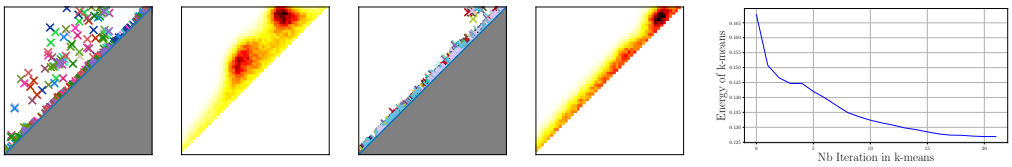


Figure 5.10: Illustration of our k-means algorithm. From left to right: 20 diagrams extracted from *horses* and *camels* plot together (one color for each diagram); the centroid they are matched with provided by our algorithm; 20 diagrams of *head* and *faces*; along with their centroid; decrease of the objective function. Running time depends on many parameters along with the random initialization of k-means. As an order of magnitude, it takes from 40 to 80 minutes with this 5000 PD dataset on a P100 GPU.

In particular, the operation $\mathbf{u} \mapsto \widehat{K}\mathbf{u}$ can be perform by only manipulating matrices in $\mathbb{R}^{d\times d}$.

Indeed, observe that:

$$\widehat{K}_{ij,kl} = e^{-(i-k)^2/\gamma} e^{-(j-l)^2/\gamma} = \mathbf{k}_{ik}\mathbf{k}_{jl},$$

so we have:

$$\begin{aligned}(\widehat{K}\mathbf{u})_{i,j} &= \sum_{k,l} K_{ij,kl}\mathbf{u}_{k,l} \\ &= \sum_{k,l} \mathbf{k}_{ik}\mathbf{k}_{jl}\mathbf{u}_{k,l} = \sum_{k} \mathbf{k}_{ik} \sum_{l} \mathbf{k}_{jl}\mathbf{u}_{kl} \\ &= \sum_{k} \mathbf{k}_{ik}(\mathbf{k}\mathbf{u}^T)_{jk} = (\mathbf{k}(\mathbf{k}\mathbf{u}^T)^T)_{i,j}.\end{aligned}$$

Thus we have in our case:

$$K(\mathbf{u}, u_\Delta) = (\widehat{K}\mathbf{u} + u_\Delta \mathbf{k}_\Delta, \langle \mathbf{u}, \mathbf{k}_\Delta \rangle + u_\Delta)$$

where $\langle a, b \rangle$ designs the Froebenius dot product between two histograms $a, b \in \mathbb{R}^{d\times d}$. These computations only involves manipulation of matrices with size $d \times d$. □

Proof of Proposition 5.4.

$$\begin{aligned}\langle \mathrm{diag}(\overrightarrow{\mathbf{u}})\widehat{K}\,\mathrm{diag}(\overrightarrow{\mathbf{v}}), \widehat{C}\rangle &= \sum_{ijkl} \mathbf{u}_{ij}\mathbf{k}_{ik}\mathbf{k}_{jl}[\mathbf{c}_{ik} + \mathbf{c}_{jl}]\mathbf{v}_{kl} \\ &= \sum_{ijkl} \mathbf{u}_{ij}\left([\mathbf{k}_{ik}\mathbf{c}_{ik}]\mathbf{k}_{jl}\mathbf{v}_{kl} + \mathbf{k}_{ik}[\mathbf{k}_{jl}\mathbf{c}_{jl}]\mathbf{v}_{kl}\right) \\ &= \sum_{ij} \mathbf{u}_{ij} \sum_{kl} \left(\mathbf{m}_{ik}\mathbf{k}_{jl}\mathbf{v}_{kl} + \mathbf{k}_{ik}\mathbf{m}_{jl}\mathbf{v}_{kl}\right)\end{aligned}$$

Thus, we finally have:

$$\langle \mathrm{diag}(\overrightarrow{\mathbf{u}})\widehat{K}\mathrm{diag}(\overrightarrow{\mathbf{v}}), \widehat{C}\rangle = \|\mathbf{u} \odot \left(\mathbf{m}(\mathbf{k}\mathbf{v}^T)^T + \mathbf{k}\mathbf{m}\mathbf{v}^T]^T\right)\|_1$$

And finally, taking the $\{\Delta\}$ bin into considerations,

$$\begin{aligned}
&\langle \mathrm{diag}(\overrightarrow{\mathbf{u}}, u_\Delta)K\mathrm{diag}(\overrightarrow{\mathbf{v}}, v_\Delta), C\rangle \\
&= \langle \begin{pmatrix} \mathrm{diag}(\overrightarrow{\mathbf{u}})\widehat{K}\mathrm{diag}(\overrightarrow{\mathbf{v}}) & v_\Delta(\overrightarrow{\mathbf{u}} \odot \overrightarrow{\mathbf{k}}_\Delta) \\ u_\Delta(\overrightarrow{\mathbf{v}}^T \odot \overrightarrow{\mathbf{k}}_\Delta^T) & u_\Delta v_\Delta \end{pmatrix}, \begin{pmatrix} \widehat{C} & \overrightarrow{\mathbf{c}}_\Delta \\ \overrightarrow{\mathbf{c}}_\Delta^T & 0 \end{pmatrix}\rangle \\
&= \langle \mathrm{diag}(\overrightarrow{\mathbf{u}})\widehat{K}\mathrm{diag}(\overrightarrow{\mathbf{v}}), \widehat{C}\rangle + u_\Delta \langle \mathbf{v}, \mathbf{k}_\Delta \odot \mathbf{c}_\Delta\rangle + v_\Delta \langle \mathbf{u}, \mathbf{k}_\Delta \odot \mathbf{c}_\Delta\rangle,
\end{aligned}$$

concluding the proof. □

Proof of Proposition 5.5. By straightforward computations, the first and second marginals of $P_t^\gamma = \mathrm{diag}(\overrightarrow{\mathbf{u}})K\mathrm{diag}(\overrightarrow{\mathbf{v}})$ are given by:

$$\left(\sum_{kl} \mathbf{u}_{ij}K_{ij,kl}\mathbf{v}_{kl}\right)_{ij} = \mathbf{u} \odot (K\mathbf{v}), \qquad \left(\sum_{ij} \mathbf{u}_{ij}K_{ij,kl}\mathbf{v}_{kl}\right)_{kl} = (\mathbf{u}K) \odot \mathbf{v}.$$

Observe that $K\mathbf{v}$ and $\mathbf{u}K$ can be computed using Proposition 5.3.

Now, the transport cost computation is:

$$\begin{aligned}
\langle F(P_t^\gamma), C\rangle &= \langle \mathrm{diag}(X)P_t^\gamma\mathrm{diag}(Y), C\rangle + \langle e_r e_c^T/\|e_c\|_1, C\rangle \\
&= \langle \mathrm{diag}(X \odot \mathbf{u})K\mathrm{diag}(Y \odot \mathbf{v}), C\rangle + \frac{1}{\|e_c\|_1}\sum_{ijkl}(e_r)_{ij}(e_c)_{kl}[\mathbf{c}_{ik} + \mathbf{c}_{jl}]
\end{aligned}$$

The first term is the transport cost induced by a rescaling of $\mathbf{u}, \mathbf{v}$ and can be computed with Proposition 5.4. Consider now the second term. Without considering the additional bin $\{\Delta\}$, we have:

$$\begin{aligned}
\sum_{ijkl}(e_r)_{ij}(e_c)_{kl}[\mathbf{c}_{ik} + \mathbf{c}_{jl}] &= \sum_{ijl}(e_r)_{ij}\sum_k \mathbf{c}_{ik}(e_c)_{kl} + \sum_{ijk}(e_r)_{ij}\sum_l \mathbf{c}_{jl}(e_c)_{kl} \\
&= \sum_{ijl}(e_r)_{ij}(\mathbf{c}e_c)_{il} + \sum_{ijk}(e_r)_{ij}(\mathbf{c}e_c^T)_{jk} \\
&= \|e_r^T\mathbf{c}e_c\|_1 + \|e_r\mathbf{c}e_c^T\|_1,
\end{aligned}$$

so when we consider our framework (with $\{\Delta\}$), it comes:

$$\begin{aligned}
\langle \begin{pmatrix} e_r \\ (e_r)_\Delta \end{pmatrix} \cdot \begin{pmatrix} e_c & (e_c)_\Delta \end{pmatrix}, C\rangle &= \langle \begin{pmatrix} e_r e_c^T & (e_c)_\Delta e_r \\ (e_r)_\Delta e_c^T & (e_r)_\Delta(e_c)_\Delta \end{pmatrix}, \begin{pmatrix} \widehat{C} & \overrightarrow{\mathbf{c}}_\Delta \\ \overrightarrow{\mathbf{c}}_\Delta^T & 0 \end{pmatrix}\rangle \\
&= \langle e_r e_c^T, \widehat{C}\rangle + (e_c)_\Delta \langle e_r, \mathbf{c}_\Delta\rangle + (e_r)_\Delta \langle e_c, \mathbf{c}_\Delta\rangle.
\end{aligned}$$

Putting things together finally proves the claim. □

Chapter 6

Linear representations of persistence diagrams

Abstract

In this chapter, we study linear representations of persistence diagrams, a common vectorization method to make us of persistence diagrams and measures in machine learning. The first section uses the characterization of convergence proved in Chapter 3 to derive a useful characterization of continuous linear representations in a very general theoretical framework. The second section proposes a neural network architecture that we call PersLay, which aims at *learning* a task-driven representation method in a supervised learning context. It is essentially based on [CCI+19].

Implementation resources.

- (To appear in `Gudhi`) implementation of PersLay. Preliminary version available on Mathieu Carriere's github page, along with a tutorial.

6.1 Continuity of linear representations

Recall that $\mathcal{M}^p$ denotes the space of persistence measures (Radon measures supported on the open half-plane Ω).

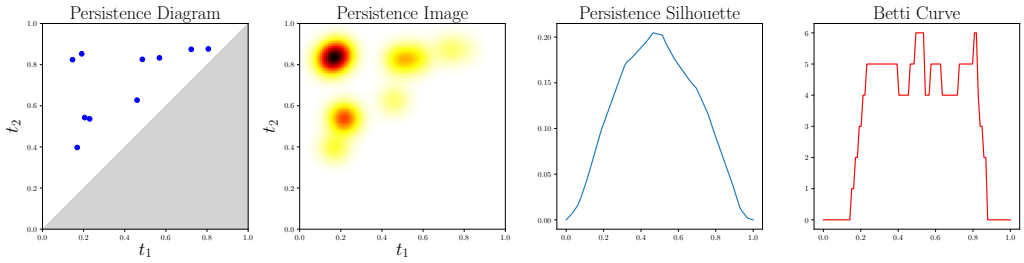


Figure 6.1: Some common linear representations of persistence diagrams. From left to right: A persistence diagram. Its persistence surface [AEK+17], which is a persistence measure. The corresponding persistence silhouette [CFL+14]. The corresponding Betti Curve [Ume17].

A *linear representation* of persistence measures (in particular persistence diagrams) is an map

$$\begin{aligned} \Phi : \mathcal{M}^p &\to \mathcal{B} \\ \mu &\mapsto \mu(f) \end{aligned}$$

for some Banach space $\mathcal{B}$ and some chosen function $f : \Omega \to \mathcal{B}$. Doing so, one can turn a sample of diagrams (or measures) into a sample of vectors, making the use of machine learning tools easier. Of course, a natural requirement is that Φ must be continuous. In practice, building a linear representations (see below for a list of examples) generally follows the same pattern: first consider a "nice" function ϕ, e.g. a Gaussian distribution, then introduce a weight with respect to the distance to the diagonal $d(\cdot, \partial\Omega)^p$, and prove that $\mu \mapsto \mu(\phi(\cdot)d(\cdot, \partial\Omega)^p)$ has some regularity properties (continuity, stability, etc.). By relying on a result proved in Chapter 3, namely Theorem 3.5, we show that this approach always gives a continuous linear representation, and that it is the only way to do.

6.1.1 Characterization of continuity

For $\mathcal{B}$ a Banach space (typically $\mathbb{R}^d$), define the class of functions:

$$\mathcal{C}^0_{b,p} = \left\{ f : \Omega \to \mathcal{B},\ f \text{ continuous and } x \mapsto \frac{f(x)}{d(x, \partial\Omega)^p} \text{ bounded} \right\} \quad (6.1)$$

Proposition 6.1. *Let $\mathcal{B}$ be a Banach space and $f : \Omega \to \mathcal{B}$ a function. The linear representation $\Phi : \mathcal{M}^p \to \mathcal{B}$ defined by $\Phi : \mu \mapsto \mu(f) = \int_\Omega f(x)\mathrm{d}\mu(x)$ is continuous with respect to OT_p if and only if $f \in \mathcal{C}^0_{b,p}$.*

Before proving this result, we recall the content of Theorem 3.5 introduced in Chapter 3: let $(\mu_n)_n$ be a sequence in $\mathcal{M}^p$, and $\mu \in \mathcal{M}^p$. One has

$$\mathrm{OT}_p(\mu_n, \mu) \to 0 \Leftrightarrow \begin{cases} \mu_n \xrightarrow{v} \mu, \\ \mathrm{Pers}_p(\mu_n) \to \mathrm{Pers}_p(\mu). \end{cases}$$

It follows that (Corollary 3.7) that $\mathrm{OT}_p(\mu_n, \mu) \to 0 \Leftrightarrow \mu_n^p \xrightarrow{w} \mu^p$, where the measure μ^p is defined, for $A \subset \Omega$ (Borel), by

$$\mu^p(A) := \int_A d(x, \partial\Omega)^p \mathrm{d}\mu(x).$$

This strong characterization of the convergence in the metric space $(\mathcal{M}^p, \mathrm{OT}_p)$ gives us a powerful tool to immediately study continuity questions, as showcased in the following proof.

Proof. Let $f \in \mathcal{C}^0_{b,p}$ and $\mu, \mu_1, \mu_2 \cdots \in \mathcal{M}^p$ be such that $\mathrm{OT}_p(\mu_n, \mu) \to 0$. Following [Nie11, Theorem 2], it is sufficient to treat the case where f is real valued ($\mathcal{B} = \mathbb{R}$). Using Corollary 3.7, it means that $\mu_n^p \xrightarrow{w} \mu^p$, and thus that

$$\int_\Omega \frac{f(x)}{d(x, \partial\Omega)^p} \mathrm{d}\mu_n^p(x) \to \int_\Omega \frac{f(x)}{d(x, \partial\Omega)^p} \mathrm{d}\mu^p(x),$$

that is

$$\Phi(\mu_n) = \int_\Omega f(x) \mathrm{d}\mu_n(x) \to \int_\Omega f(x) \mathrm{d}\mu(x) = \Phi(\mu),$$

i.e. Φ is continuous with respect to OT_p.

Conversely, let $f : \Omega \to \mathcal{B}$. Assume first that f is not continuous at some $x \in \Omega$. There exists a sequence $(x_n)_n \in \Omega^{\mathbb{N}}$ such that $x_n \to x$ but $f(x_n) \nrightarrow f(x)$. Let $\mu_n = \delta_{x_n}$ and $\mu = \delta_x$. We have $\mathrm{OT}_p(\mu_n, \mu) \to 0$, but $\mu_n(f) = f(x_n) \nrightarrow f(x_0) = \mu(f)$, so that the linear representation $\mu \mapsto \mu(f)$ cannot be continuous.
Then, assume that f is continuous but that $x \mapsto \frac{f(x)}{d(x,\partial\Omega)^p}$ is not bounded. Let thus $(x_n)_n \in \Omega^{\mathbb{N}}$ be a sequence such that $\left\| \frac{f(x_n)}{d(x_n,\partial\Omega)^p} \right\| \to +\infty$. Define the measure $\mu_n := \frac{1}{\|f(x_n)\|} \delta_{x_n}$. Observe that $\mathrm{OT}_p(\mu_n, 0) = \frac{d(x_n,\partial\Omega)^p}{\|f(x_n)\|} \to 0$ by hypothesis. However, $\|\mu_n(f)\| = 1$ for all n, allowing to conclude once again that $\mu \mapsto \mu(f)$ cannot be continuous. □

Example 6.1. Let us give some examples of such linear representations (which are thus continuous) commonly used in applications of TDA. Note that the following definitions do not rely on the fact that the input must be a persistence diagram and actually make sense for any persistence measure

in $\mathcal{M}^p$. See Figure 6.1 for an illustration, in which computations are done with $p = 1$ (and $p' = 1$ for the weighted Betti curve).

- *Persistence surface and its variations.* Let $K : \mathbb{R}^2 \to \mathbb{R}$ be a nonnegative Lipschitz continuous bounded function (e.g. $K(x, y) = \exp\left(\frac{\|x-y\|^2}{2}\right)$) and define $f : x \in \Omega \mapsto d(x, \partial\Omega)^p \times K(x, \cdot)$, so that $f(x) : \mathbb{R}^2 \to \mathbb{R}$ is a real-valued function. The corresponding representation Φ takes its values in $(C_b(\mathbb{R}^2), \|\cdot\|_\infty)$, the (Banach) space of continuous bounded functions. This representation is called the persistence surface and has been introduced with slight variations in different works [AEK+17, CWRW15, KHF16, RHBK15].

- *Persistence silhouettes.* Let $\Lambda(x, t) = \max\left(\frac{x_1-x_2}{2} - \left|t - \frac{x_1+x_2}{2}\right|, 0\right)$ for $t \in \mathbb{R}$ and $x \in \Omega$. Then, defining $f : x \in \Omega \mapsto d(x, \partial\Omega)^{p-1} \times \Lambda(x, \cdot)$, one has that $\|f(x)\|_\infty$ is proportional to $d(x, \partial\Omega)^p$, so that the corresponding representation is continuous for OT_p. This representation is called the persistence silhouette, and was introduced in [CFL+14]. The corresponding Banach space is $(C_b(\mathbb{R}), \|\cdot\|_\infty)$.

- *Weighted Betti curves.* For $t \in \mathbb{R}$, define B_t the rectangle $(-\infty, t] \times [t, +\infty)$. Let $p, p' \geq 1$, and define $f : x \in \Omega \mapsto (t \mapsto d(x, \partial\Omega)^{p-1/p'} \mathbf{1}\{x \in B_t\})$. Then $f(x) \in L_{p'}(\mathbb{R})$ with $\|f(x)\|_{p'}$ proportional to $d(x, \partial\Omega)^p$. The corresponding function Φ is the weighted Betti curve, which takes its values in the Banach space $(L_{p'}(\mathbb{R}), \|\cdot\|_{p'})$. In particular, one obtains the continuity of the classical Betti curves from $(\mathcal{M}^1, \mathrm{OT}_1)$ to $L_1(\mathbb{R})$.

6.1.2 Stability in the case $p = 1$.

Continuity is a basic expectation when embedding a set of diagrams (or measures) in some Banach space $\mathcal{B}$. One could however ask for more, e.g. some Lipschitz regularity: given a representation $\Phi : \mathcal{M}^p \to \mathcal{B}$, one may want to have $\|\Phi(\mu) - \Phi(\nu)\| \leq C \cdot \mathrm{OT}_p(\mu, \nu)$ for some constant C. This property is generally referred to as "stability" in the TDA community and is generally obtained with $p = 1$, see for example [AEK+17, Theorem 2], [CCO17, Theorem 3.3 & 3.4], [STNR+18, §4], [RHBK15, Theorem 2], etc.

Here, we still consider the case of linear representations, and show that stability always holds with respect to the distance OT_1. Informally, this is explained by the fact that when $p = 1$, the cost function $(x, y) \mapsto d(x, y)^p$ is an actual distance.

Proposition 6.2. *Define $\mathcal{L}$ the set of Lipschitz continuous functions $f : \overline{\Omega} \to \mathbb{R}$ with Lipschitz constant less than 1 and that satisfy $f(\partial\Omega) = 0$. Let $T \subset \mathbb{R}$, and consider a family $(f_t)_{t\in T}$ with $f_t \in \mathcal{L}$. Then the linear representation $\Phi : \mu \mapsto (\mu(f_t))_{t\in T}$ is 1-Lipschitz continuous in the following sense:*

$$\|\Phi(\mu) - \Phi(\nu)\|_\infty := \sup_{t\in T} |(\mu - \nu)(f_t)| \leq \mathrm{OT}_1(\mu, \nu), \tag{6.2}$$

for any measures $\mu, \nu \in \mathcal{M}^1$.

Proof. Consider $\mu, \nu \in \mathcal{M}^1$, and $\pi \in \mathrm{Opt}(\mu, \nu)$ an optimal transport plan. Let $t \in T$. We have:

$$\begin{aligned}(\mu - \nu)(f_t) &= \int_\Omega f_t(x)\mathrm{d}\mu(x) - \int_\Omega f_t(y)\mathrm{d}\nu(y) = \iint_{\overline{\Omega}\times\overline{\Omega}} (f_t(x) - f_t(y))\mathrm{d}\pi(x, y) \\ &\leq \iint_{\overline{\Omega}\times\overline{\Omega}} d(x, y)\mathrm{d}\pi(x, y) = \mathrm{OT}_1(\mu, \nu),\end{aligned}$$

and thus, $\|\Phi(\mu) - \Phi(\nu)\|_\infty \leq \mathrm{OT}_1(\mu, \nu)$. □

Remark 6.1. As observed in Section 3.4, when $p = 1$, one actually has the Kantorovich-Rubinstein duality result $\mathrm{OT}_1(\mu, \nu) = \sup_f |\mu(f) - \nu(f)|$, where the supremum is taken over 1-Lipschitz functions f that satisfy $f(\partial\Omega) = 0$.

6.2 Learning representations of persistence diagrams using PersLay

In this section, as we target numerical applications, we restrict ourselves to finite persistence diagrams and will only consider linear representation valued in $\mathbb{R}^d$.

Linear representations are pretty handy: they are pretty simple to define and to compute. For instance, if $\mu = \sum n_x\delta_x$ is a persistence diagram, and $f : \Omega \to \mathbb{R}^d$, the representation $\mu(f)$ is simply given by $\sum n_x f(x) \in \mathbb{R}^d$. Therefore, given a set of persistence diagrams $\mu_1 \dots \mu_N$, one can recover a set $\mu_1(f) \dots \mu_N(f) \in \mathbb{R}^d$, on which basically every standard machine learning pipeline can be applied to solve a given task in a *supervised setting*. However, it might happen that the chosen function f is not adapted, or at least not optimal, to perform a given task; and there is no heuristic to chose f *a priori*.

6.2.1 Model

In [CCI+19], we considered the problem of optimizing the representation function f. Motivated by Section 6.1, we considered linear representations based on a function f of the form

$$x \mapsto w(x)\phi(x),$$

where $w : \mathbb{R}^2 \to \mathbb{R}$ is the *weight function* and $\phi : \mathbb{R}^2 \to \mathbb{R}^q$ is the *representation function*. The crucial point is that we consider *parameterized* functions $w = w_{\theta_1}$ and $\phi = \phi_{\theta_2}$, where $(\theta_1, \theta_2) \in \Theta \subset \mathbb{R}^D$ are *trainable* parameters that are learned during a training phase through a gradient descent.

Remark 6.2. There is some slight redundancy in using two functions w and ϕ. Formally, this helps to clarify how our formalism encompasses most vectorizations methods used in topological data analysis literature (also, it is implemented this way). In practice, it also has the benefits in terms of interpretability: the map $w : \mathbb{R}^2 \to \mathbb{R}$ can be easily visualized, giving a qualitative information on which areas of Ω are useful to achieve a good score in a given task (see Figure 6.4).

Remark 6.3. Proposition 6.2 ensures that if $x \mapsto w(x)\phi(x)/d(x, \partial\Omega)^p$ is bounded, the vectorization $\mu \mapsto \mu(w\phi)$ would be continuous with respect to the OT_p metric. In practice, this would enforce to have to small values of w on points close to the diagonal. However, in supervised learning applications, there is no obvious reason for this, so we actually allow for more general (trainable) weight functions w.

This flexible formalism allows us to recover and generalize standard vectorization methods used in topological data analysis. To that end we define our generic neural network layer for persistence diagrams, that we call PersLay. For a (finite) persistence diagram $\mu = \sum_{i=1}^{n} \delta_{x_i}$, it is defined through the following equation:

$$\text{PersLay}(\mu) := \mu(f) = \mathtt{op}\left(\{w(x_i) \cdot \phi(x_i)\}_{1 \leq i \leq n}\right), \tag{6.3}$$

where $\mathtt{op}$ is any *permutation invariant operator* (such as minimum, maximum, sum, kth largest value...), $w : \mathbb{R}^2 \to \mathbb{R}$ is a *weight function* for the persistence diagram points, and $\phi : \mathbb{R}^2 \to \mathbb{R}^q$ is a representation function that we call *point transformation*, mapping each point (α_b, α_d) of a persistence diagram to a vector.

Remark 6.4 (Relation to other works). Note that this problem is also addressed in the concurrent work [HKN19], where authors consider a similar framework. They enforce $w = d(\cdot, \partial\Omega)^p$, and introduce essentially three

point transformation functions ϕ, which roughly correspond to Gaussian, spike, and cone functions centered on the diagram points, and only consider `op = sum`. To that respect, our approach generalizes from their and allows our model to encompass a much larger class of vectorization methods used in practice. Furthermore, we focus on providing a publicly available implementation as a module of the `Gudhi` library, also hosted at `https://github.com/MathieuCarriere/perslay`.

It is also important to note that both [CCI+19] and [HKN19] took some of their inspiration from [ZKR+17], introducing the notion of *Deep Sets*. In this work, authors proved that given a (countable, or compact) set X, a map $\Phi : X^n \to \mathbb{R}$ is *permutation invariant*—that is, it satisfies $\Phi(x_1 \dots x_n) = \Phi(x_{\sigma(1)} \dots x_{\sigma(n)})$ for any $(x_1 \dots x_n) \in \mathbb{X}^n$ and any permutation $\sigma \in \mathfrak{S}_n$—if and only if it exists a map $\rho : X \to \mathbb{R}$ such that $\Phi(x_1 \dots x_n) = \sum_{i=1}^{n} \rho(x_i)$. This expression is consistent with the one proposed by [HKN19] and with (6.3).

As explained above, in practice w and ϕ are of the form $w_{\theta_1}, \phi_{\theta_2}$ where the gradients of $\theta_1 \mapsto w_{\theta_1}$ and $\theta_2 \mapsto \phi_{\theta_2}$ are known and implemented so that back-propagation can be performed, and the parameters θ_1, θ_2 can be optimized during the training process. We emphasize that any neural network architecture ρ can be composed with PersLay to generate a neural network architecture for persistence diagrams. Let us now introduce three point transformation functions that we use and implement for parameter ϕ in (6.3).

