Barycenters and ANOVA for Point Pattern Data

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen
Doktorgrades
“Doctor rerum naturalium”
der Georg-August-Universität Göttingen

im Promotionsprogramm
“Mathematical Sciences (SMS)”
der Georg-August University School of Science (GAUSS)

vorgelegt von
Raoul Konstantin Müller
aus Dortmund

Göttingen, 2022

Betreuungsausschuss:

Prof. Dr. Dominic Schuhmacher
Institut für Mathematische Stochastik, Universität Göttingen

Prof. Dr. Anita Schöbel
Fachbereich Mathematik, Universität Kaiserslautern, und Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM, Kaiserslautern

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik, Universität Göttingen

Mitglieder der Prüfungskommission:

Referent:
Prof. Dr. Dominic Schuhmacher
Institut für Mathematische Stochastik, Universität Göttingen

Korreferentin:
Prof. Dr. Anita Schöbel
Fachbereich Mathematik, Universität Kaiserslautern, und Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM, Kaiserslautern

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik, Universität Göttingen

Prof. Dr. Axel Munk
Institut für Mathematische Stochastik, Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik, Universität Göttingen

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik, Universität Göttingen

Tag der mündlichen Prüfung: 29. April 2022

Acknowledgements

An dieser Stelle möchte ich zuerst Dr. Dennis Michaels danken. Er hat mein Interesse an diskreter Optimierung geweckt und durch zahlreiche Gespräche, Seminare und Vorlesungen stetig mein Wissen erweitert.

Ein großes Dankeschön geht an Prof. Dr. Dominic Schuhmacher, ohne den diese Arbeit nie zustande gekommen wäre. Sein Betreuungsstil hat die Arbeit am Institut sehr angenehm gemacht. Er hat sich immer Zeit genommen, wenn es etwas zu besprechen gab, und mit ihm ist ein Arbeiten auf Augenhöhe problemlos möglich.
Natürlich auch ein großes Dankeschön an Prof. Dr. Anita Schöbel, die sich, trotz großer räumlicher Distanz und zahlreichen Verpflichtungen in Kaiserslautern, Zeit für rege Diskussionen genommen hat.

Vielen Dank an Johannes und Marcel, für die zahl- und hilfreichen Korrekturen, Anmerkungen und Verbesserungsvorschläge. Sie haben dieser Arbeit den nötigen Schliff gegeben.

Außerdem vielen Dank an das RTG2088 für die finanzielle Unterstützung und das gesamte IMS für die entspannte Arbeitsatmosphäre.

Grüße gehen raus an die kompakte Allgemeinheit. Ohne euch wäre das Studium nicht dasselbe gewesen.

Preface

Data arising in modern studies, such as analyses of material composition or distribution of species in a certain habitat, is often represented as spatial point patterns. Frequently, in real life applications the observed patterns have different numbers of points (cardinalities) and, moreover, often exhibit a complex interaction structure such as inhibition or clustering of points.
To evaluate the (dis)similarity of two objects, a natural approach is to measure the distance between these objects. However, the choice of an appropriate distance measure is, then again, not always obvious. Any reasonable metric for point pattern data should likewise account for features such as the point positions and cardinality of points. Being able to calculate the distance between point patterns allows for more in-depth studies of point pattern data.

In research areas like operations research (OR) and statistics, typical representatives of a dataset are of interest. The calculation of such a central object is the topic of research since the 17th century. This central object satisfies prespecified criteria of *physical closeness* to the data. The problem of finding one point in the plane, that minimizes the sum of its Euclidean distance to three given points dates back at least to Pierre de Fermat in 1640. Since then, this problem has been extended to more general spaces, distances and objects, where minimizers are often called *Fréchet mean* or *barycenter*.

Data obtained in real life applications is usually confounded by random noise. To determine if two or more groups of observations are significantly different or if differences can be ascribed to randomness, *Analysis of Variance* (ANOVA) offers a broad variety of statistical tests. These tests require the concept of a mean, or, for point pattern data, a barycenter.

In this thesis a new metric for point patterns is developed. A heuristic algorithm is presented, that computes barycenters with respect to the new metric. The barycenters are then used in ANOVA procedures, to assess the differences between groups of point patterns. Additionally, a new statistical test is introduced, that depends solely on the pairwise distances between the observations. The last topic in this thesis is the barycenter computation of a set of points in a Euclidean space. This problem arises as a subproblem of the barycenter calculation of point patterns and is moreover a generalization of the problem introduced by Pierre de Fermat.

This dissertation is based on three publications, which can be found in the addenda. Although the points of a point pattern usually lie in a two- or three-dimensional Euclidean space, the space of point patterns is not Euclidean. Chapter 1 provides an overview of the challenges that come along with the non-Euclidean structure of point pattern data and introduces all the specific problems in more detail. The main results concerning these problems are then stated in Chapters 2-4.

Chapter 2 covers Müller et al. (2020), giving an overview of the work related to the developed point pattern metric and the barycenter algorithm. It also provides a comparison to existing algorithms and methods. In Section 2.1, a comprehensive discussion is presented about the existing literature on related metrics for point pattern data and algorithms for barycenter computations. The main results of this paper are discussed in more detail in Section 2.2.

Chapter 3 deals with the subjects of Müller et al. (2022b). It presents an overview on ANOVA procedures that can be used for point pattern data. These procedures consist of the (classical) ANOVA F-test, for detecting differences in group means, and Levene's test, for detecting differences in group variances. A new Levene's test that solely uses the pairwise distances of the data is introduced. Further reading regarding the topic of ANOVA on non-Euclidean spaces is presented in Section 3.1. The main results of this paper are summarized in Section 3.2.

Chapter 4 is about Müller et al. (2022a), in which the barycenter of a set of points is investigated. In this setting the metric is *cut off* a fixed value. Related literature on the latest research is presented in Section 4.1 and the main results of the paper are listed in Section 4.2.

Finally, Chapter 5 discusses results and future research and gives an outlook for further perspectives.

The author's contributions to the three articles are stated in Sections 2.3, 3.3 and 4.3, respectively.

Contents

List of Symbols

$\mathbb{Z}$	the set of integers
$\mathbb{Z}_{\geq 0}$	the set of non-negative integers
$\mathbb{N}$	the set of positive integers
$\mathbb{R}$	the set of real numbers
$(\mathcal{X}, d)$	a metric space
$\aleph$	an additional element $\notin \mathcal{X}$ that serves as a dummy point
ξ, η, ζ	point patterns
$\lvert\xi\rvert$	the cardinality of the point pattern ξ
$\mathfrak{N}$	the set of all point patterns
$\mathfrak{N}_{fin}$	the set of point patterns with finite cardinality
$[n]$	$\{1, 2, \ldots, n\}$, for $n \in \mathbb{Z}_{\geq 0}$
S_n	the set of all permutations on $[n]$, $n \in \mathbb{N}$
$\mathcal{A}$	a finite subset of $\mathbb{R}^k$
$\mathrm{conv}(\mathcal{A})$	$\{\sum_{i=1}^{n} \lambda_i a_i \,\vert\, a_i \in \mathcal{A}, n \in \mathbb{N}, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1\}$, the convex hull of $\mathcal{A}$

CHAPTER 1

Introduction

In many current research topics the data is naturally given by point patterns. Examples, where point patterns turn up naturally, are geo-referrenced data in the plane, in $\mathbb{R}^3$ or on graphs, e.g. street networks, but also data on manifolds or function spaces are possible. Some concrete examples are locations of accidents in a single day, burglaries in a city in one week, locations of trees in a forest or, more technological, locations of air bubbles in a mineral flotation experiment or even cells in biomedical images. For an overview of possible applications see for example Diggle (2013), Baddeley et al. (2015), Błaszczyszyn et al. (2018).