- The *triangle point transformation*

$$\phi_\Lambda : \mathbb{R}^2 \to \mathbb{R}^q, x \mapsto \left[\Lambda_x(t_1), \Lambda_x(t_2), \dots, \Lambda_x(t_q)\right]^T,$$

  where the triangle function Λ_x associated to a point $x = (b, d) \in \mathbb{R}^2$ is $\Lambda_x : t \mapsto \max\{0, d - |t - b|\}$, with $q \in \mathbb{N}$ and $t_1, \dots, t_q \in \mathbb{R}$.

- The *Gaussian point transformation*

$$\phi_\Gamma : \mathbb{R}^2 \to \mathbb{R}^q, x \mapsto \left[\Gamma_x(t_1), \Gamma_x(t_2), \dots, \Gamma_x(t_q)\right]^T,$$

  where the Gaussian function Γ_x associated to a point $x = (b, d) \in \mathbb{R}^2$ is $\Gamma_x : t \mapsto \exp\left(-\|x - t\|_2^2/(2\sigma^2)\right)$ for a given $\sigma > 0$, $q \in \mathbb{N}$ and $t_1, \dots, t_q \in \mathbb{R}^2$.

- The *line point transformation*

$$\phi_L : \mathbb{R}^2 \to \mathbb{R}^q, x \mapsto \left[L_{\Delta_1}(x), L_{\Delta_2}(x), \dots, L_{\Delta_q}(x)\right]^T,$$

where the line function L_Δ associated to a line Δ with direction vector $e_\Delta \in \mathbb{R}^2$ and bias $b_\Delta \in \mathbb{R}$ is $L_\Delta : x \mapsto \langle x, e_\Delta \rangle + b_\Delta$, with $q \in \mathbb{N}$ and $\Delta_1, \ldots, \Delta_q$ are q lines in the plane.

Formulation (6.3) is very general: despite its simplicity, it allows us to remarkably encode most persistence diagram vectorizations with a small set of point transformation functions ϕ, allowing to consider the choice of ϕ as a hyperparameter of sort. Let us show how it connects to most of the popular vectorizations and kernel methods for persistence diagrams in the literature.

- Using: $\phi = \phi_\Lambda$ with samples $t_1, \ldots, t_q \in \mathbb{R}$, $\mathtt{op} = k$th largest value, $w = 1$ (a constant weight function), amounts to evaluating the kth *persistence landscape* [Bub15] on $t_1, \ldots, t_q \in \mathbb{R}$.

- Using $\phi = \phi_\Lambda$ with samples $t_1, \ldots, t_q \in \mathbb{R}$, $\mathtt{op} = \mathrm{sum}$, arbitrary weight function w, amounts to evaluating the *persistence silhouette* weighted by w [CFL+14] on $t_1, \ldots, t_q \in \mathbb{R}$.

- Using $\phi = \phi_\Gamma$ with samples $t_1, \ldots, t_q \in \mathbb{R}^2$, $\mathtt{op} = \mathrm{sum}$, arbitrary weight function w, amounts to evaluating the *persistence surface* weighted by w [AEK+17] on $t_1, \ldots, t_q \in \mathbb{R}^2$. Moreover, characterizing points of persistence diagrams with Gaussian functions is also the approach advocated in several kernel methods for persistence diagrams [RHBK15, LY18, KHF16].

- Using $\phi = \phi_{\tilde{\Gamma}}$ where $\tilde{\Gamma}$ is a modification of the Gaussian point transformation defined with: $\tilde{\Gamma}_p = \Gamma_{\tilde{p}}$ for any $p = (x, y) \in \mathbb{R}^2$, where $\tilde{p} = p$ if $y \leq \nu$ for some $\nu > 0$, and $\left(x, \nu + \log\left(\frac{y}{\nu}\right)\right)$ otherwise, $\mathtt{op} = \mathrm{sum}$, weight function $w = 1$, is the approach presented in [HKNU17].

- Using $\phi = \phi_L$ with lines $\Delta_1, \ldots, \Delta_q \in \mathbb{R}^2$, $\mathtt{op} = k$th largest value, weight function $w = 1$, is similar to the approach advocated in [CCO17], where the sorted projections of the points onto the lines are then compared with the $\|\cdot\|_1$ norm and exponentiated to build the so-called Sliced Wasserstein kernel for persistence diagrams.

6.2.2 Numerical experiments

We describe here some numerical experiments that were run in [CCI+19, §3.2, §4], and refer the reader to Appendix C therein for complementary details (e.g. report of the hyperparameters used). The accuracies reported

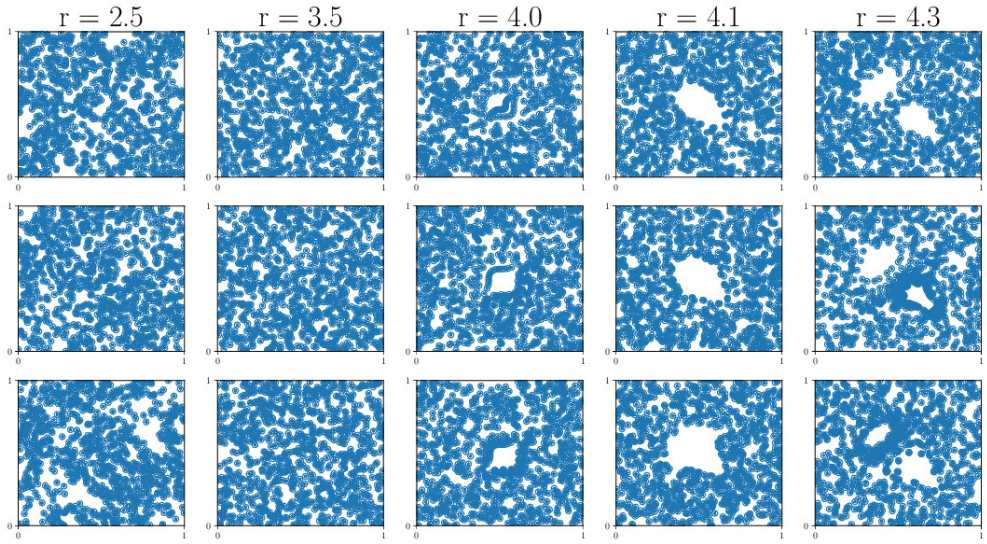


Figure 6.2: Some examples of orbits generated by the different choices of r (three simulations are represented for the different values of r).

in Table 6.1 and Table 6.2 represent a snapshot of the performances of some state-of-the-art methods at the time [CCI+19] was submitted.

Orbit classification. Our first application is on a synthetic dataset used as a benchmark in topological data analysis [LY18, CCO17, AEK+17]. It consists of sequences of points generated by different dynamical systems (see [HSW07]). Given some initial position $(x_0, y_0) \in [0,1]^2$ and a parameter $r > 0$, we generate a point cloud $(x_n, y_n)_{n=1,\dots,N}$ following:

$$\begin{cases} x_{n+1} = x_n + r y_n(1 - y_n) & \mod 1, \\ y_{n+1} = y_n + r x_{n+1}(1 - x_{n+1}) & \mod 1. \end{cases} \tag{6.4}$$

The orbits of this dynamical system heavily depend on parameter r (see Figure 6.2). More precisely, for some values of r, voids might form in these orbits, and as such, persistence diagrams are likely to perform well at attempting to classify orbits with respect to the value of r generating them. As in previous works [LY18, CCO17, AEK+17], we use the five different parameters $r = 2.5, 3.5, 4.0, 4.1$ and 4.3 to simulate the different classes of orbits, with random initialization of (x_0, y_0) and $N = 1,000$ points in each simulated orbit. These point clouds are then turned into persistence diagrams using a standard geometric filtration [CDSO14], called the `AlphaComplex`

Dataset	PSS-K	PWG-K	SW-K	PF-K	PersLay
ORBIT5K	72.38(±2.4)	76.63(±0.7)	83.6(±0.9)	85.9(±0.8)	**87.7(±1.0)**
ORBIT100K	—	—	—	—	**89.2(±0.3)**

Table 6.1: Performance table. PSS-K, PWG-K, SW-K, PF-K stand for *Persistence Scale Space Kernel* [RHBK15], *Persistence Weighted Gaussian Kernel* [KHF16], *Sliced Wasserstein Kernel* [CCO17] and *Persistence Fisher Kernel* [LY18] respectively. We report the scores given in [LY18] for competitors on ORBIT5K, and the one we obtained using PersLay for both the ORBIT5K and ORBIT100K datasets. The latter dataset, and the corresponding diagrams, are too large in practice to be handled by kernel methods on persistence diagrams.

filtration[1] in dimensions 0 and 1. We generate two datasets: The first is ORBIT5K, where for each value of r, we generate 1000 orbits, ending up with a dataset of $5,000$ point clouds. This dataset is the same as the one used in [LY18]. The second is ORBIT100K, which contains $20,000$ orbits per class, resulting in a dataset of $100,000$ point clouds—a scale that kernel for persistence diagrams cannot handle. This dataset aims to show the edge of our neural-network based approach over kernels methods when dealing with very large datasets of large diagrams, since all the previous works dealing with this data [LY18, CCO17, AEK+17] use kernel methods. Results are displayed in Table 6.1. As it can be observed, we improve on previous results for ORBIT5K, we also show with ORBIT100K that classification accuracy is further increased as more observations are made available. For consistency we use the same accuracy metric as [LY18], that is, we split observations in 70%-30% training-test sets and report the average test accuracy over 100 runs.

Graph classification. The second experiment we used to illustrate the use of PersLay as an interface to incorporate persistence diagrams in learning pipeline is a *graph classification* task. Each graphs in a given dataset was turned into a set of persistence diagrams using the so-called *Heat Kernel Signatures* (HKS) of the graph. For the sake of concision, we do not give a detailed presentation of the HKS signatures here, and refer the reader to the original paper for complementary information.

We used a series of different graph datasets commonly used as a baseline in graph classification problems.

[1] http://gudhi.gforge.inria.fr/python/latest/alpha_complex_ref.html

- `REDDIT5K, REDDIT12K, COLLAB` (from [YV15]) `IMDB-B, IMDB-M` (from [TVH18]) are composed of social graphs.
- `COX2, DHFR, MUTAG, PROTEINS, NCI1, NCI109` are graphs coming from medical or biological frameworks (also from [TVH18]).

We compared performances with five other top graph classification methods.

- Scale-variant topo [TVH18] leverages a kernel for ordinary persistence diagrams computed on point cloud used to encode the graphs.
- RetGK [ZWX+18] is a kernel method for graphs that leverages eventual *attributes* on the graph vertices and edges.
- FGSD [VZ17] is a finite-dimensional graph embedding that does not leverage attributes.
- Finally, GCNN [XC19] and GIN [XHLJ19] are two graph neural network approaches that reach top-tier results.

One could also compare our results on the `REDDIT` datasets to the ones of [HKNU17], where authors also use persistence diagrams to feed a network (using as first channel a particular case of PersLay), achieving 54.5% and 44.5% of accuracy on `REDDIT5K` and `REDDIT12K` respectively.

In this experiment, we used a voluntarily simple network architecture, namely a two-layer network. The first layer is PersLay, which processes persistence diagrams. The resulting vector is normalized and fed to the second and final layer, a fully-connected layer whose output is used for predictions. See Figure 6.3 for an illustration. We emphasize that this simplistic two-layer architecture is designed so as to produce understanding rather than achieving the best possible performances.

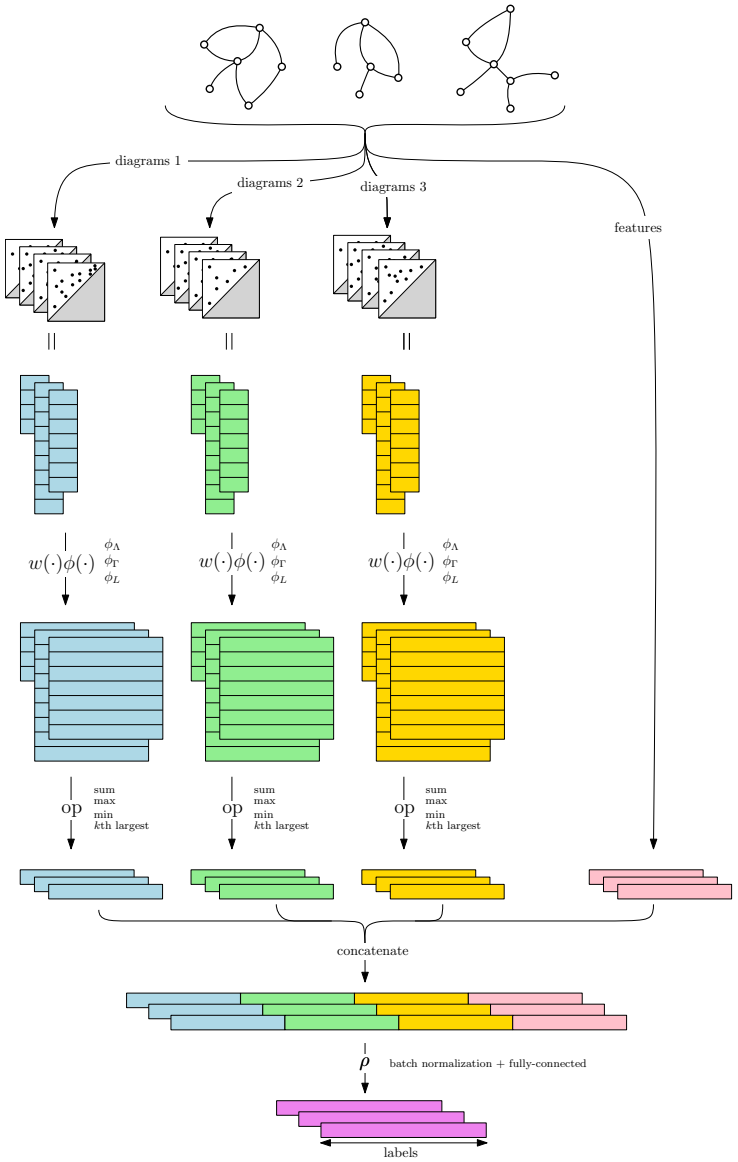


Figure 6.3: Network architecture illustrated in the case of our graph classification experiments. Each graph is encoded as a set of 4 persistence diagrams, then processed by an independent instance of PersLay. Each instance embeds diagrams in some vector space using two functions w, ϕ that are optimized during training and a fixed permutation-invariant operator `op`.

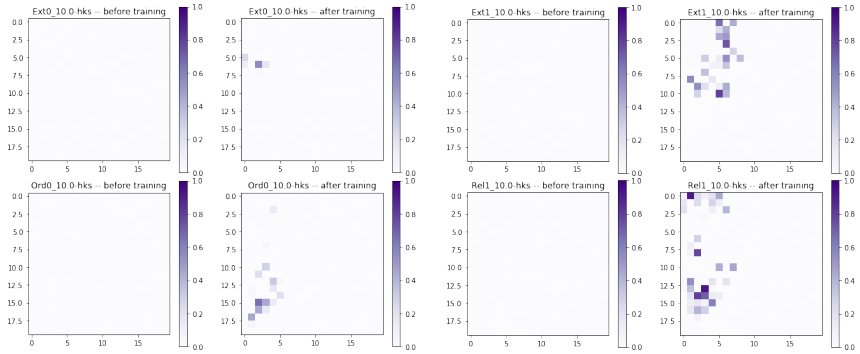


Figure 6.4: Weight function w when chosen to be a grid with size 20×20 before and after training (`MUTAG` dataset). Here, Ord0, Rel1, Ext0, and Ext1 denote the different diagrams extracted from a graphs, each of them being processed by an independent instance of PersLay.

Dataset	SV[1]	RetGK* [2]	FGSD [3]	GCNN [4]	GIN [5]	PersLay	
						Mean	Max
REDDIT5K	—	56.1	47.8	52.9	57.0	55.6	56.5
REDDIT12K	—	48.7	—	46.6	—	47.7	49.1
COLLAB	—	81.0	80.0	79.6	80.1	76.4	78.0
IMDB-B	72.9	71.9	73.6	73.1	74.3	71.2	72.6
IMDB-M	50.3	47.7	52.4	50.3	52.1	48.8	52.2
COX2*	78.4	80.1	—	—	—	80.9	81.6
DHFR*	78.4	81.5	—	—	—	80.3	80.9
MUTAG*	88.3	90.3	92.1	86.7	89.0	89.8	91.5
PROTEINS*	72.6	75.8	73.4	76.3	75.9	74.8	75.9
NCI1*	71.6	84.5	79.8	78.4	82.7	73.5	74.0
NCI109*	70.5	—	78.8	—	—	69.5	70.1

Table 6.2: Classification accuracy over benchmark graph datasets. Our results (PersLay, right hand side) are recorded from ten runs of a 10-fold classification evaluation. "Mean" is consistent with [ZWX+18][2], while "Max" should be compared to [TVH18][1], [VZ17][3], [XC19][4] and [XHLJ19][5], as it corresponds to the mean accuracy over a *single 10-fold*. The * indicates datasets that contain attributes (labels) on graph nodes and symmetrically the methods that leverage such attributes for classification purposes.

Chapter 7

Complementary examples

Abstract

This final chapter gathers three independent examples that showcase how adopting an optimal transport perspective and a measure-based formalism is helpful when dealing with persistence diagrams. Section 7.1 is dedicated to the study of *expected persistence diagrams*, that arise when considering persistence diagrams in a random setting. As those are measures with continuous support, they can be manipulated using our formalism, where standard definitions of diagram metrics would be insufficient. Section 7.2 studies the *quantization* of persistence diagrams, namely approximating a (think, large) diagram by a persistence measure with a chosen size support. Finally, Section 7.3 addresses the question of estimating *shift-invariant* distances between persistence diagrams. Using the *entropic regularization* of optimal transport (adapted to persistence diagrams), we obtain a simple and fast algorithm (although non-convex) to propose solutions to this problem.

Implementation resources.

- A tutorial for Expected persistence diagrams.
- (To appear in `Gudhi`) Quantization of persistence diagrams.
- (In progress) Shift-invariant distance.

7.1 Expected persistence diagrams and their estimators

Recall that $\mathcal{M}^p$ denotes the space of persistence measures (Radon measures supported on the open half-plane Ω) for some parameter $1 \leq p < \infty$, and $\mathcal{D}^p \subset \mathcal{M}^p$ denotes the subspace of persistence diagrams. These spaces are equipped with metric OT_p. These are Polish metric spaces (see Chapter 3) and therefore, one can consider probability distribution supported on these.

When observing a sample of persistence diagrams $\mu_1 \dots \mu_K \in \mathcal{D}^p$, a natural model is to assume that the μ_is are sampled i.i.d. with respect to some underlying probability distribution P belonging to $\mathcal{W}^p(\mathcal{D}^p)$, the Wasserstein space on $\mathcal{D}^p$ (see Section 2.2 and Chapter 3 for details). This section is dedicated to the study of a natural statistical descriptor that one can build on P: its (linear) expectation, denoted by $\mathbb{E}_P[\boldsymbol{\mu}]$ where $\boldsymbol{\mu} \sim P$, and called its *expected persistence diagram* (EPD). We start by proving a stability result on the map $P \mapsto \mathbb{E}_P[\boldsymbol{\mu}]$ (this result was part of [DL19, §5]). As in practice one only has access to a finite sample $\mu_1 \dots \mu_K \sim P$, we will study how the empirical estimator $\frac{1}{K}\sum \mu_i$ performs when it comes to model the underlying object $\mathbb{E}_P[\boldsymbol{\mu}]$.

7.1.1 Definition and properties

Let P be a probability distribution of persistence diagrams. One may want to build a statistic on P, such as defining and computing its mean. As $(\mathcal{D}^p, \mathrm{OT}_p)$ is not a vector space but only a metric space, the notion of mean is ill-defined. A natural attempt is to consider *Fréchet means*, that is minimizers of the functional $\nu \mapsto \int_{\mu \in \mathcal{D}^p} \mathrm{OT}_p^p(\nu, \mu) \mathrm{d}P(\mu)$. Algorithms to address this minimization problem have been proposed (see Chapter 5 and references therein) but those only provide an estimate and can be expensive to compute. On the other hand, one could go for a much simpler approach and define a measure $\mathbb{E}_P[\boldsymbol{\mu}]$ supported on Ω by stating, for $A \subset \Omega$ compact,

$$\mathbb{E}_P[\boldsymbol{\mu}](A) := \mathbb{E}_P[\boldsymbol{\mu}(A)], \tag{7.1}$$

where $\boldsymbol{\mu} \sim P$, and $\boldsymbol{\mu}(A)$ is the (random) number of points of $\boldsymbol{\mu}$ that belongs to A. This deterministic measure, called the *expected persistence diagram* of P, was introduced in [CD18]. In this work, Chazal and Divol proved in particular that under mild assumptions on P, the corresponding EPD admits a density with respect to the Lebesgue measure on Ω. Therefore, it is an element of $\mathcal{M}^p$, although not being a persistence diagram.

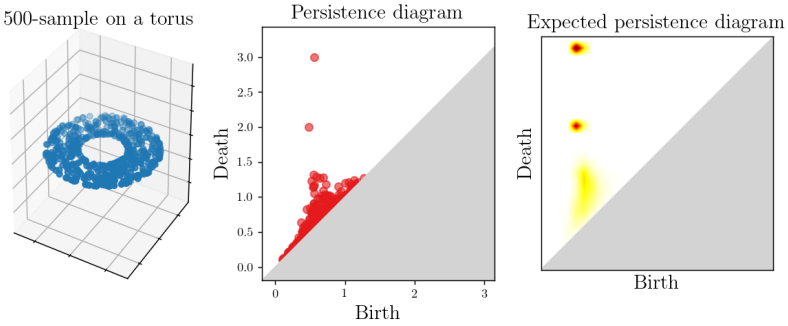


Figure 7.1: *(Left)* A n-sample on a torus, with $n = 500$. *(Middle)* The corresponding persistence diagram. *(Right)* The corresponding (empirical) expected persistence diagram, estimated by sampling $K = 100$ times n points on a torus. Note that for illustration purpose, the empirical EPD has been rendered using a 30×30 grid, and masses of each bin have been weighted according to their distance to the diagonal, to the power $p = 2$.

Intuitively, the mass $\mathbb{E}_P[\boldsymbol{\mu}]$ gives to a region $A \subset \Omega$ is simply the average number of points that falls in A when sampling diagrams with respect to P. See Figure 7.1 for an illustration.

Remark 7.1. A well-suited definition of the linear expectation in (7.1) requires technical care (basically, turning the finite sum into a Bochner integral) and is detailed in Section 7.1.3. This is required, in particular, to apply the Jensen inequality in our proof of the stability result. Nonetheless, on first read one can treat $\mathbb{E}_P[\boldsymbol{\mu}]$ as a standard expectation and take Equation (7.1) as a definition.

The main theoretical result of this section states the stability of the map

$$\begin{aligned} \mathcal{W}^p(\mathcal{D}^p) &\to \mathcal{M}^p \\ P &\mapsto \mathbb{E}_P[\boldsymbol{\mu}]. \end{aligned} \tag{7.2}$$

Proposition 7.1. *Let $P, P' \in \mathcal{W}^p(\mathcal{M}^p)$. We have*

$$\mathrm{OT}_p(\mathbb{E}_P[\boldsymbol{\mu}], \mathbb{E}_{P'}[\boldsymbol{\mu}]) \leq W_{p,\mathrm{OT}_p}(P, P').$$

The proof can be found in Section 7.1.3.

Using stability results on the distances $\mathbf{d}_p$ ($= \mathrm{OT}_p$) between persistence diagrams (see Theorem 2.2 or [CSEHM10]), one is able to obtain a more precise control between the expectations in some situations. For $\mathbb{Y}$ a sample in some metric space, denote by $\mathrm{Dgm}(\mathbb{Y})$ the persistence diagram of $\mathbb{Y}$ built with the Čech filtration.

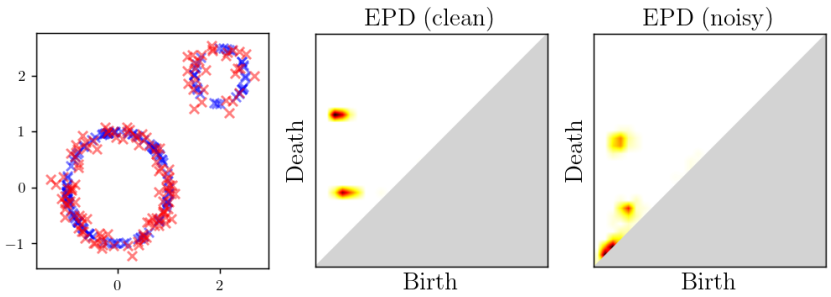


Figure 7.2: Just as standard diagrams, expected persistence diagrams are stable with respect to small perturbations. *(Left)* Two point clouds (with $n = 130$ points). Blue (dot) is (randomly, iid) sampled exactly on the union of two circles, while red (cross) is sampled on the union of these two circles in addition to some small bounded noise. *(Middle and Right* The (empirical) expected persistence diagram (EPD) for the "clean" (blue) sampling process and for the "noisy" (red) sampling process respectively. Due to the noise, the second EPD has some mass close to the diagonal; but the two EPDs remain close in terms of OT_∞ metric.

Proposition 7.2. *Let ξ, ξ' be two probability measures on $\mathbb{R}^d$. Let $\mathbb{X}_n$ (resp. $\mathbb{X}'_n$) be a n-sample of law ξ (resp. ξ'). Then, for any $k > d$, and any $p \geq k+1$,*

$$\mathrm{OT}_p^p(\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)], \mathbb{E}[\mathrm{Dgm}(\mathbb{X}'_n)]) \leq C_{k,d} \cdot n \cdot W_{p-k}^{p-k}(\xi, \xi') \tag{7.3}$$

where $C_{k,d} := C\mathrm{diam}(\mathbb{X})^{k-d}\frac{k}{k-d}$ for some constant C depending only on $\mathbb{X}$.