For all settings presented in this work, a point pattern is a subset of a metric space $(\mathcal{X}, d)$. Denote with $\mathfrak{N}$ the set of all point patterns on the metric space $(\mathcal{X}, d)$ and use $|\xi|$ to denote the total number of points in the pattern $\xi \in \mathfrak{N}$. The set of finite point patterns is defined as $\mathfrak{N}_{fin} = \{\xi \in \mathfrak{N} \mid |\xi| < \infty\}$. For $n \in \mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$ write $[n] = \{1, 2, \ldots, n\}$ (including $[0] = \emptyset$), and denote by $\mathfrak{N}_n$ the set of point patterns with exactly n points. From now on all point patterns are tacitly assumed to be finite. In this thesis finite point patterns are mostly written as $\xi = \sum_{i=1}^{n} \delta_{x_i}$, where δ_x denotes the Dirac measure with unit mass at, not necessarily distinct, points $x \in \mathcal{X}$.

This introduction is divided into four parts. First a new metric is introduced in Section 1.1, that measures distances between point patterns. The metric generalizes previous point pattern metrics.

Due to replications of experiments or due to a time series, the considered data often consists of multiple observations. For many applications a *typical representative* of these observations is of interest. On general metric spaces the *Fréchet mean*, a generalization of the (arithmetic) mean, can be seen as a typical representative, as it is the object that minimizes its sum of distances to the $q \in [1, \infty)$ to the data.

Section 1.2 focuses on the computation of Fréchet means of point patterns.

Section 1.3 gives an overview of some statistical methods that are used to distinguish groups of observations. In addition, similar methods for point pattern data are derived.

Finally, Section 1.4 takes a closer look at different problems of finding the Fréchet mean of data in a Euclidean space:

- the classical *Weber* problem of finding a point that minimizes its distance to n given points,
- the Weber problem with a distance that is cut off at a fixed value,
- the Weber problem with the cut off distance and the option to choose an *empty* solution at fixed costs per point.

More details on the subjects of Sections 1.1 and 1.2 can be found in Müller et al. (2020). Section 1.3 and Section 1.4 cover the topics of Müller et al. (2022b) and Müller et al. (2022a), respectively.

1.1 The (R)TT-metric

In this section different methods of measuring the distance between two point patterns are discussed. Two metrics are commonly used: the *spike time metric*, introduced by Victor and Purpura (1997) and Diez et al. (2012), and the *optimal subpattern assignment (OSPA) metric*, as defined by Schuhmacher and Xia (2008) and Schuhmacher et al. (2008). These metrics can be generalized by a new metric that is introduced in this section.

Definition. *Let $C > 0$ and $p \geq 1$ be two parameters, referred to as* penalty *and* order*, respectively.*

(a) For $\xi = \sum_{i=1}^{m} \delta_{x_i}, \eta = \sum_{j=1}^{n} \delta_{y_j} \in \mathfrak{N}_{fin}$ define the transport-transform (TT) metric *by*

$$\begin{aligned}\tau(\xi, \eta) &= \tau_{C,p}(\xi, \eta) \\ &= \left(\min\left((m + n - 2l)C^p + \sum_{r=1}^{l} d(x_{i_r}, y_{j_r})^p\right)\right)^{1/p},\end{aligned} \tag{1.1}$$

where the minimum is taken over equal numbers of pairwise different indices $i_1, \ldots, i_l$ *in* $[m]$ *and* $j_1, \ldots, j_l$ *in* $[n]$*, i.e. over the set*

$$\begin{aligned} S(m,n) = \big\{ (i_1, \ldots, i_l; j_1, \ldots, j_l)\, ; \\ & l \in \{0, 1, \ldots, \min\{m,n\}\}, \\ & i_1, \ldots, i_l \in [m] \textit{ pairwise different}, \\ & j_1, \ldots, j_l \in [n] \textit{ pairwise different} \big\}. \end{aligned}$$

(b) For $\xi, \eta \in \mathfrak{N}_{fin}$ *define the* relative transport-transform (RTT) metric *by*

$$\bar{\tau}(\xi,\eta) = \bar{\tau}_{C,p}(\xi,\eta) = \frac{1}{\max\{|\xi|,|\eta|\}^{1/p}}\, \tau_{C,p}(\xi,\eta). \tag{1.2}$$

In general the cardinalities of the patterns are different. The main idea of these metrics is to transform one point pattern into the other one, by *moving* a subset of l points of ξ to a subset of l points of η with costs given by the underlying metric d. This corresponds to the sum over the $d(x_{i_r}, y_{j_r})^p$ in the objective (1.1). For any $r \in \{1, \ldots, l\}$ the points x_{i_r} and y_{j_r} are called *matched*, the remaining $m + n - 2l$ points are called *unmatched.*
One single point of ξ cannot be matched with more than one point in η and vice versa. That means that for $m \neq n$ there are at least $|m - n|$ unmatched points in the patterns. These points are penalized each with the value C^p in the objective.