In particular, letting $p \to \infty$, we obtain a bottleneck stability result:

$$\mathrm{OT}_\infty(\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)], \mathbb{E}[\mathrm{Dgm}(\mathbb{X}'_n)]) \leq W_\infty(\xi, \xi'). \tag{7.4}$$

Proof. Let γ be any coupling between $\mathbb{X}_n$ a n-sample of law ξ, and $\mathbb{X}'_n$ a n-sample of law ξ'. According to Proposition 7.1,

$$\mathrm{OT}_p^p(\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)], \mathbb{E}[\mathrm{Dgm}(\mathbb{X}'_n)]) \leq \mathbb{E}_\gamma[\mathrm{OT}_p^p(\mathrm{Dgm}(\mathbb{X}_n), \mathrm{Dgm}(\mathbb{X}'_n))].$$

Theorem 2.2 states (see also [CSEHM10, Wasserstein Stability Theorem]) that

$$\mathrm{OT}_p^p(\mathrm{Dgm}(\mathbb{X}_n), \mathrm{Dgm}(\mathbb{X}'_n)) \leq C_{k,d} H(\mathbb{X}_n, \mathbb{X}'_n)^{p-k},$$

where $C_{k,d} := C\mathrm{diam}(\mathbb{X})^{k-d}\frac{k}{k-d}$ for some constant C depending only on $\mathbb{X}$, and H is the Hausdorff distance between sets. By taking the infimum on transport plans γ, we obtain

$$\mathrm{OT}_p^p(\mathbb{E}[\mathrm{Dgm}(\mathbb{X}_n)], \mathbb{E}[\mathrm{Dgm}(\mathbb{X}'_n)]) \leq C_{k,d} W_{H,p-k}^{p-k}(\xi^{\otimes n}, (\xi')^{\otimes n}),$$

where $W_{H,p}$ is the p-Wasserstein distance between probability distributions on compact sets of the manifold $\mathbb{X}$, endowed with the Hausdorff distance. Lemma 15 of [CFL+15] states that

$$W_{H,p-k}^{p-k}(\xi^{\otimes n}, (\xi')^{\otimes n}) \leq n \cdot W_{p-k}^{p-k}(\xi, \xi'),$$

concluding the proof. □

Note that this proposition illustrates the usefulness of introducing new distances OT_p: considering the proximity between linear expectations requires to extend the metrics $\mathbf{d}_p$ to Radon measures.

7.1.2 Approximating EPDs with empirical EPDs.

As in practice, one generally only have access to a sample of observations instead of the whole underlying law P and thus the corresponding expected persistence diagram $\mathbb{E}_P$, we now address the question of estimating the EPD by its empirical counterpart.

Let P be a probability measure on $(\mathcal{D}_p, \mathrm{OT}_p)$ and let $\boldsymbol{\mu}^{(1)}, \ldots, \boldsymbol{\mu}^{(K)}$ be a K-sample[1] of law P. The empirical EPD is defined as

$$\overline{\boldsymbol{\mu}}_{\mathbf{K}} := \frac{\boldsymbol{\mu}^{(1)} + \cdots + \boldsymbol{\mu}^{(K)}}{K}. \tag{7.5}$$

We first study the convergence of $\overline{\boldsymbol{\mu}}_{\mathbf{K}}$ to $\mathbb{E}_P[\boldsymbol{\mu}]$ as $K \to \infty$.

Proposition 7.3. *Let $p \geq 1$ and assume that $\mathbb{E}_P[\mathrm{Pers}_p(\boldsymbol{\mu})] < \infty$. Then, $\mathbb{E}[\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}])] \to 0$.*

Proof. We use the characterization of convergence that we proved in Theorem 3.5. It states that the convergence $\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}]) \to 0$ holds almost surely if and only if (i) $\mathrm{Pers}_p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}) \to \mathrm{Pers}_p(\mathbb{E}_P[\boldsymbol{\mu}])$ a.s. and (ii) $\overline{\boldsymbol{\mu}}_{\mathbf{K}}$ converges vaguely to $\mathbb{E}_P[\boldsymbol{\mu}]$ a.s., that is if $\overline{\boldsymbol{\mu}}_{\mathbf{K}}(\phi) \to \mathbb{E}_P[\boldsymbol{\mu}](\phi)$ a.s., for all $\phi : \Omega \to \mathbb{R}$ continuous with compact support [Kal17, Lemma 4.8]. Those two conditions hold by the strong law of large numbers. Thus, we have $\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}]) \to 0$ a.s. To obtain the convergence of the expectation, we also need the sequence of random variables $(\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}]))_K$ to be uniformly integrable. By the triangle inequality, and using that

[1] That is, K measures sampled i.i.d.

$(a+b)^p \leq 2^{p-1}(a^p + b^p)$ for $a, b \geq 0$, with $\mathbf{0}$ denoting the empty diagram,

$$\begin{aligned}\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}]) &\leq (\mathrm{OT}_p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbf{0}) + \mathrm{OT}_p(\mathbb{E}_P[\boldsymbol{\mu}], \mathbf{0}))^p \\ &\leq 2^{p-1}(\mathrm{Pers}_p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}) + \mathrm{Pers}_p(\mathbb{E}_P[\boldsymbol{\mu}])) \\ &= 2^{p-1}\frac{1}{K}\sum_{k=1}^{K}\mathrm{Pers}_p(\mu^{(\mathbf{k})}) + \mathbb{E}_P[\mathrm{Pers}_p(\boldsymbol{\mu})].\end{aligned}$$

This last quantity is uniformly integrable, and we can conclude as almost sure convergence and uniform integrability imply convergence in expectation. □

We can make this convergence result more quantitative by making more assumptions on the probability P, namely that the diagrams sampled according to P have a uniformly bounded mass (number of points) and support.

Proposition 7.4. *Consider $p \geq 1$ and let P be a probability measure on $(\mathcal{D}_p, \mathrm{OT}_p)$ such that $\boldsymbol{\mu} \sim P$ is a.s. supported on $[-L, L]^2$ and of total mass smaller than M. Then, for any $K > 0$,*

$$\mathbb{E}[\mathrm{OT}_p^p(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}])] \leq \frac{C_p M L^p}{\sqrt{K}} a_p(K). \tag{7.6}$$

with $a_p(K) = 1$ if $p > 1$ and $a_1(K) = \ln(K)$.

The proof, slightly technical, is delayed to Section 7.1.3.

Remark 7.2. Using the dual formulation (Section 3.4) for $p = 1$, the quantity $\mathrm{OT}_1(\mu, \nu)$ is equal to $\sup_f |\mu(f) - \nu(f)|$, where the supremum is taken on all Lipschitz continuous functions $f : \overline{\Omega} \to \mathbb{R}$ with $f(\partial\Omega) = 0$. Hence, Proposition 7.4 implies that

$$\mathbb{E}[\sup_f |\overline{\boldsymbol{\mu}}_{\mathbf{K}}(f) - \mathbb{E}_P[\boldsymbol{\mu}](f)|] \leq CML\frac{\ln(K)}{\sqrt{K}}. \tag{7.7}$$

Taking different choices of f induce different representations on PDs, including e.g. persistent surfaces [AEK+17], persistent silhouettes [CFL+14], the persistence scale-space kernel [RHBK15] or gaussian functions centered at random locations [HKNU17]. Hence, Proposition 7.4 gives a uniform rate of convergence for all those representations.

Let $\mathcal{M}_{M,L}$ be the set of persistence measures with support included in $[-L, L]^2$ and of mass smaller than M, and let $\mathcal{P}_{M,L}$ be the set of all

probability distributions supported on the set $\mathcal{D} \cap \mathcal{M}_{M,L}$. The following proposition shows that the minimax rate for the estimation of the EPD given K observations of law $P \in \mathcal{P}_{M,L}$ is of order $1/\sqrt{K}$, implying that the empirical EPD is a minimax estimator of the EPD. The notation $a_K \asymp b$ means that there exists a constant C such that $\forall K,\ a_K/b \leq C$ and $b/a_K \leq C$.

Proposition 7.5. *Let $1 \leq p < \infty$. One has*

$$\inf_{\hat{\mu}_{\mathbf{K}}} \sup_{P \in \mathcal{P}_{M,L}} \mathbb{E}[\mathrm{OT}_p^p(\hat{\mu}_{\mathbf{K}}, \mathbb{E}_P[\mu])] \asymp \frac{ML^p}{\sqrt{K}}, \tag{7.8}$$

where the infimum is taken on all measurable $\mathcal{M}_{M,L}$-valued functions $\hat{\mu}_{\mathbf{K}}$ of a K-sample $\mu^{(\mathbf{1})}, \ldots, \mu^{(\mathbf{K})}$ of law $P \in \mathcal{P}_{M,L}$.

7.1.3 Complementary details and delayed proofs

Formal definition of the expected persistence diagram. Define $\mathcal{M}_\pm$ the space of signed measures on Ω, i.e. a measure $\mu \in \mathcal{M}_\pm$ is written $\mu_+ - \mu_-$ for two finite measures $\mu_+, \mu_- \in \mathcal{M}_f$. The total variation distance $|\cdot|$ is a norm on $\mathcal{M}_\pm$, and $(\mathcal{M}_\pm, |\cdot|)$ is a Banach space. The Bochner integral [Boc33] is a generalization of the Lebesgue integral for functions taking their values in Banach space. We define the expected persistence measure of $P \in \mathcal{W}^p(\mathcal{M}^p)$ as the Bochner integral of some pushforward of P. More precisely, define

$$\begin{aligned} F : (\mathcal{M}^p, \mathrm{OT}_p) &\to (\mathcal{M}_\pm, |\cdot|) \\ \mu &\mapsto \mu^p. \end{aligned}$$

Note that F has an inverse G on $\mathcal{M}_f$, defined by $G(\nu)(f) := \int_\Omega \frac{f(x)}{d(x,\Omega)^p} \mathrm{d}\nu(x)$. Theorem 3.5 implies that G is a continuous function from $(\mathcal{M}_f, |\cdot|)$ to $(\mathcal{M}^p, \mathrm{OT}_p)$. In particular, as $\mathcal{M}_f$ and $\mathcal{M}^p$ are Polish spaces, F is measurable (see [Kec95, Theorem 15.1]). For $P \in \mathcal{W}^p(\mathcal{M}^p(\Omega))$, define for $\boldsymbol{\mu} \sim P$, $\mathbb{E}[\boldsymbol{\mu}]$ the *linear* expectation of P by

$$\mathbb{E}[\boldsymbol{\mu}] := G\left(\int \nu \mathrm{d}(F_\# P)(\nu)\right) \in \mathcal{M}^p, \tag{7.9}$$

where the integral is the Bochner integral on the Banach space $(\mathcal{M}_\pm, |\cdot|)$ and $F_\# P$ is the pushforward of P by F. It is straightforward to check that $\mathbb{E}[\boldsymbol{\mu}]$ defined in that way satisfies the relation

$$\forall K \subset \Omega \text{ compact},\ \mathbb{E}[\boldsymbol{\mu}](K) = \mathbb{E}[\boldsymbol{\mu}(K)].$$

Delayed proofs.

Proof of Proposition 7.1. The proof consists in applying Jensen's inequality in an infinite-dimensional setting. We first recall that the function OT_p^p is convex (Lemma 3.9). We then use the following general result:

Proposition 7.6 ([Z+00])**.** *Let $(V, \|\cdot\|)$ be a separable Banach space and $\mathcal{X}$ a closed convex subset of V with non-empty interior. Let P be a probability measure on $\mathcal{X}$ endowed with its Borelian σ-algebra. Assume that $\int \|x\| \mathrm{d}P(x) < \infty$. Let $f : \mathcal{X} \to \mathbb{R}$ be a continuous convex function so that $\int f(x)\mathrm{d}P(x) < \infty$. Then,*

$$f\left(\int x \mathrm{d}P(x)\right) \leq \int f(x)\mathrm{d}P(x).$$

Now, using the notations introduced above, take $V = \mathcal{M}_\pm \times \mathcal{M}_\pm$, $\mathcal{X} = \mathcal{M}_f \times \mathcal{M}_f$ and $f = \mathrm{OT}_p^p \circ (G, G) : \mathcal{X} \to \mathbb{R}$. The continuity of G implies that f is continuous and convexity of OT_p^p implies the convexity of f. Let now P, P' be two probability measures supported on $\mathcal{W}^p(\mathcal{M}^p)$, and γ be any coupling between P and P'. Proposition 7.6 yields

$$\mathrm{OT}_p^p(\mathbb{E}_P(\boldsymbol{\mu}), \mathbb{E}_{P'}(\boldsymbol{\mu})) \leq \iint \mathrm{OT}_p^p(\mu, \mu')\mathrm{d}\gamma(\mu, \mu'),$$

and taking the infimum over γ gives the conclusion. □

Proof of Proposition 7.4. The total variation between measures will be denoted by $|\cdot|$ in the proof. By using the triangle inequality,

$$\begin{aligned}\mathrm{OT}_p(\mu, \nu) &\leq \mathrm{OT}_p\left(\frac{|\nu|}{|\mu|}\mu, \mu\right) + \mathrm{OT}_p\left(\frac{|\nu|}{|\mu|}\mu, \nu\right) \\ &\leq |\mu - \nu|^{1/p} L + W_{p,\|\cdot\|}\left(\frac{|\nu|}{|\mu|}\mu, \nu\right),\end{aligned}$$

where at the second line, the first term is bounded by considering the transport plan which consists in taking the identity map on Ω, and by mapping the remaining mass on the diagonal, with a cost therefore bounded by $|\mu - \nu| L^p$; and the second term is bounded by observing that any off-diagonal transport plan π supported on $\Omega \times \Omega$ between $\frac{|\nu|}{|\mu|}\mu$ and ν (which do have the same total mass) is an admissible (partial) plan between these two persistence measures.

Let $\mathcal{P}_l$ be the natural partition of $[-L, L)^2$ induced by a $2^L \times 2^L$ regular grid. For any μ, ν supported on $[-L, L)^2$, we let

$$D_p(\mu, \nu) := \sum_{l \geq 1} 2^{-pl} \sum_{F \in \mathcal{P}_l} |\mu(F) - \nu(F)|. \tag{7.10}$$

This quantity was introduced by [FG15], who proved in particular that if two measures σ, τ have the same total mass, then there exists a constant κ_p such that $W^p_{p,\|\cdot\|}(\sigma,\tau) \leq \kappa_p L^p D_p(\sigma,\tau)$. Also, for $F \in \mathcal{P}_l$, we have $\left|\frac{|\nu|}{|\mu|}\mu(F) - \nu(F)\right| \leq \left|\frac{|\nu|}{|\mu|} - 1\right| \mu(F) + |\mu(F) - \nu(F)|$. Hence,

$$\begin{aligned} W_{p,\|\cdot\|}(\frac{|\nu|}{|\mu|}\mu, \nu) &\leq \kappa_p L^p D_p(\frac{|\nu|}{|\mu|}\mu, \nu) \\ &\leq \kappa_p L^p \sum_{l\geq 1} 4^{-l} \sum_{F\in\mathcal{P}_l} \left|\frac{|\nu|}{|\mu|} - 1\right| \mu(F) + \kappa_p L^p D_p(\mu,\nu) \\ &\leq \frac{\kappa_p}{3} L^p \left|\mu - \nu\right| + \kappa_p L^p D_p(\mu,\nu), \end{aligned}$$

and using that $(a+b)^p \leq 2^{p-1}(a^p + b^p)$, we obtain that $\mathrm{OT}^p_p(\mu,\nu) \leq L^p \left(2^{p-1} + \frac{\kappa_p}{3}\right) |\mu - \nu| + 2^{p-1}\kappa_p L^p D_p(\mu,\nu)$. We now apply this last inequality to $\mu = \mathbb{E}_P[\boldsymbol{\mu}]$ and $\nu = \overline{\boldsymbol{\mu}}_{\mathbf{K}}$. For $A \subset [-1,1)^2$ a measurable set, we have

$$\mathbb{E}[|\overline{\boldsymbol{\mu}}_{\mathbf{K}}(A) - \mathbb{E}_P[\boldsymbol{\mu}](A)|] \leq \sqrt{\frac{\mathbb{E}_P[\boldsymbol{\mu}(A)^2]}{K}} \leq \frac{M}{\sqrt{K}}. \tag{7.11}$$

Thus, $\mathbb{E}[|\overline{\boldsymbol{\mu}}_{\mathbf{K}} - \mathbb{E}_P[\boldsymbol{\mu}]|] \leq M/\sqrt{K}$ and, by using Jensen inequality twice,

$$\begin{aligned} &\sum_{F\in\mathcal{P}_l} \mathbb{E}_P[|\overline{\boldsymbol{\mu}}_{\mathbf{K}}(F) - \mathbb{E}_P[\boldsymbol{\mu}](F)|] \\ &\qquad \leq 2^l \left(\sum_{F\in\mathcal{P}_l} \mathbb{E}_P[(\overline{\boldsymbol{\mu}}_{\mathbf{K}}(F) - \mathbb{E}_P[\boldsymbol{\mu}](F))^2]\right)^{1/2} \\ &\qquad \leq \frac{2^l}{\sqrt{K}} \left(\sum_{F\in\mathcal{P}_l} \mathbb{E}_P[\boldsymbol{\mu}(F)^2]\right)^{1/2} \\ &\qquad \leq \frac{2^l}{\sqrt{K}} \left(M \sum_{F\in\mathcal{P}_l} \mathbb{E}_P[\boldsymbol{\mu}(F)]\right)^{1/2} \leq \frac{2^l M}{\sqrt{K}}. \end{aligned}$$

In the case $p > 1$, we obtain the conclusion as $\sum_{l\geq 1} 2^{(1-p)l} = c_p < \infty$. In the case $p = 1$, we use that $\sum_{F\in\mathcal{P}_l} \mathbb{E}_P[|\overline{\boldsymbol{\mu}}_{\mathbf{K}}(F) - \mathbb{E}_P[\boldsymbol{\mu}](F)|] \leq 2\sum_{F\in\mathcal{P}_l} \mathbb{E}_P[\boldsymbol{\mu}](F) \leq 2M$, yielding

$$\begin{aligned} \mathbb{E}[D_1(\overline{\boldsymbol{\mu}}_{\mathbf{K}}, \mathbb{E}_P[\boldsymbol{\mu}])] &\leq \sum_{l\geq 1} 2^{-l} \min\left(2M, \frac{2^l M}{\sqrt{K}}\right) \\ &\leq CM\frac{\ln(K)}{\sqrt{K}} \end{aligned}$$

for some absolute constant $C > 0$. □

Proof of Proposition 7.5. The upper bound on the minimax rate is given by the rates of convergence of the empirical EPD given in Proposition 7.4.

To obtain the lower bound, we use that if μ, ν are two persistence measures on Ω of mass smaller than M, then we can use the results proved in Section 3.2, namely Proposition 3.11 which states

$$\mathrm{OT}_p(\mu, \nu) = W_{p,\rho}(\Phi(\mu), \Phi(\nu)),$$

where

- ρ is a distance on $\tilde{\Omega} := \Omega \cup \{\partial\Omega\}$ defined by $\forall x, y \in \tilde{\Omega}$,

  $$\rho(x, y) = \min(d(x, y), d(x, \partial\Omega) + d(y, \partial\Omega)),$$

  where by convention for $x \in \Omega$, we set $d(x, \partial\Omega) := d(x, \mathbf{p}_{\partial\Omega}(x))$, with $\mathbf{p}_{\partial\Omega}(x)$ the orthogonal projection of x onto the diagonal $\partial\Omega$.

- $\Phi(\mu) := \mu + (2M - |\mu|)\delta_{\partial\Omega}$.

- $W_{p,\rho}$ denotes the Wasserstein distance between measures (with same mass) with distance ρ (see Section 2.2.2).

As Φ is a bijection, the minimax rates for the estimation of $\mathbb{E}_P[\mu]$ is therefore equal to

$$\inf_{\Phi(\hat{\mu}_{\mathbf{K}})} \sup_{P \in \mathcal{P}_{M,L}} \mathbb{E}[W^p_{p,\rho}(\Phi(\hat{\mu}_{\mathbf{K}}), \Phi(\mathbb{E}_P[\boldsymbol{\mu}]))].$$

The set $\mathcal{P}_{M,L}$ contains in particular the set of all distributions P_τ for which $\boldsymbol{\mu} \sim P$ satisfies $\Phi(\boldsymbol{\mu}) = 2M\delta_{\mathbf{x}}$, where $\mathbf{x} \sim \tau$, where τ is a probability measure on $(\tilde{\Omega}, \rho)$. For such a distribution P, one has $\Phi(\mathbb{E}_P[\boldsymbol{\mu}]) = 2M\tau$, so that, letting $\mathcal{Q}$ be the set of Borel probability measures on $(\tilde{\Omega} \cap [-L, L]^2, \rho)$, the minimax rate must be larger than

$$\inf_{\hat{\mathbf{a}}_{\mathbf{K}}} \sup_{\tau \in \mathcal{Q}} \mathbb{E}[W^p_{p,\rho}(\hat{\mathbf{a}}_{\mathbf{K}}, 2M\tau)],$$

where the infimum is taken on all measurable functions based on K observations of the form $2M\delta_{\mathbf{x}^{(\mathbf{k})}}$ with $\mathbf{x}^{(\mathbf{1})}, \ldots, \mathbf{x}^{(\mathbf{K})}$ a K-sample of law $\tau \in \mathcal{Q}$. Hence, we have shown that the minimax rate for the estimation of $\mathbb{E}_P[\boldsymbol{\mu}]$ with respect to OT_p is larger up to a factor M than the minimax rate for the estimation of τ a distribution on $(\tilde{\Omega} \cap [-L, L]^2, \rho)$ given K i.i.d. observations of τ. As the minimax rate for this problem is known to be larger than $L^p/\sqrt{K}$ [SP18, Theorem 9], we obtain the conclusion. □

7.2 Quantization of persistence diagrams

We address in this section the *quantization* of persistence diagram. By quantization, we mean a *projection* with respect to diagram metrics on a subspace of the space of persistence diagrams (or, as we will see, general persistence measures) consisting of measures with a fixed (finite) support size. Let $\mathcal{M}^p$ denote the space of persistence measures and $\mathcal{D}^p \subset \mathcal{M}^p$ be the space of persistence diagrams, for some parameter $1 \leq p \leq \infty$. Recall that $\mathcal{M}^p$ is equipped with a metric OT_p which generalizes the partial matching metric $\mathbf{d}_p$ defined on $\mathcal{D}^p$, see Section 2.1 for details. As we target numerical applications, we implicitly only consider in this section persistence diagrams and measures with *finite* support.

7.2.1 Definition and motivations

Definition 7.1. Fix some integer $k \geq 1$. Let $\mathcal{M}_k \subset \mathcal{M}^p(\overline{\Omega})$ be the persistence measures with support of size k supported on $\overline{\Omega}$, that is

$$\mathcal{M}_k := \left\{ \mu(a, X) := \sum_{i=1}^{k} a_i \delta_{x_i},\ a \in \mathbb{R}_+^k, X = (x_1 \dots x_k) \in \overline{\Omega}^k \right\}. \quad (7.12)$$

A k-quantization is a projection onto a subset of $\mathcal{X}$ of $\mathcal{M}_k$, that is a map

$$Q : \nu \mapsto Q(\nu) \in \operatorname*{arg\,min}_{\mu \in \mathcal{X}} \{\mathrm{OT}_p^p(\mu, \nu)\}, \quad (7.13)$$

In the vocabulary of quantization, the x_ks are called *codepoints* and $X = (x_1 \dots x_k)$ is called a *codebook*.

Remark 7.3. Note that *quantized* diagrams are supported on $\overline{\Omega}$ (i.e. are allowed to have points on the diagonal). This gives, for instance, a solution to the problem $\mathrm{minimize}_x\, \mathrm{OT}_p(\delta_x, \mathbf{0})$, where δ_x is a persistence diagram with a single point $x \in \overline{\Omega}$ and $\mathbf{0}$ represents the empty diagram.

As it will be used in the following, we recall that metrics OT_p can be computed as solution of a linear program (as introduced in Chapter 5). Let us consider two persistence measures with finite support $\mu = \sum_{i=1}^{n_1} a_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{n_2} b_j \delta_{y_j}$, where $a_i, b_j \geq 0$ and $x_i, y_j \in \overline{\Omega}$. Let $X = (x_i)_i, Y = (y_j)_j$ denote the locations of the points and $a = (a_i)_j, b = (b_j)_j$ be the corresponding masses. Let $\mathbf{p}_{\partial\Omega}(x)$ denote the (orthogonal) projection of $x \in \overline{\Omega}$ onto the diagonal $\partial\Omega := \{(t, t),\ t \in \mathbb{R}\}$. Let also $\mathrm{pers}(x)$ denote the distance of $x \in \overline{\Omega}$ to the diagonal, that is $\mathrm{pers}(x) := d(x, \mathbf{p}_{\partial\Omega}(x))$.

Let $C = C(X, Y)$ be the cost matrix with block structure

$$C = \begin{pmatrix} C_{XY} & C_{X\Delta} \\ C_{\Delta Y} & 0 \end{pmatrix} \in \mathbb{R}^{(n_1+1)\times(n_2+1)}, \tag{7.14}$$

with $C_{X\Delta} = (\mathrm{pers}(x_i)^p)_i$, $C_{\Delta Y} = (\mathrm{pers}(y_j)^p)_j$ and $C_{XY} = (d(x_i, y_j)^p)_{ij}$.

It reads

$$\mathrm{OT}_p^p(\mu, \nu) = \mathbf{L}_C(a, b) \tag{7.15}$$

where $\mathbf{L}_C$ is defined by the following optimization problems, dual of each other:

$$\mathbf{L}_C(a, b) = \min_{P\in\Pi(a,b)} \langle C, P\rangle = \min_{P\in\Pi(a,b)} \sum_{ij} C_{ij}P_{ij}, \tag{7.16}$$

$$= \max_{f,g\in\Psi_C} \sum_{i=1}^{n_1} a_i f_i + f_\Delta \left(\sum_j b_j\right) + \sum_{j=1}^{n_2} b_j g_j + g_\Delta \left(\sum_i a_i\right), \tag{7.17}$$

where f_Δ and g_Δ respectively denote the $(n_1 + 1)$-th and $(n_2 + 1)$-th coordinates of f and g, and the constraint sets are respectively defined as

$$\begin{aligned} \Pi(a, b) := \big\{ & P \in \mathbb{R}^{(n_1+1)\times(n_2+1)}, \\ & P\mathbf{1}_{n_2+1} = (a_1, \ldots, a_{n_1}, \sum_j b_j), \\ & P^T\mathbf{1}_{n_1+1} = (b_1, \ldots, b_{n_2}, \sum_i a_i)\big\}, \\ \Psi_C := \{(f, g) \in \mathbb{R}^{n_1+1} \times \mathbb{R}^{n_2+1}, \ & \forall i, j, f_i + g_j \le C_{ij}\}. \end{aligned}$$

Note that solving $\mathbf{L}_C$ and recovering (primal or dual) optimizers can be done using, for instance, the `POT` library [FC17].

Motivations. Quantization can be fairly useful in practice from different perspectives.

- First, it allows us to "summarize" the information contained in a persistence diagram. Indeed, persistence diagrams in practice are likely to have a large number of points (e.g. few thousands). Each point accounts for a topological feature in the filtration process generating the diagram (see Section 2.1), but when the diagram is too big, this information is overloaded. Quantizing the diagram provides a persistence diagram (or measure) close to the input while being much simpler to interpret.