Two unmatched points cost $2C^p$, hence, to comply with the triangle inequality, no points at any distance greater than $2C^p$ are matched.
Minimizing over the set $S(m,n)$ is difficult in general. Since l is unknown, the size of $S(m,n)$ grows exponential in $\min\{m,n\}$. Enumerating all possibilities is time consuming. With a slight modification of the metric space $\mathcal{X}$ and the metric d, the computation of the (R)TT-metric can be simplified. Extend the metric space $\mathcal{X}$ by an element $\aleph \notin \mathcal{X}$ and define the new metric d' by

$$d'(x,y) = \begin{cases} \min\{d(x,y), 2^{1/p}C\} & \text{if } x, y \in \mathcal{X}; \\ C & \text{if } \aleph \in \{x,y\},\ x \neq y; \\ 0 & \text{if } x = y = \aleph. \end{cases} \tag{1.3}$$

The element $\aleph$ is a new location in the metric space. Matching a point $x \in \mathcal{X}$ with $\aleph$ is always done at cost C. Let $\mathcal{X}' = \mathcal{X} \cup \{\aleph\}$, then $(\mathcal{X}', d')$ is again a metric space, cf. Müller et al. (2020).With this metric space the computations of τ and $\bar{\tau}$ become simple optimal matching problems. Denote by S_n the set of all permutations on $[n]$.

Theorem 1. *Let $\xi = \sum_{i=1}^{m} \delta_{x_i}, \eta = \sum_{j=1}^{n} \delta_{y_j} \in \mathfrak{N}_{fin}$, and suppose that $m \leq n$ (otherwise swap ξ and η). Set $x_i = \aleph$ for $m+1 \leq i \leq n$ and $\xi = \sum_{i=1}^{n} \delta_{x_i}$. Then,*

$$\tau(\xi, \eta) = \left(\min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \right)^{1/p} \quad \textit{and} \quad \bar{\tau}(\xi, \eta) = \left(\frac{1}{n} \min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \right)^{1/p}.$$

The smaller pattern is extended with points in $\aleph$ until it has the same cardinality as the larger pattern. Theorem 1 states that the calculation of the metrics τ and $\bar{\tau}$ can then be done by optimizing over all possible matchings π. This has a worst-time complexity of $\mathcal{O}(n^3)$ using the *Hungarian Algorithm*, see Kuhn (1955). In practice the *Auction Algorithm* described in Bertsekas (1988) performs better in the average case, although its worst-time complexity of $\mathcal{O}(n^3 \log(n))$ is slightly larger than that of the Hungarian Algorithm.

Both the OSPA- and the spike time metric can be computed through τ or $\bar{\tau}$. For the OSPA-metric the proposed algorithm has the same worst-time performance of $\mathcal{O}(n^3)$ while the *Incremental Matching Algorithm* that is proposed for the spike time metric in Diez et al. (2012) has a worst-time complexity of $\mathcal{O}(n^6)$. In this case the formulation of Theorem 1 allows for a huge improvement.

The *Wasserstein* distance is a metric on the space of probability measures. Mass is transported from one measure to another measure at costs modelled by some cost function c. Chizat et al. (2018) introduce the *unbalanced optimal transport* for probability measures. It is similar to the Wasserstein distance, but not all mass has to be transported. Any mass that is not transported is penalized.

Point patterns can be interpreted as discrete probability measures. Let two point patterns $\xi, |\xi| = m$ and $\eta, |\eta| = n$, $m \leq n$, be given. In the definition of the point patterns every point is assigned the same mass *one*. A probability measure, however, has total measure 1. To comply with the uniformity of the weights of the points and also the property of probability measures, that their total measure equals 1, set the weights of the points in both patterns to $1/n$.

For $n = m$ both probability measures associated with the point patterns have mass 1.

If $m < n$ then the total measure of the pattern ξ is less than 1. In terms of probability measures ξ is *incomplete*. In both cases the (R)TT-metric can be interpreted as unbalanced optimal transport metric.

1.2 Barycenters of point patterns

For a given set of data point patterns a typical representative of these point patterns is often of interest. For example the ANOVA procedures require some central object, cf. Section 1.3, but also in the location theory in operations research this is an often studied problem.

1.2.1 Introduction to Fréchet means

On general metric spaces Fréchet (1948) introduced a generalization of a *centroid* or *center of mass*, which was later named *Fréchet mean.* Let $(\mathfrak{N}, \tau)$ be the metric space of point patterns and let $\lambda_1, \ldots, \lambda_k > 0$ be weights with $\sum_{i=1}^{k} \lambda_i = 1$. The *(weighted) Fréchet-q-mean*, or *barycenter*, of order $q > 1$ for data $\xi_1, \ldots, \xi_k \in \mathfrak{N}$ is any point pattern $\zeta \in \mathfrak{N}$ that minimizes the Fréchet functional

$$\mathcal{F}^q(\zeta) = \sum_{i=1}^{k} \lambda_i \tau(\zeta, \xi_i)^q. \tag{1.4}$$

If no weights are specified, assume that $\lambda_i = 1/k$ for $1 \leq i \leq k$, leading to an *unweighted* barycenter.[1]

For $q = 2$ write Fréchet mean for short. For $q = 1$ a barycenter is also known as Fréchet median.

Note that for data in $\mathbb{R}^m$, for $q = 2$ and the Euclidean distance the arithmetic mean of the data is a minimizer of $\mathcal{F}$. Moreover, assuming $q = 1$ and the taxicab metric, the median of the data is a minimizer of $\mathcal{F}$. For $m = 1$, i.e. the metric space is the real line, also the geometric mean and harmonic mean are optimal solutions of the Fréchet functional, depending on the metric d and the power q.

Calculating the minimizer of the Fréchet functional of Euclidean data is a well studied problem. Section 1.4 gives a brief overview of related problems and Chapter 4 presents the results for a metric that is cut off.

Borgwardt and Patterson (2021) have shown that the computation of a sparse barycenter for three point patterns with equal cardinality, in $\mathbb{R}^2$, is $\mathcal{NP}$-hard for the Wasserstein-2-distance. The barycenter calculation in the (R)TT-metric is closely related to the one covered by Borgwardt and Patterson (2021). It is therefore to assume that in general the computation of a barycenter for point patterns with respect to the (R)TT-metric is practically infeasible.

[1] Here the Fréchet functional is defined on the space of point patterns $(\mathfrak{N}, \tau)$. For the definition of the Fréchet functional on a general metric space $(\mathcal{X}, d)$ replace the metric τ with d in (1.4).

To still be able to work with a central object of a set of given point patterns, Müller et al. (2020) developed a heuristic algorithm that is similar to the k-means algorithm of Lloyd (1982). The algorithm starts with a random point pattern, the *pseudo-barycenter*. Then the distance between the pseudo-barycenter and the given point patterns is computed. The calculation of the TT-metric defines clusters of points: one point of the pseudo-barycenter is matched with exactly one point (including $\aleph$) of every point pattern in the distance calculation. The next step is to compute the barycenter of every cluster. These barycenter points of the clusters then become the new pseudo-barycenter. Both steps, distance calculation and computation of the barycenters of the clusters, are iterated alternatingly until convergence. This process is illustrated for four point patterns in Figure 1.1.

In the algorithm of Müller et al. (2020) the computation of the barycenters of the clusters is done by a heuristic. The optimal solution to this problem is subject of Müller et al. (2022a) and a summary of the main results can be found in Chapter 4. Since the number of points in the barycenter is part of the optimization process, the algorithm also has routines for increasing and decreasing the number of points in the pseudo-barycenter.