- Reducing the size of diagrams has huge benefits from a computational perspective. Computing the partial matching distance $\mathbf{d}_p$ between two large diagrams is expensive. Given two diagrams μ, ν and a quantization Q, one can use $\mathrm{OT}_p(Q(\mu), Q(\nu))$ as an estimation of $\mathrm{OT}_p(\mu, \nu)$, where the error done is controlled by the quality of the quantization (using the triangle inequality).

- Finally, quantization can also be seen as a *vectorization* method (up to permutation). Indeed, an advantage of quantization in applications (in machine learning in particular) is that it turns a persistence diagram μ of arbitrary size into a vector in $\mathbb{R}^{3\times k}$ (two dimensions for the point coordinates x_i, and one for the weight a_i), allowing for the use of many machine learning algorithms. Note that the k points in the (quantized) diagrams are not ordered (that is, permuting coordinates and weights of the vector in $\mathbb{R}^3$ should be considered as representing the same diagram); this can be handled by using *permutation invariant* layers such as PersLay(see Section 6.2).

Example 7.1. Let $\mathcal{D}_k := \mathcal{M}_k \cap \mathcal{D}$ denote the space of persistence diagrams with exactly k points (supported on $\overline{\Omega}$). Let $\nu = \sum_{j=1}^{n} \delta_{y_j}$ be a persistence diagram. Assume for the sake of simplicity that $\mathrm{pers}(y_1) > \mathrm{pers}(y_2) > \cdots > \mathrm{pers}(y_n)$. Then,

$$\operatorname*{arg\,min}_{\mu\in\mathcal{D}_k} \mathbf{d}_2(\mu,\nu) = \sum_{i=1}^{k} \delta_{y_i},$$

that is, it is the diagram obtained from ν by keeping the k points of largest persistence. This quantization has the benefits of being straightforward to compute. Note that this technique, namely "only keeping points with highest persistence", is routinely used in applications (e.g. [CCI+19, CO08]). It is reinterpreted here as a solution of a natural minimization problem.

Example 7.2. Let us move temporarily out of the space of persistence diagrams. A well-known quantization problem—in the Wasserstein space $\mathcal{W}^2(\mathbb{R}^d)$, see Section 2.2—is the k-means problem. Indeed, consider a point cloud $\{y_1 \dots y_n\} \subset \mathbb{R}^d$ that is encoded by a probability measure $\nu = \frac{1}{n}\sum_{j=1}^{n} \delta_{y_j}$. Given an integer k, the k-means problem can be phrased as

$$\operatorname*{minimize}_{x\in\mathbb{R}^{d\times k}, a\in\mathbb{R}_+^k, \sum_i a_i=1} \left\{W_2^2(\mu(a,x),\nu)\right\},$$

where $\mu(a,x) := \sum_{i=1}^{k} a_i \delta_{x_i}$, and W_2 is the 2-Wasserstein distance.

Remark 7.4. Example 7.1 essentially solves the problem of quantizing "in $\mathcal{D}$", and we will thus focus on proposing quantizations whose output is *not*

a persistence diagram, in sense that mass a_i on a point x_i might not be integers. Considering this larger space of measures, which is made possible by connecting the diagram metrics $\mathbf{d}_p$ to the optimal transport metrics OT_p, allows us to obtain significantly better quantized diagrams while preserving the benefits of quantization mentioned above (optimizing on a larger space gives better minima). Note also that, although we restrict our focus to persistence diagrams in this section, the techniques presented below would apply to more general measures (e.g. empirical estimators of the EPD, see Section 7.1).

7.2.2 Optimization problem and Algorithm

General problem formulation. Let us now fix a persistence diagram $\nu = \sum_{j=1}^{n} \delta_{y_j}$ and $k \in \mathbb{N}$. Let $\mathcal{X}$ be a subset of $\mathcal{M}_k$. Our goal is to minimize the objective

$$\begin{aligned} F : \mathbb{R}_+^k \times \mathbb{R}^{k\times 2} &\to \mathbb{R} \\ (a, X) &\mapsto \mathbf{L}_{C(X,Y)}(a, b), \end{aligned} \tag{7.18}$$

under the constraint $\mu(a, X) \in \mathcal{X}$, where $\mu(a, X) = \sum_{i=1}^{k} a_i \delta_{x_i}$. For a given X, we denote by $K_X := \{a,\ \mu(a, X) \in \mathcal{X}\}$, and symmetrically, for a given a, we define $K_a := \{X,\ \mu(a, X) \in \mathcal{X}\}$.

Remark 7.5. Quantizations problems, such as k-means (Example 7.2), are in general (NP) hard to solve. The objective functional (7.18) is in general not convex. We will thus focus in the following on providing algorithms that converges to a *local minimum* of F (just as the Lloyd algorithm returns a local optimum of the k-means problem).

We propose to address this problem by *alternate minimization*: given an initialization $(a^{(0)}, X^{(0)})$, we build a sequence $(a^{(t)}, X^{(t)})_t$ by defining

$$\begin{aligned} a^{(t+1)} &= \min_{a \in K_X} F(a, X^{(t)}), \\ X^{(t+1)} &= \min_{X \in K_a} F(a^{(t+1)}, X). \end{aligned} \tag{7.19}$$

Obviously, such a sequence necessarily converges to a local minimum of F. The two minimization sub-problems will hopefully be simpler to deal with.

Mass update. Fix the location $X \in \mathbb{R}^{k\times 2}$, and consider $F_X : a \mapsto F(a, X)$. The dual formulation (7.17) yields

$$F_X(a) = \max_{f,g \in \Phi_C} \sum_{i=1}^{n_1} a_i f_i + f_\Delta \left(\sum_j b_j \right) + \sum_{j=1}^{n_2} b_j g_j + g_\Delta \left(\sum_i a_i \right) \tag{7.20}$$

Assume we have found optimal f^*, g^*. The *envelope theorem* states that a (sub-)gradient of F_X is given by

$$\nabla_a F_X = (f_i^* + g_\Delta^*)_{1 \leq i \leq k}. \tag{7.21}$$

Note that, as a supremum of convex (in fact, linear) functions, F_X is convex. Therefore, assuming K_X is closed and convex, F_X can be minimized using any projected gradient descent-based technique. For now, we only considered in our implementations the (naive) Euclidean gradient step and projection, that is for some learning rate $\lambda > 0$,

$$a \leftarrow \mathrm{proj}_{K_X}(a - \lambda \nabla_a F_X), \tag{7.22}$$

where proj_{K_X} is the projection on K_X for the Euclidean, that is $\mathrm{proj}_{K_X}(a') := \arg\min_{a'' \in K_X} \|a'' - a'\|_2^2$. More sophisticated techniques (e.g. accelerated gradient descent [Nes05]) are considered and will be incorporated in `Gudhi` if adapted. We also briefly mention the recent work of L.Chizat [Chi19] which considers optimization of discrete measures supported on a manifold (which is not $(\tilde{\Omega}, \rho)$)—it is likely that tools introduced in this work can be adapted to handle persistence diagrams and this will be investigated in future work.

Location update. Fix the weights $a \in \mathbb{R}^k$, and consider $F_a : X \mapsto F(a, X)$. The primal formulation (7.16) yields

$$\min_{X \in K_a} F_a(X) = \min_{X \in K_a} \min_{P \in \Pi(a,b)} \sum_{ij} P_{ij} C_{ij}. \tag{7.23}$$

Once again, we propose to solve this (non-convex) problem by alternating minimization between $X \in K_a$ and $P \in \Pi(a, b)$. The latter simply consists in solving the primal optimal transport problem. Now, given an optimal P^*, the former reads

$$\underset{X \in K_a}{\text{minimize}} \sum_{1 \leq i,j \leq k,n} P_{ij}^* d(x_i, y_j)^p + \sum_{i=1}^{k} P_{i\Delta}^* d(x_i, \mathbf{p}_{\partial\Omega}(x_i))^p, \tag{7.24}$$

where we recall that $\mathbf{p}_{\partial\Omega}(x_i)$ is the (orthogonal) projection of x_i onto the diagonal $\partial\Omega$.

In contrast to the update of masses, solving this minimization problem significantly depends on the nature of d, p, and K_a. We provide below some detailed examples.

Algorithm 5 Quantization of persistence diagrams.

Input: Integer k, initial masses $a^{(0)} \in \mathbb{R}^k$ and locations $X^{(0)} \in \overline{\Omega}^k$. Constraints K_a, K_X.
Output: A local minimum (a, X) of (7.18).
(Init) $a \leftarrow a^{(0)}, X \leftarrow X^{(0)}$.
while Not converged (1) **do**
  while Not converged (2) **do**
    Compute (f^*, g^*) optimal in (7.17) using `POT`
    Set $\nabla = (f_i^* + g_\Delta^*)_{i=1}^k$
    $a \leftarrow a - \lambda \nabla$
    $a \leftarrow \mathrm{proj}_{K_X}(a)$
  end while
  while Not converged (3) **do**
    Compute P^* optimal in (7.16) using `POT`
    $X \leftarrow X^*$ optimal in (7.24).
  end while
end while
return (a, X).

The meta-algorithm summarizing masses and locations updates is described in Algorithm 5.

Remark 7.6. It is worth noting that Algorithm 5 can be straightforwardly adapted to **not** update either the mass or the locations (simplifying the loop).

Remark 7.7. One can replace optimal primal P^* and dual (f^*, g^*) variables by their entropic regularized counterparts, see Section 2.2.4 for an introduction to entropic regularization of optimal transport, and Chapter 5 to see how it adapts to persistence measures. It can improves computational efficiently when k, n are getting large, but also the regularity of the algorithm, in particular with respect to the initialization. Indeed, without the entropic regularization, the optimal transport plan P^*, as a function of the current locations X, is not stable in sense that it is not unique (although being generically unique) and two close locations X', X'' can lead to two significantly different optimal transport plans $P^{*'}, P^{*''}$ and thus significant differences regarding the output of the algorithm. In contrast, the entropic regularized counterpart of P^* is unique and continuous with respect to X.

7.2.3 Examples

Let us provide examples in three natural contexts in which one may want to perform quantization of persistence diagrams. In the following, we fix an integer k and an input measure $\nu(b, Y) = \sum_{j=1}^{n} b_j Y_j$, $b \in \mathbb{R}^n, Y \subset \Omega$.

Unconstrained case. The simplest case to consider is the unconstrained one, that is we put no constraint on either locations nor masses aside from having non-negative masses, i.e. belonging to the positive orthan $\mathbb{R}^k_+$. With previous notations, and with $q = p = 2$—that is $d(x, y)^p = \|x - y\|_2^2$, it reads

- $a \geq 0$ entry-wise (formally, it yields $\forall X, K_X = R^k_+$).
- $X \in \overline{\Omega}^k$. Let $X^{(0)}, X^{(1)} \in \mathbb{R}^k$ denote the first and second coordinates of X respectively. We must have $X^{(1)} \geq X^{(0)}$ (element-wise). However, one can observe that if there is a point $x = (x^{(0)}, x^{(1)})$ in X such that $x^{(1)} < x^{(0)}$ (the point is below the diagonal), then replacing x by $x' = (x^{(1)}, x^{(0)})$ in X will decrease the objective $F(a, X)$. Therefore, this constraint is not even required explicitly. Note that $K_a = \overline{\Omega}^k$ for all $a \in \mathbb{R}^k_+$

In this context, the mass update (7.22) requires only to project onto the non-negative orthant, which simply reads

$$\mathrm{proj}_{\mathbb{R}^k_+}(x) = \max(x, 0),$$

where the maximum is meant term-wise. We thus focus on the location update. It simplifies to

$$\underset{x_1^{(0)} \ldots x_k^{(0)}, x_1^{(1)} \ldots x_k^{(1)} \in \mathbb{R}}{\text{minimize}} \sum_{ij} P^*_{ij} \left[\left(x_i^{(1)} - y_j^{(1)} \right)^2 + \left(x_i^{(0)} - y_j^{(0)} \right)^2 \right] + \frac{1}{2} \sum_i P^*_{i\Delta} \left(x^{(1)} - x^{(0)} \right)^2 + \sum_j P^*_{\Delta j} (y_j^{(1)} - y_j^{(0)})^2.$$

Proposition 7.7. *Assume that $a > 0$ element-wise. For $1 \leq i \leq k$, let $\mathbf{P}_i := \sum_{j=1}^{n} P^*_{ij}$ and $\overline{y}_i := \frac{1}{\mathbf{P}_i} \sum_{j=1}^{n} P^*_{ij} y_j$. Also, let $\tau_i := \frac{\mathbf{P}_i}{a_i} \leq 1$. The optimality condition reads*

$$\begin{cases} x_i^{(0)} = x_i^{(1)} & \text{if } \mathbf{P}_i = 0, \\ x_i = \tau_i \overline{y}_i + (1 - \tau_i)\, \mathbf{p}_{\partial\Omega}(\overline{y}_i) & \text{otherwise.} \end{cases} \tag{7.25}$$

Here, $\mathbf{P}_i$ represents the quantity of mass initially located on x_i that is off-diagonally transported, and $\overline{y}_i$ represents a barycenter of the $(y_j)_j$ weighted by the proportion of mass $P^*_{ij}/\mathbf{P}_i$ transported from x_i to y_j. The new position of x_i is eventually a barycenter between $\overline{y}_i$ and its orthogonal projection onto $\partial\Omega$, where the weight τ_i represents the fraction of mass that is off-diagonally transported.

Proof. Let $\mathcal{F}$ denote the function to minimize. Let $1 \leq i \leq k$, $\alpha \in \{0,1\}$, and $\alpha' = 1-\alpha$. The partial derivative of this function with respect to x_i^α is given by

$$\partial_{x_i^\alpha}\mathcal{F} = 2\sum_{1\leq j\leq n} P^*_{ij}(x_i^\alpha - y_j^\alpha) + P^*_{i\Delta}(x_i^\alpha - x_i^{\alpha'}).$$

Recall the relation $\mathbf{P}_i + P_{i\Delta} = a_i$ (marginal constraint on the transportation polytope). Equalizing this quantity to 0 gives

$$\begin{aligned} & 2\mathbf{P}_i x_i^\alpha - 2\mathbf{P}_i\overline{y}_i^\alpha + P_{i\Delta}(x_i^\alpha - x_i^{\alpha'}) = 0, \\ \text{that is} \quad & a_i x_i^\alpha - \mathbf{P}_i\overline{y}_i^\alpha = P_{i,\Delta}\left(\frac{x_i^\alpha + x_i^{\alpha'}}{2}\right), \end{aligned}$$

and same relation holds interchanging α and α', so that we have

$$\begin{cases} x_i^\alpha + x_i^{\alpha'} = \overline{y}_i^\alpha + \overline{y}_i^{\alpha'} \\ a_i x_i^\alpha - \mathbf{P}_i\overline{y}_i^\alpha = a_i x_i^{\alpha'} - \mathbf{P}_i\overline{y}_i^{\alpha'} \end{cases} \Rightarrow \begin{cases} x_i^\alpha + x_i^{\alpha'} = \overline{y}_i^\alpha + \overline{y}_i^{\alpha'} \\ x_i^\alpha - x_i^{\alpha'} = \tau_i(\overline{y}_i^\alpha - \overline{y}_i^{\alpha'}) \end{cases}$$

which gives

$$\begin{cases} 2x_i^\alpha = (1-\tau_i)(\overline{y}_i^\alpha + \overline{y}_i^{\alpha'}) + 2\tau_i y_i^\alpha, \\ 2x_i^{\alpha'} = (1-\tau_i)(\overline{y}_i^{\alpha'} + \overline{y}_i^\alpha) + 2\tau_i y_i^{\alpha'}, \end{cases} \tag{7.26}$$

which proves the claim, using that $\mathbf{p}_{\partial\Omega}(\overline{y}_i) = \left(\frac{\overline{y}_i^{(0)}+\overline{y}_i^{(1)}}{2}, \frac{\overline{y}_i^{(0)}+\overline{y}_i^{(1)}}{2}\right)$. □

Mass constrained. The unconstrained case presents the advantage of being fairly simple to solve. One might however want the quantized diagram to have more properties. For instance, one may ask that, although having k points, the quantized diagram should have the same total mass as the input diagram. Since the mass of the input diagram corresponds to its number of points (counted with multiplicity), and each point accounts for a topological event occurring in the filtration process, this constraint can be interpreted as asking for the quantized diagram to account for the same

"total number of events" observed in the filtration process. Formally, the constraint on the masses a simply reads (for any location X)

$$\sum_{i=1}^{k} a_i = a^T \mathbf{1}_k = n,$$

that is, a must lies on a re-scaled version of the probability simplex of dimension k, $\{u \in \mathbb{R}^k,\ u^T\mathbf{1}_k = 1, u \geq 0\}$. The (Euclidean) projection onto the probability simplex can be found using, for instance, [WCP13, Algorithm 1]. Cuturi and Doucet [CD14, §4.2] propose to rely on a Bregman projected gradient descent. In that context, a gradient step reads

$$a \leftarrow n \frac{a \odot \exp(-\lambda \nabla_a F)}{\|a \odot \exp(-\lambda \nabla_a F_X)\|_1},$$

where $\odot$ denotes element-wise multiplication, $\lambda > 0$ is a learning rate, and $\nabla_a F_X = (f_i^* + g_\Delta^*)_i$ for an optimal pair of dual potentials (f^*, g^*).

Note that the locations remain unconstrained, and one can use Proposition 7.7 to update the locations.

Persistence constrained. We finally consider a more complicated variation where we ask the quantized diagram to share the same *total persistence* as the input diagram ν. Recall that for a parameter p, the total persistence of $\mu = \sum_{i=1}^k a_i \delta_{x_i}$ is defined as

$$\mathrm{Pers}_p(\mu) = \sum_{i=1}^{k} a_i \mathrm{pers}(x_i)^p = a^T \mathrm{pers}(X),$$

where $\mathrm{pers}(X) := (\mathrm{pers}(x_i)^p)_i = (d(x_i, \mathbf{p}_{\partial\Omega}(x_i))^p)_i$. Informally, the total persistence can be understood as the total quantity of topological information contained in a diagram. Preserving the total persistence appears thus as a natural constraint.

Let $\rho = \mathrm{Pers}_p(\nu)$ be the total persistence of the input diagram ν. For a given X, the constraint on a is given by $K_X = \{u \geq 0\} \cap \{u,\ u^T \mathrm{pers}(X) = \rho\}$, which is the intersection of two convex sets: the non-negative orthant and an affine plan, it is thus convex. Furthermore, if X has at least one off-diagonal point, this intersection is not empty (assume $\mathrm{pers}(x_i) > 0$, and take $a = (0 \dots 0, \rho/\mathrm{pers}(x_i), 0 \dots 0)$). The (Euclidean) projection on this set can be found using a Linear Program solver.

Let us fix the mass a. The subtlety is that now, the locations $X = (x_i)_{i=1}^k$ are also constrained to satisfy $a^T \mathrm{pers}(X) = \rho$. In particular, this constraint is not convex if we set $p = 2$, yielding $d(x,y)^p = \|x-y\|_2^2$. We would face a Quadratically Constrained Quadratic Program, with non-convex constraints. Such a problem cannot be handle by a standard optimizer solver and would require to do optimization on the manifold $\left\{\frac{1}{2}\sum_{i=1}^n (x_i^{(1)} - x_i^{(0)})^2 = \rho\right\} = \{\mu \in \mathcal{M}_k,\ \mathrm{OT}_2^2(\mu, \mathbf{0}) = \rho\}$, which corresponds to a sphere of radius $\sqrt{\rho}$ centered on the empty diagram $\mathbf{0}$ on $(\mathcal{M}_k, \mathrm{OT}_2)$. Optimization on manifold is a wide topic for which techniques exist [AMS09, Bou14]. Those have not been implemented at the time I am writing these lines and are left for future work.

In order to recover a convex problem, we instead consider in the following the case $p = q = 1$, that is $d(x,y)^p = \|x-y\|_1$. The constraints now read

$$\sum_{i=1}^k a_i(x_i^{(1)} - x_i^{(0)}) = \rho, \quad X^{(1)} \geq X^{(0)} \tag{7.27}$$

and are thus convex. Meanwhile, the objective functional reads, for an optimal transport plan P^*,

$$\begin{aligned} F_a(X) = &\sum_{ij} P^*_{ij}\left(|x_i^{(0)} - y_j^{(0)}| + |x_i^{(1)} - y_j^{(1)}|\right) \\ &+ \sum_i P^*_{i\Delta}(x_i^{(1)} - x_i^{(0)}) + \sum_j P^*_{\Delta j}(y^{(1)} - y^{(0)}), \end{aligned}$$

which is convex. Minimizing F_a under the constraints (7.27) problem can be reformulated as a Linear Program, and thus finding an optimal X can be handled by a convex optimization solver such as `GLPK` interfaced with `cvxpy` [DB16], as done in our implementation. See Figure 7.3 for an illustration.

7.3 Shift-invariant distance between persistence diagrams

7.3.1 Generalities

Recall that a persistence diagram μ built on top of an object X come from a *filtration* $f : X \to \mathbb{R}$, which is denoted by $\mu = \mathrm{Dgm}(X, f)$—see Section 2.1 for details. In particular, the points $(x_i)_i$ of $\mu = \sum_{i=1}^n \delta_{x_i}$ are linked with some *critical values* of the map f. Roughly speaking[2],

[2]This can be made formal by observing that the persistence module of $(X, \varphi \circ f)$ indexed by $t \in \mathbb{R}$ is exactly the one of (X, f) indexed by $\varphi(t) \in \mathbb{R}$.

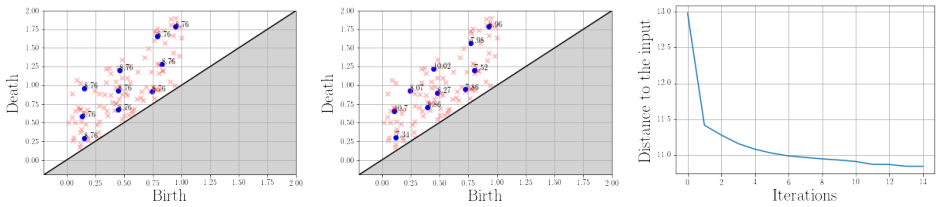


Figure 7.3: Persistence constrained quantization with $k = 10$. *(Left)* Initial state of the quantization algorithm. Initialization on location is performed using a simple k-mean algorithm on the points of the input diagram ν (shaded red). Initialization of weights is with $a_i = \rho / \sum_i \mathrm{pers}(x_i)$, so that $\mathrm{Pers}_1(\mu(a_{\mathrm{init}}, X_{\mathrm{init}})) = \rho$. *(Middle)* Output of the algorithm after convergence with constrained persistence. *(Right)* Evolution of $\mathcal{E}_t := \mathrm{OT}_1(\mu^{(t)}, \nu)$ through iterations t. Stopped when $|\mathcal{E}_{t+1}/\mathcal{E}_t - 1| \leq 0.001$.

the coordinates $(b_i, d_i) \in \Omega$ of a point x_i in the diagram $\mu = \mathrm{Dgm}(X, f)$ must correspond to local minima or maxima of f reached at b_i and d_i. As such, if one compose f with a (strictly) increasing map $\varphi : \mathbb{R} \to \mathbb{R}$, the resulting diagram $\mathrm{Dgm}(X, \varphi \circ f)$—that will also be denoted by $\varphi(\mu)$—is simply described by $\sum_{i=1}^n \delta_{\varphi(x_i)}$, where $\varphi(x_i) := (\varphi(b_i), \varphi(d_i)) \in \Omega$, with $x_i = (b_i, d_i) \in \Omega$. Note that in particular $\mathrm{Dgm}(X, \varphi \circ f)$ and $\mathrm{Dgm}(X, f)$ have the same number of points n, and one has a natural correspondence between those points (although this might not be the correspondence given by an optimal partial matching between the two diagrams). To sum up, the two diagrams are the same up to a transformation of the ground space Ω.

Example 7.3. A standard example, that will be deepen in the next section, is the case where f is shifted by a constant term $\tau \in \mathbb{R}$, that is $\varphi : t \mapsto t + \tau$ (written $\varphi \circ f = f + \tau$ for short). Intuitively, this should not change the topological information inferred from X. In practice, the resulting diagram $\mathrm{Dgm}(X, f + \tau)$ will be the diagram obtained from $\mathrm{Dgm}(X, f)$ where points have been shifted by $x_i \mapsto x_i + (\tau, \tau)$, that is a shift parallel to the diagonal. Similarly, one could consider $\varphi : t \mapsto \lambda t$ for some $\lambda > 0$, modeling a change of scale in the filtration. See also Example 7.4 below.

A natural question is thus: given two diagrams μ, ν, is it possible to compute their distance up to such a transformation of the ground space? Formally, it would read

$$\mathbf{d}_2^{\mathrm{shift},2} := \inf_{\varphi : \mathbb{R} \to \mathbb{R}} \mathbf{d}_2^2(\mu, \varphi(\nu)), \tag{7.28}$$

where the infimum is taken over strictly increasing functions φ. Note that,

for the sake of simplicity, we restricted ourselves to the case $p = 2$. Of course, such an infinite-dimensional optimization problem is in general intractable in practice, so we will instead assume that φ is of the form φ_θ, where $\theta \in \Theta \subset \mathbb{R}^d$ is a (finite-dimensional) set of parameters. In this context, (7.28) becomes

$$\mathbf{d}_2^{\mathrm{shift},2} := \inf_{\theta\in\Theta} \mathbf{d}_2^2(\mu, \varphi_\theta(\nu)), \tag{7.29}$$

For instance, one could have $\theta = (\lambda, \tau) \in (0, +\infty) \times \mathbb{R}$, and $\varphi_\theta : x \mapsto \lambda x + \tau$ (affine transformation of the filtration).

This optimization problem can be addressed in two ways, depending on the properties we have on φ_θ: by alternate minimization or by performing a gradient descent. Indeed, using the optimal transport formulation of the diagram metric $\mathbf{d}_2^2$, it turns into

$$\mathbf{d}_2^{\mathrm{shift},2} := \min_\theta F(\theta) = \min_\theta \min_P \sum_{ij} P_{ij} C_{ij}(\theta), \tag{7.30}$$

where the minimum over P is taken over transportation plans between the two diagrams—including the diagonal—(we refer to (7.16) in §7.2 for details), and the cost matrix $C(\theta) \in \mathbb{R}^{(n+1)\times(m+1)}$ is defined by the block structure

$$C(\theta) = \begin{pmatrix} C_{XY}(\theta) & C_{X\Delta} \\ C_{\Delta Y}(\theta) & 0 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(m+1)}, \tag{7.31}$$

with $C_{X\Delta} = (\|x_i - \mathbf{p}_{\partial\Omega}(x_i)\|_2^2)_i$, $C_{\Delta Y}(\theta) = (\|\mathbf{p}_{\partial\Omega}(\varphi_\theta(y_j)) - \varphi_\theta(y_j)\|_2^2)_j$ and $C_{XY}(\theta) = (\|\varphi_\theta(y_j) - x_i\|_2^2)_{ij}$, where as usual $\mathbf{p}_{\partial\Omega}$ denotes the orthogonal projection of $x \in \Omega$ onto the diagonal $\partial\Omega$.

At fixed θ, an optimal P^* can be found easily by solving the corresponding linear program, using for instance the (python) optimal transport library `POT` [FC17]. Once this is done, one can use an optimal θ^* in closed-form (if known) and then find the new P^* (alternate minimization). Otherwise, the *envelope theorem* ensures that the gradient of $\theta \mapsto F(\theta)$ is given by $\sum_{ij} P^*_{ij} \nabla_\theta C_{ij}(\theta)$, and then perform a gradient descent step.

Note that, in general, $\theta \mapsto F(\theta)$ is not convex and both methods can only guarantee to return a local optimum of the map $\theta \mapsto F(\theta)$. In the following subsection, we will study in detail the case where φ_θ simply induces a translation parallel to the diagonal. The techniques and remarks developed therein (in particular the use of entropic regularization) might easily be adapted to other contexts.

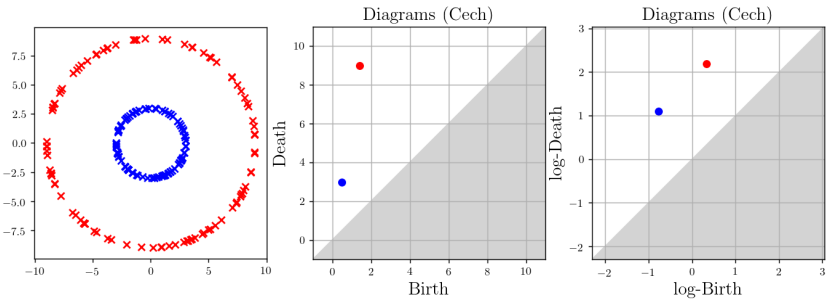


Figure 7.4: *(Left)* Two circles, the largest obtained by scaling the smallest by a factor $\lambda = 3$. *(Middle)* The standard Čech diagrams of both circles. *(Right)* The Čech diagrams in log-scale. They are obtained from each other by a translation (parallel to the diagonal) by a factor $\tau = \log(\lambda)$. As such, their distance *up to translation* is 0.

7.3.2 Shift parallel to the diagonal

Preliminary remarks

Let $\tau \in \mathbb{R}$ and consider the simple shift $t \mapsto t + \tau$ (at the filtration level) which induces (at the ground space level) a translation parallel to the diagonal. The diagram obtained from an input μ by shifting its points by a factor τ parallel to the diagonal will be denoted by $\tau(\mu)$. In this context, (7.28) yields

$$\mathbf{d}_2^{\tau,2}(\mu,\nu) := \inf_{\tau\in\mathbb{R}} f(\tau) := \inf_{\tau\in\mathbb{R}} \mathbf{d}_2^2(\mu,\tau(\nu)). \tag{7.32}$$

Although being extremely simple, this problem is probably the most standard one might consider when dealing with persistence diagrams up to a shift. We give in Example 7.4 and Example 7.5 some contexts where this problem arise.

Example 7.4. Let $X \subset \mathbb{R}^d$ be a (finite) point cloud centered at 0, and λX be the point cloud obtained from X by scaling it with a factor $\lambda > 0$. One can build the *Čech* diagrams (see Example A.2 for details) on top of X and λX, using distances in log-scale. This would respectively give two diagrams μ and $\nu = \tau(\mu)$, where $\tau = \log(\lambda)$. If the scaling factor λ is far from 1, μ and ν are two very different diagrams (in terms of the diagram metric d_p). However, one could argue that describing topology of a point cloud should be scale-independent. It invites us to consider distance between diagram up to translation, yielding (7.32). See Figure 7.4 for an illustration.

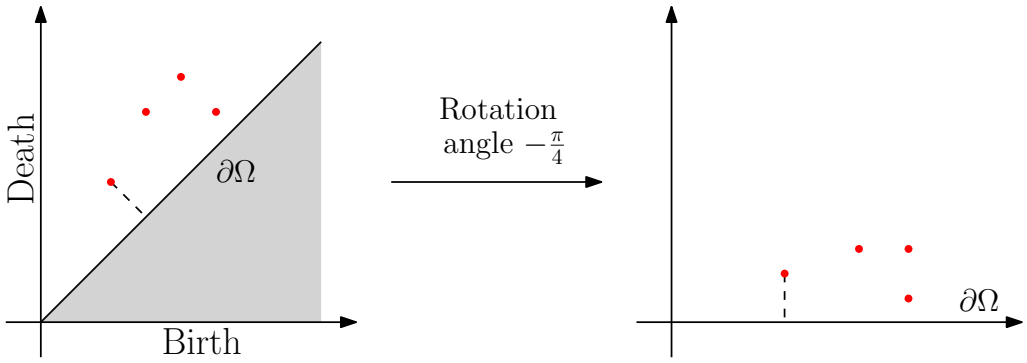


Figure 7.5: Rotation of the ground space used to make computations simpler.

Example 7.5. It might also happen that the filtration we want to consider is intrinsically defined up to a constant. In *symplectic topology* for instance, recent advances in the field were brought by considering persistence diagrams built using a Hamiltonian as a filtration. However, such filtrations are defined up to a constant term; the resulting persistence diagram is therefore only defined up to a translation parallel to the diagonal, see [RSV18, Remark 34] or the paragraph **barcodes** in [BHS18, §1.1] for more details. The natural metric that arises in this context to compare these topological descriptors is then exactly $\mathbf{d}_2^{\tau,2}$.

Remark 7.8. For the ease of notations and computations, in the remainder this section we rotate the ground space Ω by $-\frac{\pi}{4}$. In particular, the distance between a point $x = (x^{(0)}, x^{(1)})$ and the diagonal $\partial\Omega$ is now simply given by $x^{(1)}$, see Figure 7.5.

Remark 7.9. In standard optimal transport (see Section 2.2) on the Euclidean space $\mathbb{R}^d$, the analog problem is trivial: one just has to center the measures considered. Namely, if μ, ν are two probability measures in $\mathbb{R}^d$. We have that[3] $\inf_{\tau\in\mathbb{R}^d} W_2(\mu, \tau(\nu))$ is reached for $\tau = \mathbb{E}(\mu) - \mathbb{E}(\nu)$, where W_2 denotes the 2-Wasserstein distance on the Wasserstein space $\mathcal{W}^2(\mathbb{R}^d)$.

However, including the diagonal, even if the set of translations considered here is only one-dimensional ($\tau \in \mathbb{R}$), the problem becomes significantly harder as it is, in particular, not convex, see Figure 7.6.

Limitation of naive variational approaches. Intuitively, one could try to (locally) minimize the map $f : \tau \mapsto \mathbf{d}_2^2(\mu, \tau(\nu))$ by using either an alternate minimization or by doing a gradient descent as explained previously. Unfortunately, f is likely to have many plateaus (intervals on which it

[3]With $\tau(\nu)$ the *push-forward* of ν by the map $x \mapsto x + \tau$.

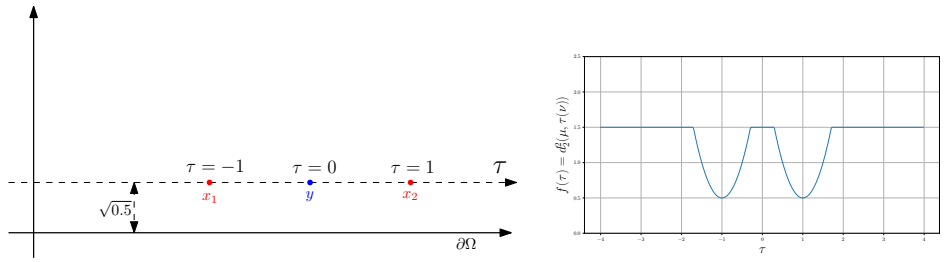


Figure 7.6: Illustration of the non-convexity of $\tau \mapsto \mathbf{d}_2^2(\mu, \tau(\nu))$ in a simple case. Here, $\mu = \delta_{x_1} + \delta_{x_2}$ (red points), and $\nu = \delta_y$ (blue point). Here, f has two (global) minima: $\tau = -1$ and $\tau = 1$. However, $f(0)$ is not a minimum. On the left, the graph of the map $\tau \mapsto f(\tau)$, which illustrates (a) non convexity of f, (b) presence of points of non-differentiability, (c) presence of plateaus.

takes a constant value) or points of non-differentiability due to the peculiar role played by the diagonal, see for instance Figure 7.6. As a consequence, if initialized in such plateaus, both methods would get stuck and return a clearly unsatisfactory result.

Remark 7.10. We point out that a similar problem was already considered by Sheehy et al. in [SKC18]. Their method consists in adopting a combinatorial approach to provide a *global* minimizer of F. Their work differs from the one we present here in the following ways:

- Sheehy et al. consider the bottleneck distance (case $p = \infty$, i.e. turning the sum into a max), while we consider the Wasserstein distance with $p = 2$ (although our approach works for any $p \geq 1$ using as ground metric $d(x, y) = \|x - y\|_p$).

- They did not focus on tractability of their algorithm: theoretical complexity is in $\mathcal{O}(n^{3.5})$ (where n is the cardinality of the diagrams in input) and they report a running time of 2 minutes for $n = 128$. Our implementation can handle much larger instances.

- Although developed only in the context of persistence diagrams, our approach straightforwardly adapts to any discrete measure supported on Ω and, theoretically, even to measures with a non-discrete support, adopting the formalism developed in Chapter 3, allowing to compute shift-invariant distance between persistence surfaces (see Section 6.1 and [AEK+17]), expected persistence diagrams (see Section 7.1) or quantized diagrams (see Section 7.2) for instance.

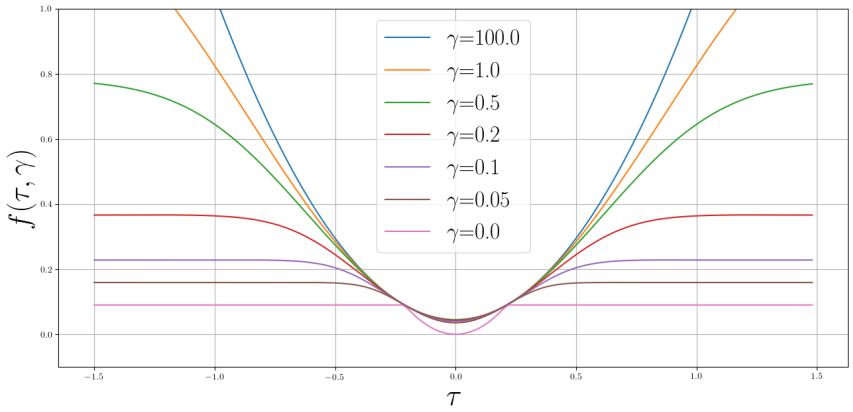


Figure 7.7: Illustration of the role of the regularization γ when $\mu = \nu = \delta_x$ is a diagram with a single point. When $\gamma = 0$, $f(\tau, 0) = \mathbf{d}_2^2(\mu, \tau(\nu))$ takes a constant value as long as $|\tau| > 2\|x - \mathbf{p}_{\partial\Omega}(x)\|_2^2$—in which case it becomes cheaper to transport the point of both diagram to the diagonal. Increasing the value of γ helps to mitigate this phenomenon.

Entropic optimal transport to get non-zero gradients

In order to mitigate the effect of these plateaus, we propose to use (once again) entropic optimal transport. This will also have benefits on differentiability and computational efficiency. With the notations above, let $X = (x_1 \dots x_n) \in R^{n\times 2}$ and $Y = (y_1 \dots y_m) \in R^{m\times 2}$ be the points appearing in μ and ν respectively. Define $a = (1 \dots 1, m)/(n+m) \in \mathbb{R}^{n+1}$ and $b = (1 \dots 1, n)/(n+m) \in \mathbb{R}^{m+1}$. Note that here, we renormalized a, b so that they belong to the probability simplex, allowing the use of standard (regularized) optimal transport results off-the-shelf.

We recall that we work with $p = 2$, and we thus we define $c(x, y) = \|x - y\|_2^2$ and $c(x, \partial\Omega) := \|x - \mathbf{p}_{\partial\Omega}(x)\|_2^2 = (x^{(1)})^2$, where $x = (x^{(0)}, x^{(1)}) \in \Omega$ (recall that we performed a rotation of the ground space Ω by $-\pi/4$ to simplify computations, see Figure 7.5). Consider $C(\tau) := c(x_i, y_j + (\tau, 0))$ the cost matrix where locations Y of ν get translated by τ, parallel to $\partial\Omega$. Note that $c(\partial\Omega, y) = c(\partial\Omega, y + (\tau, 0))$. Our goal becomes to minimize the function $f(\cdot, \gamma) : \mathbb{R} \to \mathbb{R}$ defined for a parameter $\gamma \geq 0$ by:

$$f(\tau, \gamma) = \min_{P \in \Pi(a,b)} E(P, \tau, \gamma) \tag{7.33}$$

with

$$E(P, \tau, \gamma) := \langle P, C(\tau) \rangle + \gamma \mathrm{KL}(P | a \otimes b) \tag{7.34}$$

which corresponds to the primal formulation of entropic regularized optimal transport problem (see Section 2.2 and Chapter 5). Here, $\Pi(a,b)$ denotes the transportation polytope between a and b and $\mathrm{KL}(P|a\otimes b) := \sum_{ij} P_{ij}\log\left(\frac{P_{ij}}{a_i b_j}\right)$ is the Kullback-Leibler divergence.

We denote by $P(\tau,\gamma)$ the minimizer of $P \mapsto E(P,\tau,\gamma)$, which is unique if $\gamma > 0$, and by convention $P(\tau,0)$ denotes the minimizer of $P \mapsto E(P,\tau,0)$ of largest entropy. It is well known [PC17, Prop. 4.1] that $P(\tau,\gamma) \to P(\tau,0)$ as $\gamma \to 0$. We also denote by $P_{\partial\Omega}$ the "trivial" transport plan that maps every point to (its orthogonal projection onto) the diagonal $\partial\Omega$, formally defined as

$$P_{\partial\Omega} := \begin{pmatrix} & 0 & & \frac{1}{n+m} \\ & & & \vdots \\ & & & \frac{1}{n+m} \\ \frac{1}{n+m} & \cdots & \frac{1}{n+m} & 0 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(m+1)}.$$

By definition, we know that for all τ,γ,

$$\begin{aligned} f(\tau,\gamma) \leq & E(P_{\partial\Omega},\tau,\gamma) = \frac{1}{n+m}(\mathrm{Pers}(\mu)+\mathrm{Pers}(\nu)) + \gamma\mathrm{KL}(P_{\partial\Omega}|a\otimes b) \\ =: & E_{\partial\Omega}(\gamma), \end{aligned}$$

where $\mathrm{Pers}(\mu) = \sum_{i=1}^n \|x_i - \mathbf{p}_{\partial\Omega}(x_i)\|_2^2$ is the *total persistence* of μ ($\mathrm{Pers}(\nu)$ is defined likewise). Importantly, the latter quantity does not depend on τ, and thus $f(\cdot,\gamma)$ is bounded for all $\gamma \geq 0$.

Some intuition. Why could it help to add this entropic regularization? Aside from its general strengths (faster computations, access to gradients, etc.), in our context, the key idea is that the (entropic regularized) optimal transport plan $P(\tau,\gamma)$ is *dense*, that is $P(\tau,\gamma)_{ij} > 0$ for all i,j. To be more precise, there exist two diagonal matrices u,v (unique up to a scaling factor) such that

$$P(\tau,\gamma) = u \cdot \exp\left(-\frac{C(\tau)}{\gamma}\right) \cdot v.$$

Therefore, in contrast to the non-regularized optimal transport plan (which is, in this case, a permutation matrix (Proposition 2.6 and Proposition 3.3), the smoothed transport plan takes account of interactions between *all* pairs of points x,y even if those are far away. In particular, the situation where everything is transported onto the diagonal $\partial\Omega$ cannot occur when adding this entropic regularization[4]. See Figure 7.7 for an illustration.

[4]Note that however, in practice, if $C(\tau)/\gamma$ has some large entries at indices i,j, it numerically yields $\exp(-C(\tau)_{ij}/\gamma) \simeq 0$, which would impose $P_{ij} = 0$.

We now state asymptotic properties (in both γ and τ) of f.

Proposition 7.8 (Asymptotic behavior)**.** *One has*

$$\lim_{\gamma\to 0} f(\tau,\gamma) = \min_{P\in\Pi(a,b)} \langle P, C(\tau)\rangle =: f(\tau,0). \tag{7.35}$$

and

$$\lim_{\gamma\to\infty} f(\tau,\gamma) = \langle a\otimes b, C(\tau)\rangle =: f(\tau,\infty). \tag{7.36}$$

In particular, $f(\tau,\infty)$ admits a unique minimizer

$$\tau(\infty) := \frac{1}{m}\sum_j y_j^0 - \frac{1}{n}\sum_i x_i^0. \tag{7.37}$$

Similarly, one has

$$\lim_{|\tau|\to\infty} f(\tau,\gamma) = E_{\partial\Omega}(\gamma). \tag{7.38}$$

Proof. Equations (7.35) and (7.36) are now standard results in optimal transport theory, see [Gen19, Theorem 10] for instance. Note that

$$\begin{aligned}\langle a\otimes b, C(\tau)\rangle =& \frac{1}{(n+m)^2}\sum_{1\le i,j\le n,m} [(\tau + y_j^{(0)} - x_i^{(0)})^2 + (y_j^{(1)} - x_i^{(1)})^2] \\ &+ \frac{n}{(n+m)^2}\mathrm{Pers}(\nu) + \frac{m}{(n+m)^2}\mathrm{Pers}(\mu),\end{aligned}$$

which give the value of $\tau(\infty)$.

For the asymptotic in τ, let $P(\tau,\gamma)$ denote the (unique) minimizer of f. Assume that for some $1\le i,j\le n,m$, we do not have $P_{ij}(\tau,\gamma)C_{ij}(\tau)$ bounded as $\tau\to\infty$. As the Kullback-Leibler divergence is always non-negative, we have $f(\tau,\gamma)\ge\sum_{ij} C_{ij}(\tau)P_{ij}(\tau,\gamma)\to\infty$. This contradicts the fact that f must remain bounded. Therefore, as $C_{ij}(\tau)\to\infty$, we necessarily have $P_{ij}(\tau,\gamma)\to 0$ and thus $P(\tau,\gamma)\to P_{\partial\Omega}$ (due to marginal constraints) and (as E is continuous) $f(\tau,\gamma)\to E_{\partial\Omega}(\gamma)$. □

Intuitively, the larger the regularization γ, the simpler it is to minimize $f(\cdot,\gamma)$, to the extreme case $\gamma=\infty$ where $f(\cdot,\infty)$ is even convex with a minimizer $\tau(\infty)$ that can be found in closed-form. Of course, a (global) minimizer $\tau(\gamma)$ of $f(\cdot,\gamma)$ has no reason to be a good approximation of some $\tau(0)$ (a minimizer of our initial objective, $f(\tau,0)$). There is a natural trade-off between a large γ (making the optimization problem simple) and a small one, that would give satisfactory minimizers. We first make this last point more precise.

Proposition 7.9 (Convergence of minima and minimizers)**.** *We have convergence of minimum values:*

$$\lim_{\gamma\to 0}\min_{\tau\in\mathbb{R}} f(\tau,\gamma)=\min_{\tau\in\mathbb{R}} f(\tau,0). \tag{7.39}$$

Now, let $\tau(\gamma)$ *be in* $\arg\min f(\tau,\gamma)$ *for* $\gamma>0$*. One has*

$$\begin{cases}\lim_{\gamma\to 0}\tau(\gamma)\in\arg\min_\tau f(\tau,0)\\ f(\tau(\gamma),\gamma)\to\min_\tau f(\tau,0)\end{cases} \tag{7.40}$$

In order to prove this proposition, we introduce the following lemma.

Lemma 7.10 (Variations)**.** *One has, for any* (τ,γ) *and* (τ',γ')*:*

$$|f(\tau,\gamma)-f(\tau,\gamma')|\leq 2\,|\gamma-\gamma'|\log(n+m). \tag{7.41}$$

In particular, the map $f(\cdot,\gamma)$ *uniformly converges to* $f(\cdot,0)$ *as* $\gamma\to 0$*.*

Similarly,

$$|f(\tau',\gamma)-f(\tau,\gamma)|\leq\frac{1}{n+m}\|C(\tau)-C(\tau')\|_1 \tag{7.42}$$

Proof. Recall that $P(\tau,\gamma)$ denotes the unique minimizer of $P\mapsto E(P,\tau,\gamma)$ (or, in the particular case $\gamma=0$, the one with largest entropy).

By definition,

$$\begin{aligned}f(\tau,\gamma)&=\langle P(\tau,\gamma),C(\tau)\rangle+\gamma\mathrm{KL}(P(\tau,\gamma)|a\otimes b)\\&\leq\langle P(\tau,\gamma'),C(\tau)\rangle+\gamma\mathrm{KL}(P(\tau,\gamma')|a\otimes b)\\&=f(\tau,\gamma')+(\gamma-\gamma')\mathrm{KL}(P(\tau,\gamma')|a\otimes b).\end{aligned}$$

Now, observe that $\mathrm{KL}(P|a\otimes b)=\sum_{ij}P_{ij}\log(P_{ij})-\sum_{ij}P_{ij}\log(a_ib_j)$. The

first term is non-positive. Furthermore,

$$\begin{aligned}
-\sum_{ij} P_{ij}\log(a_i b_j) =& \sum_{1\le i,j\le n,m} P_{ij}\log((n+m)^2) \\
&+ \sum_{1\le i\le n} P_{i,m+1}\log((n+m)^2/n) \\
&+ \sum_{1\le j\le m} P_{n+1,j}\log((n+m)^2/m) \\
&+ P_{n+1,m+1}\log((n+m)^2/nm) \\
=& 2\log(n+m) - (\log(n)\sum_{1\le i\le n} P_{i,m+1} \\
&+ \log(m)\sum_{1\le j\le m} P_{n+1,j} + \log(nm)P_{n+1,m+1}) \\
=& 2\log(n+m) - \frac{n}{n+m}\log(n) - \frac{m}{n+m}\log(m) \\
\le& 2\log(n+m).
\end{aligned}$$

It proves the first inequality.

The second inequality is obtained in a similar way. By definition,

$$\begin{aligned}
f(\tau,\gamma) &\le \langle P(\tau',\gamma), C(\tau)\rangle + \gamma \mathrm{KL}(P(\tau',\gamma)|a\otimes b) \\
&= f(\tau',\gamma) + \langle P(\tau',\gamma), C(\tau) - C(\tau')\rangle\,.
\end{aligned}$$

Note that for $i = n+1$ or $j = m+1$, C does not depend on τ. Using the Hölder inequality on the indices $1 \le i,j \le n,m$, we can write

$$\langle P(\tau',\gamma), C(\tau) - C(\tau')\rangle \le \|P\|_q \|C(\tau) - C(\tau')\|_{q'}$$

where $1/q + 1/q' = 1$. Taking $q = \infty$, and observing that (on the aforementioned indices) $P(\cdot) \le \frac{1}{n+m}$ (due to the marginal constraints) gives the result. □

Proof of Proposition 7.9. Given uniform convergence of the maps (Lemma 7.10), convergences of minima and minimizers follow from standard result in Γ-convergence literature (see [DM12]). More precisely, (7.39) is given by [DM12, Theorem 7.4], and (7.40) is given by [DM12, Corollary 7.20]. □

Proposition 7.9 supports that it is reasonable to minimize $f(\cdot,\gamma)$ for some small $\gamma > 0$. Actually, it even suggests that one can start by minimizing $f(\cdot,\gamma)$ with some large γ, obtaining a minimizer $\tau(\gamma)$, and progressively decreasing γ. These techniques cannot guarantee to reach a global optimum

of $f(\cdot, 0)$, but significantly improve on doing a naive alternate minimization on $f(\cdot, 0)$ directly. We give in the following the formula to optimize τ doing either an alternate minimization or a gradient step. Recall that computing $P(\tau, \gamma)$ can be done using the Sinkhorn algorithm—see Section 2.2—and is handle in practice by the library `POT` [FC17] (for instance).

Proposition 7.11 (Optimization steps)**.** *Let $\tau \in \mathbb{R}$ and $\gamma > 0$. Let $P \in \Pi(a, b)$ be a transport plan and assume that $\sum_{1 \leq i,j \leq n,m} P_{ij} > 0$ (i.e. there is some off-diagonal mass that is transported). Then, alternate minimization can be performed by using the following formula:*

$$\arg\min_{\tau} E(P, \tau, \gamma) = \frac{\sum_{1 \leq i,j \leq n,m} P_{ij}(x_i^{(0)} - y_j^{(0)})}{\sum_{1 \leq i,j \leq n,m} P_{ij}}. \tag{7.43}$$

Now, let $P(\tau, \gamma)$ be the unique minimizer of $P \mapsto E(P, \tau, \gamma)$. Then gradient descent can be performed by using the following formula

$$\frac{\partial f(\tau, \gamma)}{\partial \tau} = 2 \sum_{1 \leq i,j \leq n,m} P(\tau, \gamma)_{ij}(\tau + y_j^{(0)} - x_i^{(0)}). \tag{7.44}$$

Proof. Equation (7.43) follows from straightforward computations. Indeed,

$$\begin{aligned} E(P, \tau, \gamma) = &\sum_{ij} P_{1 \leq ij \leq n,m}[(\tau + y_j^{(0)} - x_i^{(0)})^2 + (y_j^{(1)} - x_i^{(1)})^2] \\ &+ \sum_{i=1}^{n} P_{i,m+1} \|x_i - \mathbf{p}_{\partial\Omega}(x_i)\|_2^2 + \sum_{j=1}^{m} P_{n+1,j} \|y_j - \mathbf{p}_{\partial\Omega}(y_j)\|_2^2 \\ &+ \gamma \mathrm{KL}(P | a \otimes b) \end{aligned}$$

which is a quadratic (in particular, convex) map in τ (as long as $\exists (1 \leq i, j \leq n, m) | P_{ij} > 0$) which is minimized for

$$\sum_{ij} P_{1 \leq ij \leq n,m}(\tau + y_j^{(0)} - x_i^{(0)}) = 0,$$

which proves the formula.