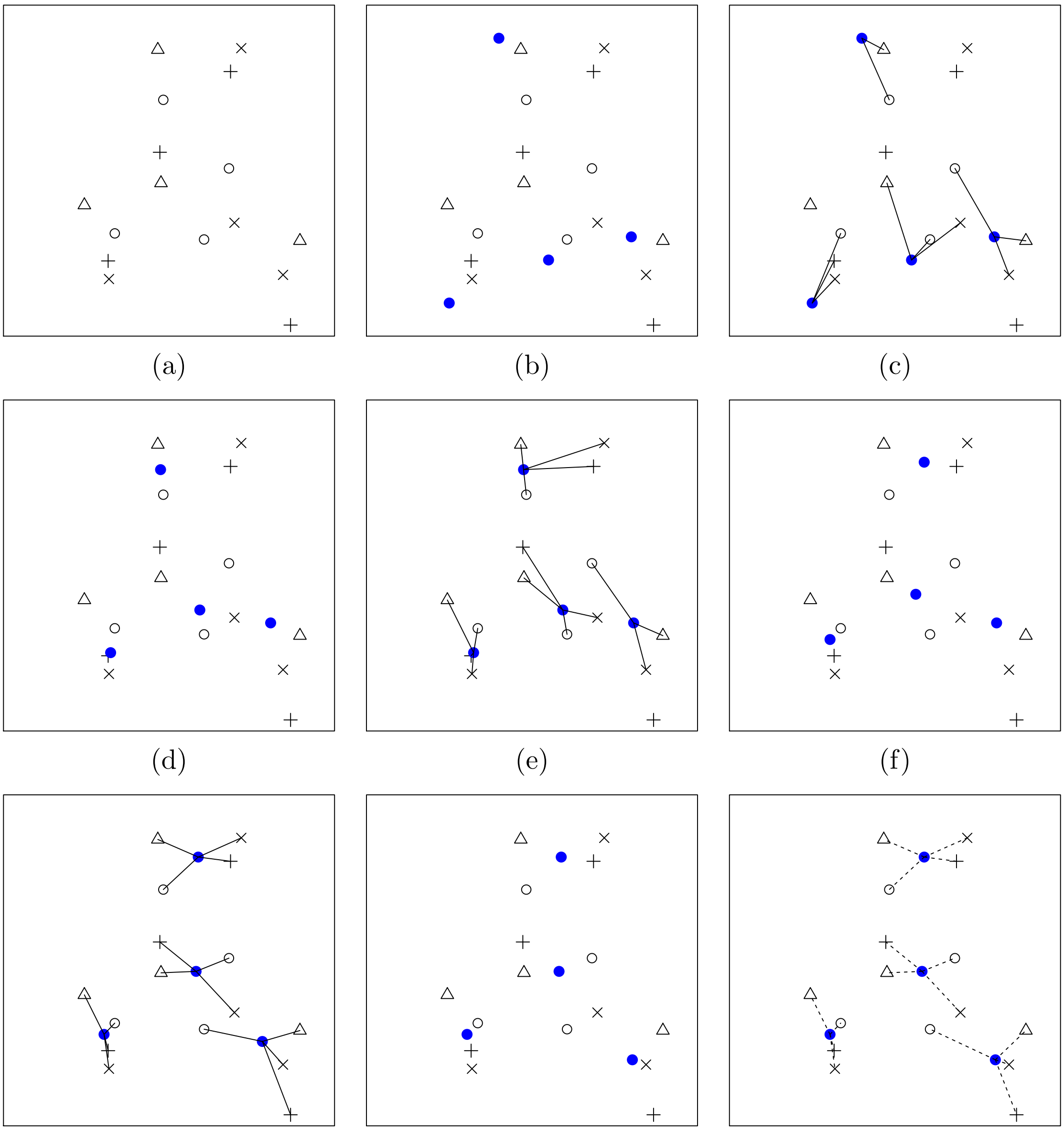
For more details on the different routines and the algorithm itself see Müller et al. (2020), Section 4. Moreover, in Section 4.2 an improvement is provided that reduces the runtime of the algorithm by a factor of 2 for large instances.

1.2.2 Comparison and simulation study

For Wasserstein-2 barycenters and finite support probability measures, similar algorithms had been introduced by Cuturi and Doucet (2014), del Barrio et al. (2019) and Borgwardt (2020). In Müller et al. (2020), Section 5, their algorithm is compared to the algorithm of Cuturi and Doucet (2014).