Equation (7.44) is obtained using the envelope theorem which here simply reads

$$\begin{aligned} \frac{\partial}{\partial \tau} f(\tau, \gamma) = &\sum_{ij} P(\tau, \gamma)_{ij} \frac{\mathrm{d}}{\mathrm{d}\tau} C(\tau)_{ij} \\ = &2 \sum_{1 \leq i,j \leq n,m} P(\tau, \gamma)_{ij}(\tau + y_j^{(0)} - x_i^{(0)}), \end{aligned}$$

concluding the proof. □

This approach is eventually summarized in Algorithm 6. We propose to initialize on $\tau(\infty)$ as defined in (7.37). We initialize γ at 1, but other values might be satisfactory too. We run our experiments with a preset series of decreasing values of γ for illustration purpose but other heuristics such as $\gamma \leftarrow \gamma/2$ would make sense. Note also that we have access to $\frac{\partial f}{\partial \gamma}$, which is simply given by $\mathrm{KL}(P(\tau, \gamma)|a \otimes b)$ (once again, this follows from the envelope theorem).

Algorithm 6 Shift-invariant distance between persistence diagram.

Input: Initial diagrams locations X (fixed) and Y (optimized). Update rule for γ. If gradient descent, learning rate $\lambda > 0$.
Output: A local minimum τ^* of $f(\tau, \gamma)$.
(Init) $\tau = \tau(\infty)$, $\gamma = 1$.
while Not converged (1) **do**
  while Not converged (2) **do**
    Compute $P(\tau_t, \gamma_{t'})$ optimal using `POT` .
    if Gradient descent **then**
      Set $\nabla = \frac{\partial f}{\partial \tau}$ using (7.44).
      $\tau_{t+1} \leftarrow \tau_t - \lambda \nabla$.
    end if
    if Alternate minimization **then**
      Set $\tau_{t+1} \leftarrow \arg\min_\tau E(P(\tau_t, \gamma_{t'}), \tau, \gamma_{t'})$ using (7.43).
    end if
  end while
  Update γ
end while
return (a, X).

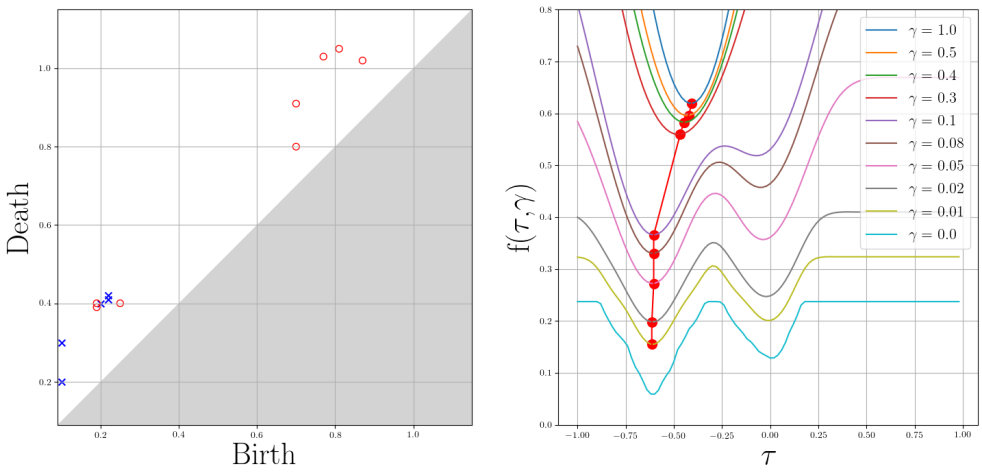


Figure 7.8: Illustration of the shift-invariant distance algorithm on a simple example. *(Left)* The two diagrams we want to "align", at $\tau = 0$. Intuitively, one should shift the red diagram by $\tau \simeq -0.6$; however the presence of a red cluster close to some of the blue points makes $\tau \simeq 0$ a local optimum of $\tau \mapsto \mathbf{d}_2^2(\mu, \tau(\nu)) =: f(\tau, 0)$. *(Right)* The graphs of $\tau \mapsto f(\tau, \gamma)$ for different regularization parameters γ. In red dots, the values $(\tau(\gamma_t), f(\tau(\gamma_t), \gamma_t))$ for γ_t decreasing from 1.0 to 0.01. In this (lucky) configuration, our approach does converge (close) to the global optimum, while directly working with $\gamma = 0$ (or even a small γ) would get stuck in the local optimum around $\tau = 0$.

Conclusion

The similarities between the metrics used to compare persistence diagrams in topological data analysis and those used to compare probability measures in optimal transport were known for long. However, an explicit connection between these two formalisms was yet to come. In this thesis, we established such a connection. It turned out to be very prolific, especially when targeting geometric and statistical applications but also in the conception of new algorithms based on optimal transport techniques that could be of great use when dealing with persistence diagrams.

The contributions of this work were presented in two different parts. Part I is dedicated to the theoretical results. Chapter 3 is the engine of the machinery we developed. We show how persistence diagram metrics can be reformulated as optimal partial transport problems. Of major interest, our formalism handles any kind of non-negative measures, not only those with discrete support such as persistence diagrams. We study the properties of the resulting space in detail, proving some powerful results such as a characterization of convergence in the space. Chapter 4 studies Fréchet means in this space from a theoretical perspective. We have established a strong connection between Fréchet means of persistence diagrams and their counterpart in optimal transport theory. It allows us to prove their existence as well as a consistency result in a great generality.

Part II is dedicated to applications. Following Chapter 4, Chapter 5 is dedicated to the estimation of Fréchet means of persistence diagrams in practice. By switching from a Lagrangian approach to a Eulerian one, we retrieved a convex problem, a major improvement over previous approaches. Furthermore, building on entropic optimal transport, we obtain a GPU-friendly parallelizable implementation, especially efficient on large sets of large diagrams. Chapter 6 is dedicated to linear representations of persistence diagrams, for which we provide an exhaustive criterion to prove their continuity. We then propose a neural network architecture, named PersLay, that will *learn* a task-driven linear representation. Our

publicly available implementation is made to encompass most preexisting vectorization techniques used when incorporating persistence diagrams in machine learning pipelines. Finally, Chapter 7 provides some complementary examples that showcase the benefits of our work. We study persistence diagrams in a random setting for which measures with continuous support naturally appear. We prove stability properties of persistence diagrams in this context, a result that only makes sense after extending diagram metrics to measures with continuous support. This is done thanks to our optimal transport based approach. Similarly, we propose an algorithm to perform quantization on persistence diagrams, and a method to estimate a shift-invariant distance between them.

We believe this work opens the way for faithful use of optimal transport tools in combination with topological descriptors such as persistence diagrams, both in theory and in pratice. Of course, many tracks remain to be explored. Pretty much any aspect of the optimal transport literature may lead to an interesting counterpart in topological data analysis. We think for instance of gradient flows [AGS08], convergence rates based on geometric properties of the space [GPRS19] and using different regularization terms [BSR17], to name a few.

Bibliography

[AC11] Martial Agueh and Guillaume Carlier. Barycenters in the wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011. *Cited (3) on pages 49, 92, and 93*

[AEK+17] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: a stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18(8):1–35, 2017. *Cited (12) on pages 16, 27, 44, 125, 126, 132, 134, 138, 139, 140, 150, and 169*

[AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008. *Cited on page 180*

[AKP17] Stephanie Alexander, Vitali Kapovitch, and Anton Petrunin. Alexandrov geometry. *Book in preparation*, 2017. *Cited on page 221*

[AMS09] P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009. *Cited on page 164*

[AWR17] Jason Altschuler, Jonathan Weed, and Philippe Rigollet. Near-linear time approximation algorithms for optimal transport via sinkhorn iteration. In *Advances in Neural Information Processing Systems*, pages 1961–1971, 2017. *Cited (3) on pages 118, 119, and 122*

[BBB+01] Dmitri Burago, Iu D Burago, Yuri Burago, Sergei A Ivanov, and Sergei Ivanov. *A course in metric geometry*, volume 33. American Mathematical Soc., 2001. *Cited on page 221*

[BE19] Peter Bubenik and Alex Elchesen. Universality of persistence diagrams and the bottleneck and wasserstein distances. *arXiv preprint arXiv:1912.02563*, 2019. *Cited on page 79*

[BGMP14] Andrew J Blumberg, Itamar Gal, Michael A Mandell, and Matthew Pancia. Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces. *Foundations of Computational Mathematics*, 14(4):745–789, 2014. *Cited on page 74*

[BHS18] Lev Buhovsky, Vincent Humiliere, and Sobhan Seyfaddini. The action spectrum and c^ 0 symplectic topology. *arXiv preprint arXiv:1808.09790*, 2018. *Cited on page 168*

[BLPY19] Greg Bell, Austin Lawson, C. Neil Pritchard, and Dan Yasaki. The space of persistence diagrams fails to have yu's property a, 2019. *Cited on page 27*

[BM19] Christophe AN Biscio and Jesper Møller. The accumulated persistence function, a new useful functional summary statistic for topological data analysis, with a view to brain artery trees and spatial point process applications. *Journal of Computational and Graphical Statistics*, (just-accepted):1–20, 2019. *Cited on page 44*

[BMEWL11] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 2011. *Cited (2) on pages 9 and 20*

[Boc33] Salomon Bochner. Integration von funktionen, deren werte die elemente eines vektorraumes sind. *Fundamenta Mathematicae*, 20(1):262–176, 1933. *Cited on page 151*

[Bou14] Nicolas Boumal. *Optimization and estimation on manifolds.* PhD thesis, Catholic University of Louvain, Louvain-la-Neuve, Belgium, 2014. *Cited on page 164*

[BSR17] Mathieu Blondel, Vivien Seguy, and Antoine Rolet. Smooth and sparse optimal transport. *arXiv preprint arXiv:1710.06276*, 2017. *Cited on page 180*

[BT03] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003. *Cited on page 125*

[Bub15] Peter Bubenik. Statistical topological data analysis using persistence landscapes. *The Journal of Machine Learning Research*, 16(1):77–102, 2015. *Cited (4) on pages 16, 27, 44, and 138*

[BV18] Peter Bubenik and Tane Vergili. Topological spaces of persistence modules and their properties. *Journal of Applied and Computational Topology*, pages 1–37, 2018. *Cited (3) on pages 27, 41, and 74*

[BW19] Peter Bubenik and Alexander Wagner. Embeddings of persistence diagrams into hilbert spaces. *arXiv preprint arXiv:1905.05604*, 2019. *Cited (2) on pages 16 and 27*

[CB15] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 14(05):1550066, 2015. *Cited on page 211*

[CB18] M. Carriere and U. Bauer. On the Metric Distortion of Embedding Persistence Diagrams into Reproducing Kernel Hilbert Spaces. *ArXiv e-prints*, June 2018. *Cited (2) on pages 16 and 27*

[CCI+19] Mathieu Carrière, Frédéric Chazal, Yuichi Ike, Théo Lacombe, Martin Royer, and Yuhei Umeda. PersLay: A Neural Network Layer for Persistence Diagrams and New Graph Topological Signatures. *arXiv preprint arXiv:1904.09378*, 2019. *Cited (8) on pages 19, 29, 131, 136, 137, 138, 139, and 157*

[CCO17] Mathieu Carrière, Marco Cuturi, and Steve Oudot. Sliced wasserstein kernel for persistence diagrams. In *34th International Conference on Machine Learning*, 2017. *Cited (8) on pages 16, 27, 44, 69, 134, 138, 139, and 140*

[CCSG+09a] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J Guibas, and Steve Y Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the*

twenty-fifth annual symposium on Computational geometry, pages 237–246. ACM, 2009. *Cited (2) on pages 34 and 212*

[CCSG+09b] Frédéric Chazal, David Cohen-Steiner, Leonidas J Guibas, Facundo Mémoli, and Steve Y Oudot. Gromov-hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403. Wiley Online Library, 2009. *Cited on page 217*

[CD14] Marco Cuturi and Arnaud Doucet. Fast computation of wasserstein barycenters. In *International Conference on Machine Learning*, pages 685–693, 2014. *Cited (8) on pages 50, 112, 114, 115, 116, 124, 125, and 163*

[CD18] Frédéric Chazal and Vincent Divol. The Density of Expected Persistence Diagrams and its Kernel Based Estimation. In Bettina Speckmann and Csaba D. Tóth, editors, *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. *Cited (2) on pages 43 and 146*

[CDSGO16] Frédéric Chazal, Vin De Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. Springer, 2016. *Cited (4) on pages 32, 211, 212, and 216*

[CDSO14] Frédéric Chazal, Vin De Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, 2014. *Cited on page 139*

[CE10] G. Carlier and I. Ekeland. Matching for teams. *Economic Theory*, 42(2):397–418, 2010. *Cited on page 92*

[CFG+15] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015. *Cited (2) on pages 9 and 20*

[CFL+14] Frédéric Chazal, Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, and Larry Wasserman. Stochastic convergence of persistence landscapes and silhou-

ettes. In *Proceedings of the thirtieth annual symposium on Computational geometry*, page 474. ACM, 2014. *Cited (4) on pages 132, 134, 138, and 150*

[CFL+15] Frédéric Chazal, Brittany Fasy, Fabrizio Lecci, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Subsampling methods for persistent homology. In *International Conference on Machine Learning*, pages 2143–2151, 2015. *Cited on page 149*

[CFP01] Francesca Cagliari, Massimo Ferri, and Paola Pozzi. Size functions from a categorical viewpoint. *Acta Applicandae Mathematica*, 67(3):225–235, 2001. *Cited on page 22*

[CHF12] Matthew A Cleveland, John M Hickey, and Selma Forni. A common dataset for genomic analysis of livestock populations. *G3: Genes, Genomes, Genetics*, 2(4):429–435, 2012. *Cited (2) on pages 9 and 20*

[Chi17] Lenaic Chizat. *Transport Optimal de mesures positives: modèles, méthodes numériques, applications.* PhD thesis, Université Paris-Dauphine, 11 2017. *Cited on page 51*

[Chi19] Lenaic Chizat. Sparse optimization on measures with over-parameterized gradient descent. *arXiv preprint arXiv:1907.10300*, 2019. *Cited on page 159*

[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*, pages 232–241. ACM, 2008. *Cited on page 157*

[COO15a] Guillaume Carlier, Adam Oberman, and Edouard Oudet. Numerical methods for matching for teams and wasserstein barycenters. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1621–1642, 2015. *Cited (3) on pages 92, 104, and 114*

[COO15b] Mathieu Carrière, Steve Y. Oudot, and Maks Ovsjanikov. Stable topological signatures for points on 3d shapes. *Computer Graphics Forum*, 34(5):1–12, 2015. *Cited on page 128*

[CP16] Marco Cuturi and Gabriel Peyré. A smoothed dual approach for variational wasserstein problems. *SIAM Journal on Imaging Sciences*, 9(1):320–343, 2016. *Cited on page 50*

[CPSV15] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: geometry and kantorovich formulation. *arXiv preprint arXiv:1508.05216*, 2015. *Cited on page 51*

[CSEH07] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007. *Cited on page 34*

[CSEHM10] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko. Lipschitz functions have l p-stable persistence. *Foundations of computational mathematics*, 10(2):127–139, 2010. *Cited (3) on pages 36, 147, and 148*

[Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013. *Cited (2) on pages 54 and 123*

[CWRW15] Yen-Chi Chen, Daren Wang, Alessandro Rinaldo, and Larry Wasserman. Statistical analysis of persistence intensity functions. *arXiv preprint arXiv:1510.02502*, 2015. *Cited on page 134*

[DB16] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016. *Cited on page 164*

[DGK18] Pavel Dvurechensky, Alexander Gasnikov, and Alexey Kroshnin. Computational optimal transport: Complexity by accelerated gradient descent is better than by sinkhorn's algorithm. *arXiv preprint arXiv:1802.04367*, 2018. *Cited on page 118*

[DL19] Vincent Divol and Théo Lacombe. Understanding the topology and the geometry of the persistence diagram space via optimal partial transport. *arXiv preprint arXiv:1901.03048*, 2019. *Cited (5) on pages 18, 19, 28, 29, and 146*

[DM12] Gianni Dal Maso. *An introduction to Γ-convergence*, volume 8. Springer Science & Business Media, 2012. *Cited on page 174*

[EH10] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. American Mathematical Soc., 2010. *Cited (2) on pages 31 and 71*

[FC17] R'emi Flamary and Nicolas Courty. Pot python optimal transport library, 2017. *Cited (3) on pages 156, 166, and 175*

[FG10] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *Journal de mathématiques pures et appliquées*, 94(2):107–130, 2010. *Cited (12) on pages 52, 64, 65, 67, 76, 77, 80, 83, 84, 88, 89, and 101*

[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. *Cited on page 153*

[Fig10] Alessio Figalli. The optimal partial transport problem. *Archive for rational mechanics and analysis*, 195(2):533–560, 2010. *Cited on page 51*

[Fro92] Patrizio Frosini. Discrete computation of size functions. 1992. *Cited on page 22*

[FSV+18] Jean Feydy, Thibault Séjourné, François-Xavier Vialard, Shun-Ichi Amari, Alain Trouvé, and Gabriel Peyré. Interpolating between optimal transport and mmd using sinkhorn divergences. *arXiv preprint arXiv:1810.08278*, 2018. *Cited (3) on pages 55, 56, and 57*

[Gab72] Peter Gabriel. Unzerlegbare darstellungen i. *Manuscripta mathematica*, 6(1):71–103, 1972. *Cited on page 211*

[GBR+12] Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(Mar):723–773, 2012. *Cited on page 57*

[Gen19] Aude Genevay. *Entropy-regularized optimal transport for machine learning*. PhD thesis, 2019. *Cited (2) on pages 55 and 172*

[GPC18] Aude Genevay, Gabriel Peyre, and Marco Cuturi. Learning generative models with sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617, 2018. *Cited (2) on pages 55 and 57*

[GPRS19] Thibaut Le Gouic, Quentin Paris, Philippe Rigollet, and Austin J Stromme. Fast convergence of empirical barycenters in alexandrov spaces and the wasserstein space. *arXiv preprint arXiv:1908.00828*, 2019. *Cited (2) on pages 50 and 180*

[GUD15] GUDHI Project. *GUDHI User and Reference Manual*. GUDHI Editorial Board, 2015. *Cited on page 19*

[Hal86] Marshall Hall. *Combinatorial theory (2nd ed)*. John Wiley & Sons, 1986. *Cited on page 73*

[HKN19] Christoph D. Hofer, Roland Kwitt, and Marc Niethammer. Learning representations of persistence barcodes. *Journal of Machine Learning Research*, 20(126):1–45, 2019. *Cited (2) on pages 136 and 137*

[HKNU17] Christoph Hofer, Roland Kwitt, Marc Niethammer, and Andreas Uhl. Deep learning with topological signatures. In *Advances in Neural Information Processing Systems*, pages 1634–1644, 2017. *Cited (3) on pages 138, 141, and 150*

[HSW07] Jan-Martin Hertzsch, Rob Sturman, and Stephen Wiggins. Dna microarrays: design principles for maximizing ergodic, chaotic mixing. *Small*, 3(2):202–218, 2007. *Cited on page 139*

[Kal17] Olav Kallenberg. *Random measures, theory and applications*. Springer, 2017. *Cited (2) on pages 149 and 219*

[Kan42] L Kantorovich. On the transfer of masses (in russian). In *Doklady Akademii Nauk*, volume 37, pages 227–229, 1942. *Cited on page 45*

[Kec95] A.S. Kechris. *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Springer-Verlag, 1995. *Cited on page 151*

[KFH17] Genki Kusano, Kenji Fukumizu, and Yasuaki Hiraoka. Kernel method for persistence diagrams via kernel embedding and weight factor. *The Journal of Machine Learning Research*, 18(1):6947–6987, 2017. *Cited on page 44*

[KH+09] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009. *Cited (2) on pages 9 and 20*

[KHF16] Genki Kusano, Yasuaki Hiraoka, and Kenji Fukumizu. Persistence weighted gaussian kernel for topological data analysis. In *International Conference on Machine Learning*, pages 2004–2013, 2016. *Cited (3) on pages 134, 138, and 140*

[KHN+15] Roland Kwitt, Stefan Huber, Marc Niethammer, Weili Lin, and Ulrich Bauer. Statistical topological data analysis-a kernel perspective. In *Advances in neural information processing systems*, pages 3070–3078, 2015. *Cited on page 69*

[KKM+16] Kristian Kersting, Nils Kriege, Christopher Morris, Petra Mutzel, and Marion Neumann. Benchmark data sets for graph kernels, 2016. *Cited (2) on pages 9 and 20*

[KMN17] Michael Kerber, Dmitriy Morozov, and Arnur Nigmetov. Geometry helps to compare persistence diagrams. *Journal of Experimental Algorithmics (JEA)*, 22(1):1–4, 2017. *Cited (2) on pages 71 and 116*

[KMV+16] Stanislav Kondratyev, Léonard Monsaingeon, Dmitry Vorotnikov, et al. A new optimal transport distance on the space of finite radon measures. *Advances in Differential Equations*, 21(11/12):1117–1164, 2016. *Cited on page 51*

[KVT19] Max Kontak, Jules Vidal, and Julien Tierny. Statistical parameter selection for clustering persistence diagrams. In *2019 IEEE/ACM HPC for Urgent Decision Making (UrgentHPC)*, pages 7–12. IEEE, 2019. *Cited on page 116*

[LCO18] Théo Lacombe, Marco Cuturi, and Steve Oudot. Large scale computation of means and clusters for persistence diagrams using optimal transport. In *Advances in Neural Information Processing Systems*, 2018. *Cited (4) on pages 19, 29, 71, and 109*

[Léo13] Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *arXiv preprint arXiv:1308.0215*, 2013. *Cited on page 54*

[LGL16] Thibaut Le Gouic and Jean-Michel Loubes. Existence and consistency of wasserstein barycenters. *Probability Theory and Related Fields*, pages 1–17, 2016. *Cited (4) on pages 49, 93, 94, and 96*

[Luc10] Yves Lucet. What shape is your conjugate? a survey of computational convex analysis and its applications. *SIAM review*, 52(3):505–542, 2010. *Cited (2) on pages 119 and 124*

[LY18] Tam Le and Makoto Yamada. Persistence Fisher kernel: a Riemannian manifold kernel for persistence diagrams. In *Advances in Neural Information Processing Systems*, pages 10027–10038, 2018. *Cited (4) on pages 44, 138, 139, and 140*

[MKPY16] Nikolay Makarenko, Maksat Kalimoldayev, Ivan Pak, and Ainur Yessenaliyeva. Texture recognition by the methods of topological data analysis. *Open Engineering*, 6(1), 2016. *Cited on page 125*

[MMH11] Yuriy Mileyko, Sayan Mukherjee, and John Harer. Probability measures on the space of persistence diagrams. *Inverse Problems*, 27(12):124007, 2011. *Cited (6) on pages 38, 41, 42, 67, 68, and 96*

[Mon84] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais*. Histoire de l'Académie Royale des Sciences de Paris, 1784. *Cited on page 45*

[Mor40] Marston Morse. Rank and span in functional topology. *Annals of Mathematics*, pages 419–454, 1940. *Cited on page 22*

[Mor65] Jean-Jacques Moreau. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société mathématique de France*, 93:273–299, 1965. *Cited on page 119*

[Mun84] James R Munkres. *Elements of algebraic topology*, volume 2. Addison-Wesley Menlo Park, 1984. *Cited (4) on pages 12, 22, 201, and 204*

[Nes05] Yu Nesterov. Smooth minimization of non-smooth functions. *Mathematical programming*, 103(1):127–152, 2005. *Cited on page 159*

[Nie11] Lance Nielsen. Weak convergence and banach space-valued functions: improving the stability theory of feynman's operational calculi. *Mathematical Physics, Analysis and Geometry*, 14(4):279–294, 2011. *Cited on page 133*

[OHK18] Ippei Obayashi, Yasuaki Hiraoka, and Masao Kimura. Persistence diagrams with linear machine learning models. *Journal of Applied and Computational Topology*, 1(3-4):421–449, 2018. *Cited on page 125*

[Oud15] Steve Y Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society, 2015. *Cited on page 31*

[OY19] Ippei Obayashi and Michio Yoshiwaki. Field choice problem in persistent homology. *arXiv preprint arXiv:1911.11350*, 2019. *Cited on page 204*

[PC17] Gabriel Peyré and Marco Cuturi. *Computational Optimal Transport*. Number 2017-86. December 2017. *Cited (7) on pages 43, 45, 55, 71, 116, 119, and 171*

[PMK19] Jose A Perea, Elizabeth Munch, and Firas A Khasawneh. Approximating continuous functions on persistence diagrams using template functions. *arXiv preprint arXiv:1902.07190*, 2019. *Cited on page 76*

[Pra07] Aldo Pratelli. On the equality between monge's infimum and kantorovich's minimum in optimal mass transportation. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 43, pages 1–13. Elsevier, 2007. *Cited on page 48*

[PWZ+17] Shirui Pan, Jia Wu, Xingquan Zhu, Guodong Long, and Chengqi Zhang. Task sensitive feature exploration and learning for multitask graph classification. *IEEE transactions on cybernetics*, 47(3):744–758, 2017. *Cited (2) on pages 9 and 20*

[RHBK15] Jan Reininghaus, Stefan Huber, Ulrich Bauer, and Roland Kwitt. A stable multi-scale kernel for topological machine learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4741–4748, 2015. *Cited (5) on pages 44, 134, 138, 140, and 150*

[Rob99] Vanessa Robins. Towards computing homology from finite approximations. In *Topology proceedings*, volume 24, pages 503–532, 1999. *Cited on page 22*

[RSV18] Frédéric Le Roux, Sobhan Seyfaddini, and Claude Viterbo. Barcodes and area-preserving homeomorphisms. *arXiv preprint arXiv:1810.03139*, 2018. *Cited on page 168*

[RTC17] Aaditya Ramdas, Nicolás García Trillos, and Marco Cuturi. On wasserstein two-sample testing and related families of nonparametric tests. *Entropy*, 19(2):47, 2017. *Cited (2) on pages 55 and 57*

[San15] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 2015. *Cited (4) on pages 45, 71, 78, and 119*

[Sch31] Erwin Schrödinger. *Über die umkehrung der naturgesetze*. Verlag der Akademie der Wissenschaften in Kommission bei Walter De Gruyter u . . . , 1931. *Cited on page 54*

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003. *Cited (2) on pages 104 and 105*

[SDGP+15] Justin Solomon, Fernando De Goes, Gabriel Peyré, Marco Cuturi, Adrian Butscher, Andy Nguyen, Tao Du, and Leonidas Guibas. Convolutional wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Transactions on Graphics (TOG)*, 34(4):66, 2015. *Cited (2) on pages 50 and 121*

[SFV+19] Thibault Séjourné, Jean Feydy, François-Xavier Vialard, Alain Trouvé, and Gabriel Peyré. Sinkhorn divergences for unbalanced optimal transport. *arXiv preprint arXiv:1910.12958*, 2019. *Cited (2) on pages 55 and 58*

[SKC18] Don Sheehy, Oliver Kisielius, and Nicholas J Cavanna. Computing the shift-invariant bottleneck distance for persistence diagrams. In *CCCG*, pages 78–84, 2018. *Cited on page 169*

[SP04] Robert W Sumner and Jovan Popović. Deformation transfer for triangle meshes. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 399–405. ACM, 2004. *Cited on page 128*

[SP18] Shashank Singh and Barnabás Póczos. Minimax distribution estimation in wasserstein distance. *arXiv preprint arXiv:1802.08855*, 2018. *Cited on page 154*

[STNR+18] Anirudh Som, Kowshik Thopalli, Karthikeyan Natesan Ramamurthy, Vinay Venkataraman, Ankita Shukla, and Pavan Turaga. Perturbation robust representations of topological persistence diagrams. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 617–635, 2018. *Cited on page 134*

[TMMH14] Katharine Turner, Yuriy Mileyko, Sayan Mukherjee, and John Harer. Fréchet means for distributions of persistence diagrams. *Discrete & Computational Geometry*, 52(1):44–70, 2014. *Cited (11) on pages 38, 41, 42, 71, 110, 112, 113, 115, 126, 127, and 128*

[Tur13] Katharine Turner. Means and medians of sets of persistence diagrams. *arXiv preprint arXiv:1307.8300*, 2013. *Cited (2) on pages 42 and 96*

[TVH18] Quoc Hoan Tran, Van Tuan Vo, and Yoshihiko Hasegawa. Scale-variant topological information for characterizing complex networks. *arXiv preprint arXiv:1811.03573*, 2018. *Cited (2) on pages 141 and 143*

[Ume17] Yuhei Umeda. Time series classification via topological data analysis. *Information and Media Technologies*, 12:228–239, 2017. *Cited on page 132*

[VBT19] Jules Vidal, Joseph Budin, and Julien Tierny. Progressive wasserstein barycenters of persistence diagrams. *IEEE transactions on visualization and computer graphics*, 26(1):151–161, 2019. *Cited on page 116*

[Vil08] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008. *Cited (9) on pages 45, 47, 48, 49, 67, 76, 87, 88, and 89*

[VZ17] Saurabh Verma and Zhi-Li Zhang. Hunt for the unique, stable, sparse and fast feature learning on graphs. In *Advances in Neural Information Processing Systems*, pages 88–98, 2017. *Cited (2) on pages 141 and 143*

[Wag19] Alexander Wagner. Nonembeddability of persistence diagrams with $p > 2$ wasserstein metric. *arXiv preprint arXiv:1910.13935*, 2019. *Cited (2) on pages 16 and 27*

[WCP13] Weiran Wang and Miguel A Carreira-Perpinán. Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application. *arXiv preprint arXiv:1309.1541*, 2013. *Cited on page 163*

[XC19] Zhang Xinyi and Lihui Chen. Capsule graph neural network. In *International Conference on Learning Representations*, 2019. *Cited (2) on pages 141 and 143*

[XHLJ19] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *ICLR 2019*, 2019. *Cited (2) on pages 141 and 143*

[YV15] Pinar Yanardag and S.V.N. Vishwanathan. Deep graph kernels. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 1365–1374, New York, NY, USA, 2015. ACM. *Cited (3) on pages 9, 20, and 141*

[Z+00] August M Zapała et al. Jensen's inequality for conditional expectations in banach spaces. *Real Analysis Exchange*, 26(2):541–552, 2000. *Cited on page 152*

[ZC05] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005. *Cited on page 22*

[ZKR+17] Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander Smola. Deep sets. In *Advances in Neural Information Processing Systems*, pages 3391–3401, 2017. *Cited on page 137*

[ZWX+18] Zhen Zhang, Mianzhi Wang, Yijian Xiang, Yan Huang, and Arye Nehorai. RetGK: Graph Kernels based on Return Probabilities of Random Walks. In *Advances in Neural Information Processing Systems*, pages 3968–3978, 2018. *Cited (2) on pages 141 and 143*

[ZZJS16] Matthias Zeppelzauer, Bartosz Zieliński, Mateusz Juda, and Markus Seidl. Topological descriptors for 3d surface analysis.

In *International Workshop on Computational Topology in Image Context*, pages 77–87. Springer, 2016. *Cited on page 125*

Index

Appendix A

Complementary notions of homology and persistence theory

This chapter recalls the algebraic root of topological data analysis: *homology* theory, which gives a formal meaning to the notions of "topological properties". It then presents (one-dimensional) persistence theory, the modern bridge between homology and data science, and its main tool: the *persistence diagram*.

A.1 Homology theory

For the sakes of simplicity and concision, we focus in this section on *simplicial homology*, which is the version of homology used in applications. We briefly mention at the end of this section an extension to more general objects. We refer the interested reader to [Mun84] for more details.

Simplicial complexes

Simplicial homology is the formal way to describe the topology of a simplicial complex, a generalization of the notion of graph that can be conveniently encoded in a computer and used to make computations in applications.

Definition A.1. Let $E = \{v_0 \dots v_N\}$ be a finite set. A *simplex* on E is an element of $\mathcal{P}(E)$, whose elements are called its *vertices*. A *simplicial complex* X on E is a (non-empty) subset of $\mathcal{P}(E)$ that satisfies the following:

$$\forall \sigma \in X, \forall \tau \subset \sigma \text{ with } \tau \neq \emptyset, \text{ we have } \tau \in X.$$

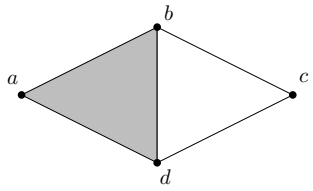


Figure A.1: An example of simplicial complex.

The *dimension* of a simplex σ is $\dim(\sigma) := \text{card}(\sigma) - 1$. Simplices of dimension 0 will be called *points* or *vertices*, those of dimension 1 are called *edges*, dimension 2 and above are generally called *faces*. The dimension of a simplicial complex is the largest dimension of one of the simplices it contains. The k-simplices of X are the simplices of dimension k in X, and we write X_k for the set of k-simplices of X.

Note that simplicial complexes are purely combinatorial structures: they do not need do be embedded in some (Euclidean) space.

We now introduce the notion of *orientation of a simplex*. Intuitively, it formalizes the idea that an edge $\{a, b\}$ can be oriented either "from a to b", denoted by $[a, b]$, or "from b to a", denoted by $[b, a]$.

Definition A.2 (Orientation)**.** Let $\sigma = \{v_0 \dots v_k\}$ be a k-simplex. Two orderings $(v_i)_{i=1}^k$ of σ are equivalent if they differ from one another by an even permutation. It defines two equivalence classes, called the *orientations* of σ. We will write $\sigma = [v_0 \dots v_k]$ to specify an orientation of σ.

From now on, we will implicitly assume that our simplices are oriented.

Example A.1. The complex represented in Figure A.1 is defined by

$$\begin{aligned} X = \{&\{a\}, \{b\}, \{c\}, \{d\}, && \text{simplices of dim 0} \\ &\{ab\}, \{ad\}, \{bc\}, \{bd\}, \{cd\} && \text{simplices of dim 1} \\ &\{abd\}\} && \text{simplices of dim 2.} \end{aligned}$$

Simplicial homology

Computing the *simplicial homology* of a simplicial complex consists, roughly speaking, in encoding its topological features. Informally, topological features are connected components (features of "dimension 0"), loops (dimension 1), cavities (dimension 2), and so on... To formalize this intuition, we

need to put an algebraic structure on our simplicial complex that will allow us, both in theory and in practice, to identify these topological features.

Let X be a (finite) simplicial complex. Let $\mathbb{Z}$ denote the ring of integers.

Definition A.3 (k-chains)**.** The space of k-chains of a simplicial complex X is the free abelian group generated by the set of k-simplices $X_k = \{\sigma_1 \dots \sigma_{|X_k|}\}$:

$$C_k(X, \mathbb{Z}) := \left\{ \sum_{i=1}^{|X_k|} \alpha_i \sigma_i : \alpha_i \in \mathbb{Z} \right\}.$$

boundary operator. Now that we have a group structure on the $(X_k)_k$, we can build morphisms between these.

Definition A.4. Given a k-simplex $[v_0 \dots v_k] \in X_k$, we denote by $[v_0 \dots, \widehat{v_j}, \dots v_k] \in X_{k-1}$ the $(k-1)$-simplex (which is a face of $[v_0 \dots v_k]$) where the vertex v_j has been removed.

The *boundary operator* ∂_k for $k \geqslant 1$ over k-dimensional chains is defined as the morphism:

$$\begin{aligned} \partial_k : C_k(X, \mathbb{Z}) &\to C_{k-1}(X, \mathbb{Z}) \\ \sigma = [v_0 \dots v_k] &\mapsto \sum_{j=0}^{k} (-1)^j \underbrace{[v_0 \dots \widehat{v_j}, \dots, v_k]}_{\in X_{k-1}} \\ (\sigma + \lambda\sigma') &\mapsto \partial_k \sigma + \lambda \partial_k \sigma' \end{aligned}$$

By convention, $\partial_0 = 0$.

Note that $\forall k > 0$ we have $\partial_{k-1} \circ \partial_k = 0$, that is, the boundary of a boundary is zero.

Homology groups. We have the following sequence[1] of applications:

$$C_k(X, \mathbb{Z}) \xrightarrow{\partial_k} C_{k-1}(X, \mathbb{Z}) \xrightarrow{\partial_{k-1}} C_{k-2}(X, \mathbb{Z}) \xrightarrow{\partial_{k-2}} \dots$$

We define:

- *k-cycles* : $Z_k := \ker(\partial_k) = \{\sigma \in C_k,\ \partial_k \sigma = 0\}$

[1] Called a *chain complex*, as if $A \xrightarrow{f} B \xrightarrow{g} C$, we have $\mathrm{Im}(f) \subset \ker(g)$.

- *k-boundaries* : $B_k := \mathrm{Im}(\partial_{k+1}) = \{\partial_{k+1}\sigma,\ \sigma \in C_{k+1}\}$

As $\partial_{k-1} \circ \partial_k = 0$, we know that $\mathrm{Im}(\partial_{k+1}) \subset \ker(\partial_k)$, which states that boundaries of $(k+1)$-simplices are k-cycles. It allows us to define the following quotient.

Definition A.5. The k-th *homology group* of X is

$$H_k(X, \mathbb{Z}) := \frac{\ker(\partial_k)}{\mathrm{Im}(\partial_{k+1})}. \tag{A.1}$$

This definition can be interpreted in the following way: the homology group of a complex X represents the "space of its cycles modulo its boundaries".

The collection of the homology groups of X is denoted by $H_*(X, \mathbb{Z})$.

Remark A.1. In practice, simplicial homology is computed by replacing the ring $\mathbb{Z}$ by a finite field $\mathbb{K}$ such as $\mathbb{Z}/p\mathbb{Z}$ for some prime number p (typically $p = 2$). It follows that $C_k(X, \mathbb{K})$ and thus $H_k(X, \mathbb{K})$ have a *finite-dimensional* vector space structure. In this context, the dimension β_k of $H_k(X, \mathbb{K})$ is called the k-th Betti number of X (for the field of coefficients $\mathbb{K}$). A basis $v_1 \dots v_{\beta_k}$ of $H_k(X, \mathbb{K})$, which are (representing) linear combinations of elements in X, can be understood as a maximal family of independent topological features of dimension k. In particular, β_k encodes the number of such topological features.

From a theoretical perspective, the choice of the field of coefficients $\mathbb{K}$ has some consequence on the computation of $H_k(X, \mathbb{K})$ only when $H_k(X, \mathbb{Z})$ has *torsion*, see [Mun84] for more details. In applications however, it seems that the choice of $\mathbb{K}$ has barely any impact [OY19].

Invariance properties. We briefly discuss here why homology is a relevant descriptor to describe the topology of structured objects. First, let us introduce a notion of equivalence between topological spaces.

Definition A.6. Let $\mathcal{X}$ and $\mathcal{Y}$ be two topological spaces. Two continuous maps $u, v : \mathcal{X} \to \mathcal{Y}$ are *homotopic* if there exists $\varphi : [0,1] \times \mathcal{X} \to \mathcal{Y}$ continuous, such that $\varphi(0, \cdot) = u$ and $\varphi(1, \cdot) = v$. This will be denoted by $u \sim v$.

Let $\mathcal{X}$ and $\mathcal{Y}$ be two topological spaces. They are said to have the same *homotopy type* if there exist $f : \mathcal{X} \to \mathcal{Y}$ and $g : \mathcal{Y} \to \mathcal{X}$ such as $f \circ g \sim \mathrm{id}_Y, g \circ f \sim \mathrm{id}_X$. We will denote this by $\mathcal{X} \sim \mathcal{Y}$.

Let X be a (finite) simplicial complex. A *geometric realization* of X is a map $\psi : X \to \mathbb{R}^{2n+1}$ such that for any simplex $\sigma \in X$, with $\sigma = \{v_1 \dots v_k\}$, one has $\psi(\sigma) = \text{conv}(\{\psi(v_1) \dots \psi(v_k)\})$ where the $(\psi(v_i))_i$ must be affinely independent and $\text{conv}(A)$ denotes the convex hull of a set A, and such that for any $\sigma_1, \sigma_2 \in X$, one has $\psi(\sigma_1) \cap \psi(\sigma_2) = \psi(\sigma_1 \cap \sigma_2)$ (with $\psi(\emptyset) = \emptyset$). A simplicial complex X of dimension n always admits a geometric realization in the Euclidean space $\mathbb{R}^{2n+1}$ (by Whitney's Embedding theorem).

The important result to note is the following one:

Theorem. *Let X, Y be two simplicial complexes, and $r(X), r(Y)$ be two geometric realizations of those. If $r(X) \sim r(Y)$ then $H_*(X, \mathbb{K}) = H_*(Y, \mathbb{K})$.*

This theorem can be rephrased by stating that two complexes that have isomorphic homotopy type have the same homology groups, or equivalently that homology is a homotopy invariant.

Singular homology. We end this section by mentioning that homology theory can be extended to a wider class of objects than just simplicial complexes. Namely, given a topological space $\mathcal{X}$, we can say that $\Sigma \subset \mathcal{X}$ is a *singular simplex* of X of dimension k if it is the image of the geometric realization of a (oriented) k-simplex $\text{conv}(\{v_0 \dots v_k\})$ by a continuous map (Figure A.2). All the definitions and properties stated above can be adapted to general topological spaces: the space of k-chains is the free abelian group generated by the (oriented) singular simplices of X, and all the subsequent definitions (boundary operator, cycles, boundaries, homology groups, homotopy invariance) extend straightforwardly. Note that the (simplicial) homology groups of X and the (singular) homology groups of a geometric realization $r(X)$ do coincide, making singular homology a satisfactory extension of simplicial homology to general topological spaces. Note that singular homology is also a homotopy invariant.

A.2 Filtrations and persistence modules

Although simplicial homology allows us to describe the topological features of a given object, it cannot be used faithfully in statistical and machine learning applications, for several reasons:

- The topology of a simplicial complex is based on a purely combinatorial representation of the object, without (explicitly) taking the extra information available (such as the geometry of the ambient space) into account.

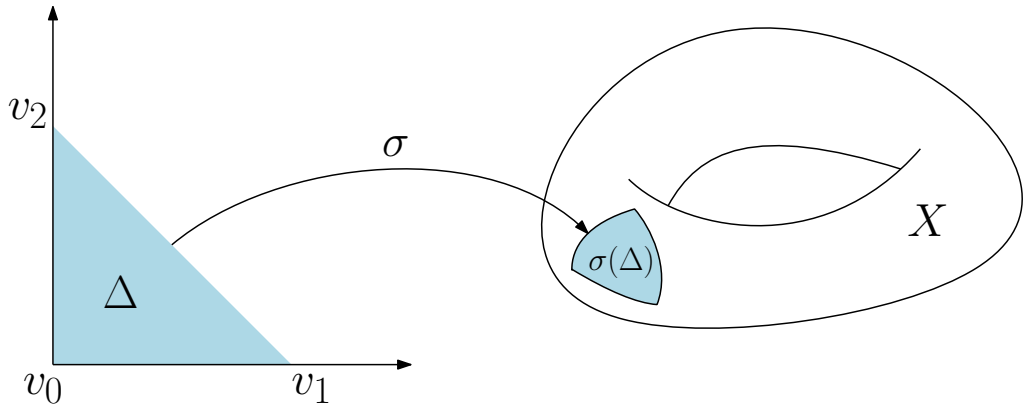


Figure A.2: *(Left)* A geometric realization Δ of the 2-simplex in $\mathbb{R}^2$. *(Right)* A torus X and a singular simplex (the image of Δ by a continuous map σ).

- The homology groups of a point cloud $P \subset \mathbb{R}^d$ with n points will simply be given by $H_0(P, \mathbb{Z}) = \mathbb{Z}^n$ and $H_k(P, \mathbb{Z}) = 0$ for $k \geq 1$ (n connected components and nothing else); even if the points are perfectly distributed on, say, a circle.

- The information encoded by homology might not be quantitative enough. Although $H_1(X, \mathbb{Z})$ might account for the presence of loops, for instance, we do not have access to some quantitative information about these loops: are there some loops more significant, or more reliable, than others?

To improve on these, we will rely on *persistence theory*: namely we will study the evolution of the homology groups of an object at various scales using a *filtration*.

Preliminary definitions

Definition A.7. Let $\mathcal{X}$ be a topological space and $T \subset \mathbb{R}$. A *filtration* $(\mathcal{F}_t)_{t \in T}$ is a familly of topological spaces such that

$$\forall (s,t) \in T^2, \quad s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t.$$

A standard way to build filtration is to look at the *sublevel sets* of a map $f : \mathcal{X} \to \mathbb{R}$, i.e. the subsets of $\mathcal{X}$ of the form

$$\mathcal{F}_t := f^{-1}((-\infty, t]).$$

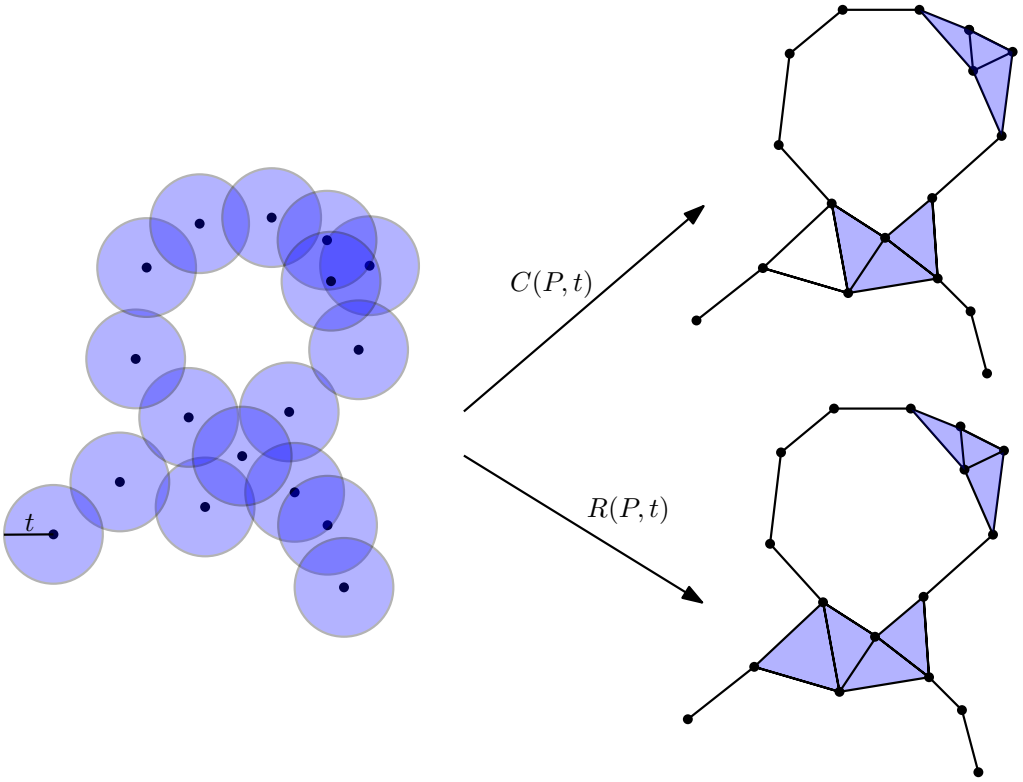


Figure A.3: Example of Čech and Rips complexes built on a point cloud.

Example A.2. Let $P = \{x_1 \dots x_n\}$ be a point cloud in $\mathbb{R}^d$. The *offsets filtration* on P is the filtration described by the map

$$f = \mathrm{dist}_P : x \mapsto \min_{1 \leq i \leq n} \|x - x_i\|_2.$$

Here, one has

$$\mathcal{F}_t = \bigcup_{i=1}^{n} B(x_i, t),$$

where $B(x_i, t)$ is the Euclidean ball centered in x_i of radius t for $t \geq 0$, set to be the empty set if $t < 0$. See Figure A.3 for an example. Note that in practice, computing the homology groups of the space $\mathcal{F}_t$ directly is not handy as algorithms take simplicial complexes as input. In this particular case however, one has two natural options to build a simplicial complex based on $\mathcal{F}_t$:

- The *Čech complex* $C(P,t)$ of parameter t is given, for $t \geq 0$, by

$$\{x_{i_1} \dots x_{i_k}\} = \sigma \in C(P,t) \Leftrightarrow \bigcap_{j=0}^{k} B(x_{i_j}, t) \neq \emptyset,$$

  that is the balls have a common intersection. The *nerve theorem* ensures that $C(P,t)$ and $\mathcal{F}_t$ have the same homotopy type and thus the same homology groups (see Appendix A). Therefore, working in practice with the simplicial complexes $(C(P,t))_t$ allows us to exactly recover the topology of $(\mathcal{F}_t)_t$.

- Similarly, the *Rips complex* $R(P,t)$ of parameter t is given, for $t \geq 0$, by

$$\{x_{i_1} \dots x_{i_k}\} = \sigma \in R(P,t) \Leftrightarrow \forall j, j', B(x_{i_j}, t) \cap B(x_{i_{j'}}, t) \neq \emptyset,$$

  that is the balls have pairwise intersections. Note that we have for all $t \geq 0$:

$$C(P,t) \subset R(P,t) \subset C(P,2t). \tag{A.2}$$

  Furthermore, $R(P,t)$ is much simpler to compute than $C(P,t)$ as one has to check if the balls have pairwise intersections instead of common intersections.

Persistence modules

Now, fix a field $\mathbb{K}$. It turns out that $(H_*(\mathcal{F}_t, \mathbb{K}))_t$ is not a family of independent vector spaces but—thanks to the increasing property $s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t$—has much more structure, which is encoded as a *persistence module*.

Definition A.8. A *persistence module* over $T \subset \mathbb{R}$ is a family $\mathbb{V} = (V_t)_{t \in T}$ of $\mathbb{K}$-vector spaces, endowed with linear applications $v_t^{t'} : V_t \to V_{t'}$ such that:

$$\begin{aligned} &\forall t \in T, \quad v_t^t = id \\ &\forall t \leq t' \leq t'' \in T, \quad v_{t'}^{t''} \circ v_t^{t'} = v_t^{t''}. \end{aligned}$$

Note that the linear applications are generally made implicit in the notation, that is we will write $\mathbb{V} = (V_t)_t$ to denote a persistence module over a family of $\mathbb{K}$-vector spaces $(V_t)_t$ endowed with maps $(v_t^{t'})_{t \leq t'}$.

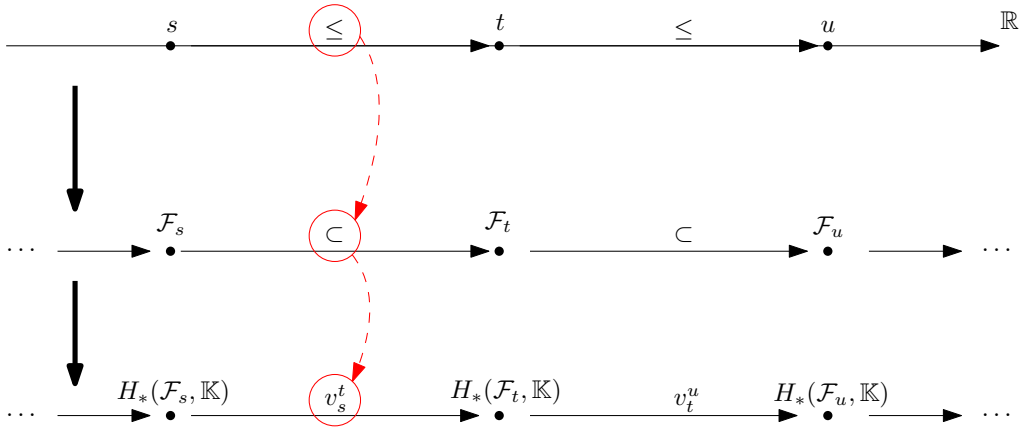


Figure A.4: Global picture of the algebraic pipeline behind the construction of persistence modules.

Persistence module of a simplicial filtration. Let X, Y, Z be three simplicial complexes such that $X \subset Y \subset Z$. Let i_X^Y and i_Y^Z denote the inclusion maps between (X, Y) and (Y, Z) respectively, along with $i_X^Z = i_Y^Z \circ i_X^Y$, the inclusion between X and Z. Let also H_X, H_Y, H_Z denote their respective homology groups. The inclusion maps are naturally turned into linear maps $v_{H_X}^{H_Y}, v_{H_Y}^{H_Z}, v_{H_X}^{H_Z}$ respectively, and one has $v_{H_Y}^{H_Z} \circ v_{H_X}^{H_Y} = v_{H_X}^{H_Z}$, see Figure A.5.

Thus, if we have $\mathcal{F} = (F_t)_t$ a filtration of simplicial complexes (called a simplicial filtration), we have:

- $\forall t \in T$, we define $V_t := H_*(F_t, \mathbb{K})$
- $\forall t \leq t'$, let $v_t^{t'}$ be the linear application induced by the canonical inclusion $F_t \hookrightarrow F_{t'}$.

This defines the (canonical) persistence module over $\mathcal{F}$. Without introducing the terminology of category theory, which is out of the scope covered by this manuscript, it can be useful to have the following global picture (Figure A.4) in mind: a real number s is turned into a topological space $\mathcal{F}_s$ which is then turned into a vector space $H_*(\mathcal{F}_s, \mathbb{K})$, while inequalities $s \leq t$ are turned into inclusions $\mathcal{F}_s \subset \mathcal{F}_t$ which are then turned into linear maps $v_s^t : H_*(\mathcal{F}_s, \mathbb{K}) \to H_*(\mathcal{F}_t, \mathbb{K})$.