For $p = q = 1$ the barycenter problem for point patterns has been studied by Schoenberg and Tranbarger (2008), Diez et al. (2012) and Mateu et al. (2015) who used the spike time metric. Due to the large computational cost of this metric they, however, were only able to cope with rather small data sets.

In Figure 1.2 are three scenarios of three data point patterns each (in black), indicated by the different symbols (triangle, cross, upside down triangle). The computed pseudo-barycenter is presented in blue. It is clearly visible that the pseudo-barycenter well preserves the structure of the data, which is often desired from an application point of view. This is best visible in the second and third image, where the data patterns consist of points arranged in circular shapes. The pseudo-barycenter does not only represent the data by lying *between* the data points but



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 1.1: The steps in the evolution of the pseudo-barycenter, as calculated by the algorithm of Müller et al. (2020) for four point patterns inside the unit square. Each point pattern is represented by one black symbol (panel (a)). The pseudo-barycenter is pictured in blue.

Each step of the algorithm is pictured in one image. Starting with a random pseudo-barycenter (panel (b)), the first matching between the points of the patterns and the pseudo-barycenter is calculated through the TT-distance. This is depicted in panel (c). Points with a Euclidean distance of at most 0.25 to the pseudo-barycenter are connected to a matched blue point with a black line (panels c, e, g). For the clusters of points that are connected through a blue point is then a new centerpoint calculated (panels d, f, h). The two steps, calculating the matching through the TT-distance and new centerpoints, are repeated until the final barycenter is calculated. The calculated matching in the last panel is the same as the previous matching (in (g)). The objective function value does not decrease and the algorithm stops.

To save space, two steps between (f) and (g) are not depicted.

also the circular shape of the data is visible in the pseudo-barycenter point pattern.

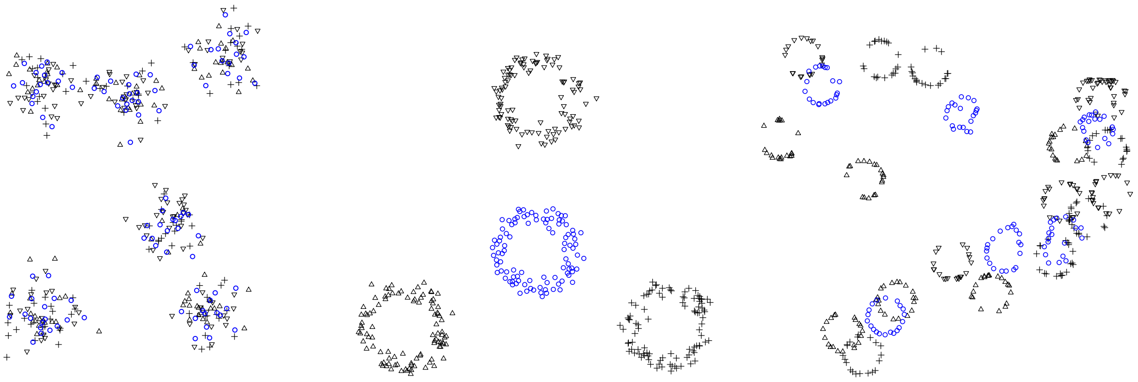


Figure 1.2: An example of barycenters computed by our algorithm for three different data sets. In each panel there are three data point patterns indicated by the three different symbols (black). The resulting (pseudo-)barycenter pattern with respect to squared Euclidean distance is given by the blue circles ($p = q = 2$).

In Section 5 of Müller et al. (2020), a large study on the robustness of the algorithm is presented, as well as a runtime analysis. The study is divided into different scenarios for the distribution of the points in the data point patterns. The scenarios are further divided by the number of points per pattern, as well as the number of patterns per group, for which a pseudo-barycenter is computed. The algorithm is started multiple times on the same data with different starting patterns and the resulting costs are compared. For smaller instances, 20 point patterns with an expected number of 20 points each, the computed costs deviated roughly 5% from the smallest computed value. For larger instances, 100 patterns with an expected number of 100 points each, the computed costs deviated roughly 2.5% from the smallest computed value. This low variance in the computed costs indicates the robustness of the algorithm. The computed pseudo-barycenters are not only reliable, but they visually represent the data well and their computation is moreover fast. For example, a pseudo-barycenter of 100 patterns with around 100 points each can be calculated in about 2 to 3 seconds.