Example A.3. Figure A.5 gives an example of an increasing sequence of simplicial complexes on which we derive the corresponding persistence module, that is $H_0(F_t) \oplus H_1(F_t), t \in \{1 \dots 6\}$. Linear applications v_t^{t+1} :

$H_0(F_t) \oplus H_1(F_t) \to H_0(F_{t+1}) \oplus H_1(F_{t+1})$ are represented as matrices at each step.

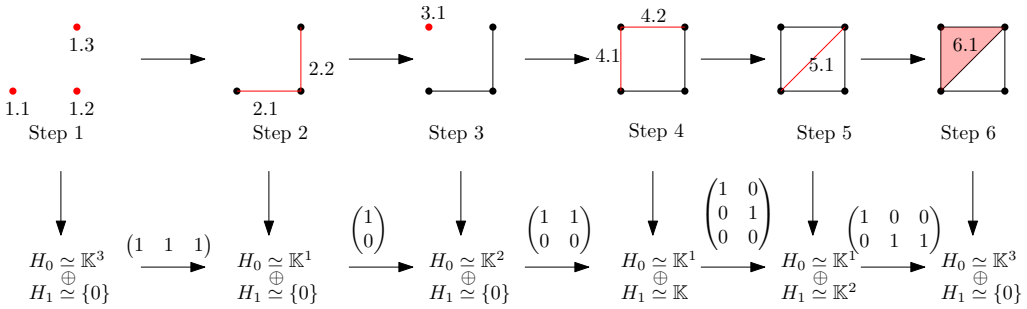


Figure A.5: An example of filtration over a simplicial complex and the corresponding persistence module.

A.3 Persistence diagrams

Decompositions of persistence modules

Fix $T \subset \mathbb{R}$ and a field $\mathbb{K}$.

Definition A.9. An *interval* of T is a subset $I \subset T$ such as:

$$\forall t \leq t' \leq t'' \in T, \quad t, t'' \in I \Rightarrow t' \in I$$

An *interval module* over an interval $I \subset T$ is a persistence module $\mathbb{V}$ defined by:

- $V_t = \mathbb{K}$ if $t \in I$, $V_t = \{0\}$ otherwise.
- $\forall t \leq t' \in T, v_t^{t'} = \mathrm{id}_{\mathbb{K}}$ if $t, t' \in I$, $v_t^{t'} = 0$ otherwise.

For $I = [b, d]$, the corresponding interval module is denoted by $\mathbb{I}_{[b,d]}$. It can be described as

$$\mathbb{I}_{[b,d]} := \underbrace{\{0\} \xrightarrow{0} \ldots \xrightarrow{0} \{0\}}_{t<b} \xrightarrow{0} \underbrace{\mathbb{K} \xrightarrow{\mathrm{id}} \ldots \xrightarrow{\mathrm{id}} \mathbb{K}}_{b \leq t \leq d} \xrightarrow{0} \underbrace{\{0\} \xrightarrow{0} \ldots \xrightarrow{0} \{0\}}_{d<t}$$

Analogous definitions and notations stand for $I = (a, b], (a, b), [a, b)$.

Interpretation. Interval modules are the "base blocks" of persistence modules, intuitively in the same way as prime numbers are the base block of natural integers, or (maybe more precisely) as real lines are the base blocks of real vector spaces. Just as any real vector space can be decomposed as a unique (up to isomorphism) direct sum of real lines, we will see that (under mild assumption), any persistence module can be uniquely (up to isomorphism) decomposed as a direct sum (see below) of interval modules.

Definition A.10. Given $\mathbb{V} = (V_t)_t, \mathbb{W} = (W_t)_t$ two persistence modules with linear applications $(v_t^{t'})_{t,t'}, (w_t^{t'})_{t,t'}$, we define the *direct sum of modules* $\mathbb{V} \oplus \mathbb{W}$ by:

- $\mathbb{V} \oplus \mathbb{W} = (V_t \oplus W_t)_t$
- The corresponding linear applications are denoted by:

$$\begin{aligned}(v \oplus w)_t^{t'} : V_t \oplus W_t &\to V_{t'} \oplus W_{t'} \\ (x, y) &\mapsto (v_t^{t'}(x), w_t^{t'}(y))\end{aligned}$$

This definition extends naturally to a family of persistence modules $(\mathbb{V}_j)_{j \in J}$, denoted by $\mathbb{V} := \bigoplus_{j \in J} \mathbb{V}_j$.

Note that $\mathbb{V} \oplus \mathbb{W}$ is also a persistence module (so is the sum over a family).

Theorem A.1 ([CDSGO16, Theorem 2.8])**.** *A persistence module $\mathbb{V}$ can be decomposed as a direct sum of interval modules, written as:*

$$\mathbb{V} \simeq \bigoplus \mathbb{I}_{[b_j, d_j]},$$

in the following sufficient cases:

1. *If T is finite [Gab72].*
2. *When all the vector spaces V_t are finite-dimensional [CB15].*

When it exists, the decomposition is unique (up to isomorphism and ordering of the terms).

When we have such a decomposition, the knowledge of the intervals $([b_j, d_j])_j$ gives a complete description of the structure of $\mathbb{V}$. This information can be encoded as a (multi-)set[2] of points, called the *persistence*

[2]That is, a set where elements can be repeated. The number of times an element x appears in a multiset is called its *multiplicity*.

diagram $\mathrm{Dgm}(\mathbb{V})$ of the persistence module, supported on

$$\overline{\mathbb{R}}^2_{\geq} := \{(b,d) \in \overline{\mathbb{R}}^2, \quad b \leq d\}, \tag{A.3}$$

where $\overline{\mathbb{R}}$ is the extended real line $\mathbb{R} \cup \{\pm\infty\}$. Each point in the persistence diagram accounts for an interval module in the decomposition of $\mathbb{V}$.

It might happen that $\mathbb{V}$ is not decomposable. However, one can still define a persistence diagram of $\mathbb{V}$ assuming that $\mathbb{V}$ is *q-tame*.

Definition A.11. A persistence module $\mathbb{V}$ is said to be *q-tame* if $\forall t < t' \in T$, $\mathrm{rank}(v_t^{t'})$ is finite.

Extending the notion of persistence diagram for a persistence module $\mathbb{V}$ that cannot be decomposed can be done using a measure-theoretic formalism. Intuitively, the idea is the following: when $\mathbb{V}$ can be decomposed, the number of points of $\mathrm{Dgm}(\mathbb{V})$ that fall into a rectangle $[a,b] \times [c,d] \subset \overline{\mathbb{R}}^2_{\geq}$ with $a < b \leq c < d$ corresponds to the number of intervals that start between a and b and end between c and d. It turns out to be equal to $\mathrm{rank}(v_a^d) - \mathrm{rank}(v_a^c) + \mathrm{rank}(v_b^c) - \mathrm{rank}(v_b^d)$.

Therefore, for a q-tame persistence module $\mathbb{V}$, let us define

$$\mu_{\mathbb{V}} : [a,b] \times [c,d] \mapsto \mathrm{rank}(v_a^d) - \mathrm{rank}(v_a^c) + \mathrm{rank}(v_b^c) - \mathrm{rank}(v_b^d),$$

which defines, modulo some technical details (see [CDSGO16, Chapter 3]), a measure on $\overline{\mathbb{R}}^2_{\geq}$. It turns out that this measure has some strong structure properties.

Theorem A.2 ([CCSG+09a, CDSGO16])**.** *Let $\mathbb{V}$ be a q-tame persistence module. Then $\mu_{\mathbb{V}}$ is a point measure, i.e. is integer-valued and supported on a locally finite set $X \subset \overline{\mathbb{R}}^2_{\geq}$. It can thus be written as*

$$\mu_{\mathbb{V}} = \sum_{x \in X} n_x \delta_x,$$

where δ_x denotes the Dirac mass located at $x \in \overline{\mathbb{R}}^2_{\geq}$, and $n_x \in \mathbb{N}$ denotes the multiplicity *of x.*

Furthermore, if $\mathbb{V}$ can be decomposed as $\mathbb{V} = \bigoplus_j \mathbb{I}_{[b_j,d_j]}$, then $\mu_{\mathbb{V}} = \mathrm{Dgm}(\mathbb{V})$, in sense that $\mu_{\mathbb{V}} = \sum_j \delta_{(b_j,d_j)}$.

This theorem is the first example in which adopting a measure-theoretic perspective provides a powerful framework to study topological descriptors. We will adopt this approach in chapters 3 and 4 and showcase its strength to deal with various statistical problems occurring in topological data analysis.

Computation

Let us now mention how persistence diagrams can be computed. In practice, algorithms take as input a *simplicial filtration*, that is a filtration over a simplicial complex K which satisfies:

- $T = \{0, 1, \dots, m\}$ (finite, so we have a decomposable persistence module).
- $K_0 = \emptyset, K_m = K$.
- $\forall t \in T, K_t$ is a finite simplicial complex, which is a sub-complex of K_{t+1}.

We can also assume that we actually only add one simplex at each step, that is $K_{t+1} \backslash K_t = \{\sigma_t\}$.

Persistence diagrams can be computed using Gaussian elimination, taking in addition the ordering of the simplices into account. The pipeline reads as follows

1. Let M be the matrix of the boundary operator ∂.
2. Define
$$\text{low}(j) := \begin{cases} \max\{i | M_{ij} \neq 0\} \\ 0 \text{ if } M_{ij} = 0 \text{ for all } i \end{cases}$$
3. Use Gaussian elimination from the left to the right (see Algorithm 7).

Algorithm 7 Compute the barcode corresponding to a simplicial filtration

for $j = 1 \dots m$ **do**
  while $\exists i < j$ s.t. $\text{low}(i) = \text{low}(j) \neq 0$ **do**
    $c_j \leftarrow c_j - \frac{M[\text{low}(i),j]}{M[\text{low}(i),i]} c_i$
  end while
end for
return The reduced matrix.

4. Building the diagram: after reduction each column with only 0 entries accounts for the birth of a cycle. Any other column j (with non 0) induces a boundary which trivializes (fills in) the cycle $i = \text{low}(j)$. Thus:

- Each column i with only 0 entries encodes the beginning of one interval module in the decomposition of the persistence module $\mathbb{I}_{[t(i),?]}$, where $t(i)$ denotes the insertion time of the corresponding simplex.
- Finite intervals are $\mathbb{I}_{[t(i),t(j)]}$ with $i = \text{low}(j)$
- Infinite intervals are $\mathbb{I}_{[t(i),+\infty)}$ when there is no j such that $i = \text{low}(j)$.

Note that in practice, we generally build one diagram for each homology dimension k.

Example A.4. Let $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$. The pipeline is illustrated in Figure A.6. After reduction, the column of e_5 is empty, it means that inserting e_5 created a cycle (of dimension $\dim(e_5) = 1$, that is a loop). As $\text{low}(\sigma_1) = e_5$, it means that inserting σ_1 at $t = 6$ killed (filled) the cycle created by e_5. It is recorded in the persistence diagram by adding a point with the corresponding insertion times $(5, 6)$. Similarly, e_4 created a cycle. As there is no σ such that $\text{low}(\sigma) = e_4$, the cycle did not get killed at the end of the filtration process, resulting in a point of coordinates $(4, +\infty)$ in the persistence diagram.

Thus, the decomposition of the persistence module is given by

$$H_0 \simeq \mathbb{I}_{[1,+\infty)} \oplus \mathbb{I}_{[1,2)} \oplus \mathbb{I}_{[1,2)} \oplus \mathbb{I}_{[3,4)}$$
$$H_1 \simeq \mathbb{I}_{[4,+\infty)} \oplus \mathbb{I}_{[5,6)},$$

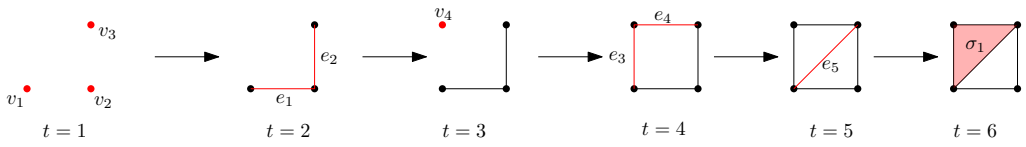
and the corresponding persistence diagrams in homology dimension $k = 0$ and $k = 1$ reads, adopting the measure notation,

$$\text{Dgm}_0 = \delta_{(1,+\infty)} + 2\delta_{(1,2)} + \delta_{(3,4)}$$
$$\text{Dgm}_1 = \delta_{(4,+\infty)} + \delta_{(5,6)}.$$

Interleaving and bottleneck distances

Using the bottleneck distance to compare persistence diagrams is motivated by algebraic considerations. The bridge between the bottleneck distance and persistence modules is made by re-interpreting this distance in terms of *interleaving*.

Definition A.12. Let $\varepsilon > 0$. Let $\mathbb{V} = (V_t)_{t\in\mathbb{R}}$ and $\mathbb{W} = (W_t)_{t\in\mathbb{R}}$ be two persistence modules on $\mathbb{R}$. $\mathbb{V}$ and $\mathbb{W}$ are said to be ε-*interleaved* if there are two families of applications $\varphi = (\varphi_t)_t, \psi = (\psi_t)_t$ where $\varphi_t : V_t \to W_{t+\varepsilon}$ and $\psi_t : W_t \to V_{t+\varepsilon}$ such that for all $t \leq t'$, the following diagrams commute:



Leading to (matrix reduction):

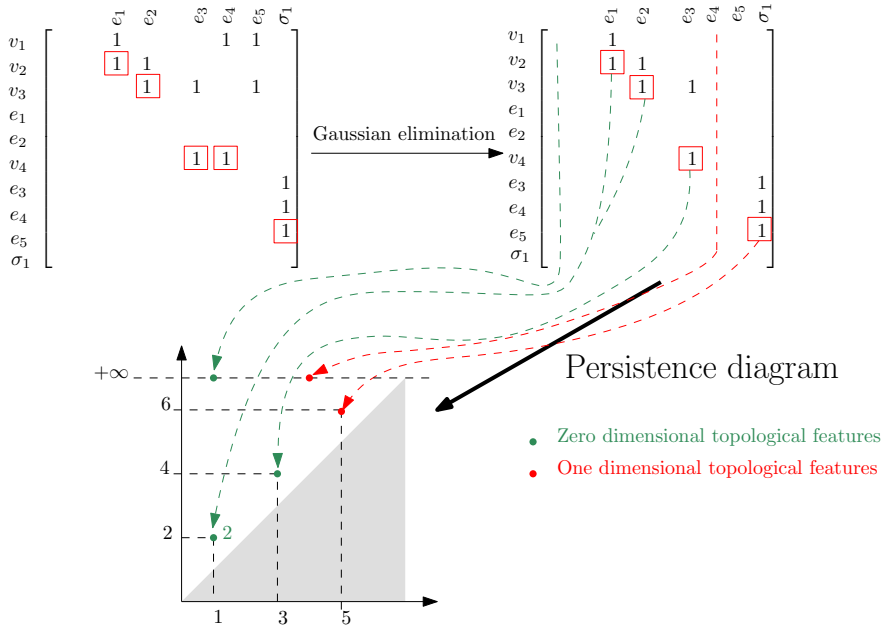


Figure A.6: Pipeline to build a persistence diagram from a simplicial filtration. *(Top)* The simplicial filtration. *(Middle)* On the left, the matrix of the boundary operator ∂. Note that simplices must be ordered with respect to their insertion time t. On the right, the matrix after reduction. *(Bottom)* The persistence diagram deduced from the reduced matrix.

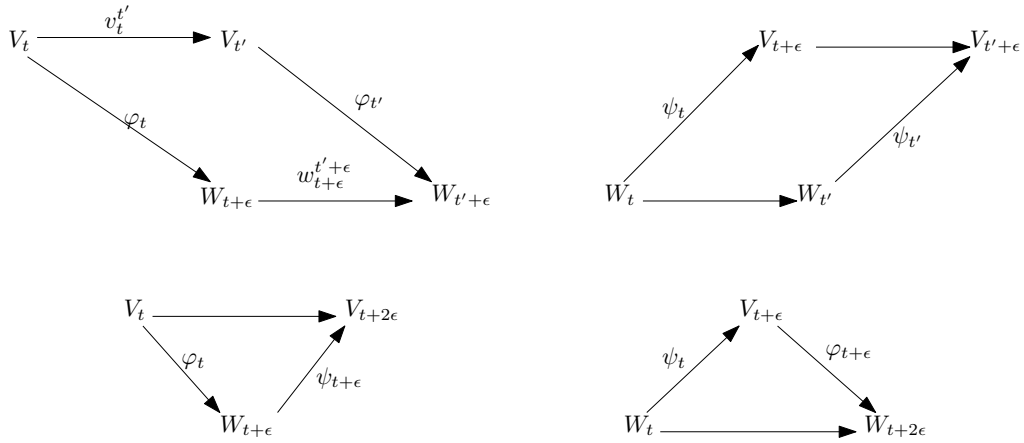


Figure A.7: Schematic representation of interleaving between persistence module.

Then, we define the *interleaving distance* as

$$d_i(\mathbb{V}, \mathbb{W}) := \inf\{\varepsilon \geqslant 0 | \mathbb{V}, \mathbb{W} \text{ are } \varepsilon\text{-interleaved } \} \in \mathbb{R}_+ \cup \{+\infty\}.$$

The bottleneck distance between persistence diagrams and the interleaving distance between persistence modules are strongly related by the following result.

Theorem A.3 ([CDSGO16, Theorem 5.14])**.** *If $\mathbb{V}, \mathbb{W}$ are q-tame, then*

$$\mathbf{d}_\infty(\mathrm{Dgm}(\mathbb{V}), \mathrm{Dgm}(\mathbb{W})) = d_i(\mathbb{V}, \mathbb{W}).$$

A complementary example on stability properties of the bottleneck distance.

Example A.5. Let X, Y be two point clouds in $\mathbb{R}^d$. Consider the two filtrations $d_X = \mathrm{dist}_X$ and $d_Y = \mathrm{dist}_Y$. Recall (Example A.2) that f and g induce the Čech complexes of X and Y respectively. Let thus $\mathrm{Dgm}(d_X)$ and $\mathrm{Dgm}(d_Y)$ denote the *Čech diagrams* of X and Y.

Simple computations show that

$$\begin{aligned}
\|d_X - d_Y\|_\infty &= \sup_{z \in \mathbb{R}^d} \left| \min_{x \in X} \|x - z\| - \min_{y \in Y} \|y - z\| \right| \\
&= \max\{\max_{y \in Y} \min_{x \in X} \|x - y\|, \max_{x \in X} \min_{y \in Y} \|x - y\|\} \\
&=: d_H(X, Y),
\end{aligned}$$

the latter quantity being known as the *Hausdorff distance* between X and Y, a standard metric between point clouds (and more generally between compact metric spaces).

In this context, the stability theorem implies that

$$\mathbf{d}_\infty(\mathrm{Dgm}(d_X), \mathrm{Dgm}(d_Y)) \leq d_H(X, Y), \tag{A.4}$$

that is the (bottleneck) distance between the two diagrams is controlled by the (Hausdorff) distance between the point clouds (see Figure 2.2).

This example can be extended to general compact metric spaces (not only point clouds), in which case the stability is expressed in terms of the *Gromov-Hausdorff distance* [CCSG+09b].

This result has some consequences in applications. For instance, consider a (unknown) compact set $X \subset \mathbb{R}^d$, and a point cloud $P \subset X$ (for which the diagram can be computed in practice). If one can ensure that $d_H(P, X) \leq \varepsilon$ (we say that P is a ε-sample of X), then it follows that

$$\mathbf{d}_\infty(\mathrm{Dgm}(d_P), \mathrm{Dgm}(d_X)) \leq \varepsilon.$$

Under some geometric assumptions on X, not detailed here, $\mathrm{Dgm}(d_X)$ actually reflects the homology of X, and so does $\mathrm{Dgm}(d_P)$ if P is close enough to X.

Appendix B

Elements of measure theory

In the following, let Ω be a locally compact *Polish metric space* (i.e. a Polish space equipped with a distinguished Polish metric).

Definition B.1 (*Radon measures*)**.** The space $\mathcal{M}(\Omega)$ of Radon measures supported on Ω is the space of Borel measures which give finite mass to every compact set of Ω. The *vague topology* on $\mathcal{M}(\Omega)$ is the coarsest topology such that the applications $\mu \mapsto \mu(f)$ are continuous for every $f \in C_c(\Omega)$, the space of continuous functions with compact support in Ω.

Definition B.2 (*Weak topology*)**.** Denote by $\mathcal{M}_f(\Omega)$ the space of finite Borel measures on Ω. The weak topology on $\mathcal{M}_f(\Omega)$ is the coarsest topology such that the applications $\mu \mapsto \mu(f)$ are continuous for every $f \in C_b(\Omega)$, the space of continuous bounded functions in Ω.

We denote by $\xrightarrow{v}$ the vague convergence and $\xrightarrow{w}$ the weak convergence.

Definition B.3. A set $F \subset \mathcal{M}(\Omega)$ is said to be *tight* if, for every $\varepsilon > 0$, there exists a compact set K with $\mu(\Omega \backslash K) \leq \varepsilon$ for every $\mu \in F$.

The following propositions are standard results. Corresponding proofs can be found for instance in [Kal17, Section 15.7].

Proposition B.1. *A set $F \subset \mathcal{M}(\Omega)$ is relatively compact for the vague topology if and only if for every compact set K included in Ω,*

$$\sup\{\mu(K),\ \mu \in F\} < \infty.$$

Proposition B.2 (Prokhorov's theorem)**.** *A set $F \subset \mathcal{M}_f(\Omega)$ is relatively compact for the weak topology if and only if F is tight and $\sup_{\mu \in F} \mu(\Omega) < \infty$.*

Proposition B.3. *Let $\mu, \mu_1, \mu_2, \dots$ be measures in $\mathcal{M}_f(\Omega)$. Then, $\mu_n \xrightarrow{w} \mu$ if and only if $\mu_n(\Omega) \to \mu(\Omega)$ and $\mu_n \xrightarrow{v} \mu$.*

Proposition B.4 (Portmanteau theorem)**.** *Let $\mu, \mu_1, \mu_2, \dots$ be measures in $\mathcal{M}(\Omega)$. Then, $\mu_n \xrightarrow{v} \mu$ if and only if one of the following propositions holds:*

- *for all open sets $U \subset \Omega$ and all* bounded *closed sets $F \subset \Omega$,*

$$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F) \text{ and } \liminf_{n\to\infty} \mu_n(U) \geq \mu(U).$$

- *for all* bounded *Borel sets A with $\mu(\partial A) = 0$, $\lim_{n\to\infty} \mu_n(A) = \mu(A)$.*

Definition B.4. The set of *point measures* on Ω is the subset $\mathcal{D}(\Omega) \subset \mathcal{M}(\Omega)$ of Radon measures with discrete support and integer mass on each point, that is of the form

$$\sum_{x\in X} n_x \delta_x$$

where $n_x \in \mathbb{N}$ and $X \subset \Omega$ is some locally finite set.

Proposition B.5. *The set $\mathcal{D}(\Omega)$ is closed in $\mathcal{M}(\Omega)$ for the vague topology.*

Definition B.5. Let μ, ν be two non-negative Radon measures on a space Ω. We say that μ is *absolutely continuous* with respect to ν, denoted by $\mu \ll \nu$ if for all Borel $A \subset \Omega$ such that $\nu(A) = 0$, one has $\mu(A) = 0$. We say that μ and ν are *singular* (to each other), denoted by $\mu \perp \nu$, if there exists a Borel set $E \subset \Omega$ such that for all Borel A, $\mu(A\backslash E) = 0$ and $\nu(A\backslash E^c) = 0$.

Proposition B.6 (Radon-Nikodym derivative and Lebesgue decomposition theorem.)**.** *Let μ, ν be two non-negative Radon measures supported on Ω. There exists a unique pair of measures (μ_1, μ_2) such that $\mu = \mu_1 + \mu_2$, $\mu_1 \ll \nu$, and $\mu_2 \perp \nu$.*

Furthermore, there exists a measurable non-negative function denoted by $\frac{\mathrm{d}\mu_1}{\mathrm{d}\nu}$, unique ν-ae, such that $\mathrm{d}\mu_1 = \frac{\mathrm{d}\mu_1}{\mathrm{d}\nu}\mathrm{d}\nu$. This function is called the Radon-Nikodym derivative of μ with respect to ν.

Appendix C

Elements of metric geometry

Here, $\mathcal{X}$ denotes a locally compact Polish metric space, equipped with a distance d. The following definitions are standard and the interested reader can find more content in [BBB+01] and the recent book [AKP17].

Definition C.1 (Paths, lengths, geodesics)**.** A *path* between x, y in $\mathcal{X}$ is a continuous map $\gamma : [0, 1] \to \mathcal{X}$ such that $\gamma(0) = x, \gamma(1) = y$. The *length* of a path γ is

$$L(\gamma) := \inf_{0=t_1 \leq \cdots \leq t_n = 1} \sum_{i=1}^{n} d(\gamma(t_i), \gamma(t_{i+1})).$$

Let $P(x, y)$ denote the set of paths between two points x and y of $\mathcal{X}$.

Definition C.2 (Geodesics)**.** The space $\mathcal{X}$ is said to be a *length space* if

$$\inf_{\gamma \in P(x,y)} L(\gamma) = d(x, y).$$

A path realizing this infimum, should it exist, is said to be a *geodesic* between x and y.

If for any pair $x, y \in \mathcal{X}$, there exists a geodesic between x and y, $\mathcal{X}$ is said to be a *geodesic space*.

The notion of curvature in geodesic spaces is of major importance to understand the behavior of Fréchet means, among other things. For the sake of simplicity, we only expose here the notion of "non-negatively curved" Alexandrov spaces, as it covers the case of spaces of persistence diagrams. Intuitively, it means that two geodesics emanating from a same point tends

to "get closer", as it would happen on a sphere for instance. This is measured by the notion of angle, mimicking the formula one would expect in the Euclidean setting.

Definition C.3. Let p, x, y be three points in a geodesic space $\mathcal{X}$. The *angle* $\measuredangle_p(x, y)$ between x and y at p is defined as

$$\cos(\measuredangle_p(x, y)) = \frac{d(p, x)^2 + d(p, y)^2 - d(x, y)^2}{2d(p, x)d(p, y)}.$$

Definition C.4. A geodesic space $\mathcal{X}$ is said to be a *non-negatively curved Alexandrov space* if for any triplet of points (p, x, y), and for any geodesics $\gamma_x \in P(p, x)$ and $\gamma_y \in P(p, y)$, the function

$$(s, t) \mapsto \cos(\measuredangle_p(\gamma_x(s), \gamma_y(t)))$$

is increasing in both of its arguments.