1.3 ANOVA

A popular application of barycenters finds use in the ANOVA. The theory of classical ANOVA dates back to Fisher (1925), but variants to generalized settings are still part of research in recent times, cf. Ramón et al. (2016), Anderson (2017), Hamidi et al. (2019). ANOVA offers a number of different methods for detecting differences between groups of data. This section recaps two very common methods, one for detecting differences in the group means and one for detecting differences in the group variances.

Let $x_{ij} \in \mathbb{R}$ be observations, where $1 \leq i \leq k$ denotes the k different groups and $1 \leq j \leq n_i$ denotes the n_i observations of group i. The observations stem from k independent distributions $P_1, \ldots, P_k$. Let $n = \sum_{i=1}^{k} n_i$ be the total number of observations, $\bar{x}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$ the *group mean* of group i and let $\bar{x}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}$ be the *overall mean*. Define

$$\begin{aligned} \mathrm{TSS} &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{\cdot\cdot})^2 && \textit{(total sum of squares)} \\ \mathrm{MSS} &= \sum_{i=1}^{k} n_i (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 && \textit{(model sum of squares)} \\ \mathrm{RSS} &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot})^2 && \textit{(residual sum of squares).} \end{aligned}$$

These sums are proportional to variance estimates of the complete data or groups of data. For example $\frac{1}{n-1}\mathrm{TSS}$ is a variance estimator of the whole dataset. The basic idea of ANOVA is to compare these variance estimates, e.g. in the F-test (1.5) the MSS is compared to the RSS. Suppose the group means $\bar{x}_{i\cdot}$ are equal, then the MSS is 0. If the difference between the groups means is large, then MSS is large. The F-statistic

$$F = \frac{n-k}{k-1} \frac{\mathrm{MSS}}{\mathrm{RSS}} = \frac{n-k}{k-1} \frac{\mathrm{TSS} - \mathrm{RSS}}{\mathrm{RSS}} \tag{1.5}$$

compares the model sum of squares to the residual sum of squares. If all distributions $P_1 = \ldots = P_k$ are Gaussian then the statistic (1.5) is F-distributed with $k-1$ and $n-k$ degrees of freedom. This fact is used to assess the equality of the group means. Given Gaussian data with equal variances, the value of the computed F-statistic can be compared to the $(1-\alpha) \cdot 100\%$-quantile of the F distribution for some $\alpha \in (0,1)$. If the value of the F-statistic is larger than the quantile, then the hypothesis that all the distributions have equal means can be rejected at the significance level α.

The basic theory related to statistical testing is not subject of this thesis. For an introduction on statistical testing, including statistical tests, hypotheses and significance levels, cf. Lehmann et al. (2005) or Rice (2006).

To detect differences in group variances a common method is *Levene's test.* Additional to the definitions above define $z_{ij} = |x_{ij} - \bar{x}_{i\cdot}|$, $\bar{z}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} z_{ij}$ and $\bar{z}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} z_{ij}$. Levene's test statistic is then equal to the usual F-statistic if the observations x_{ij} are replaced by their absolute deviations from their respective group mean:

$$\widetilde{F} = \frac{n-k}{k-1} \cdot \frac{\sum_{i=1}^{k} n_i (\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot})^2}. \tag{1.6}$$

Both statistics (1.5) and (1.6) are originally introduced for Euclidean data. In general, point pattern data is neither Euclidean nor Gaussian, even if the points of the patterns lie in a Euclidean space. It is still possible to transfer the ideas of ANOVA, i.e. to decide whether differences between the observations are random or significantly dfferent, to the space of point patterns. Dubey and Müller (2019) present a statistic for data in any metric space and combine the methods of the classical ANOVA F-statistic and Levene's test. However, their statistic is based on distances between observations and Fréchet means, whose computation for point patterns is in general computationally infeasible. Even with the fast heuristic presented in Müller et al. (2020), the calculation of the statistic is slow compared to statistics for Euclidean data.

To remedy this issue a transformation to the data is applied, to map the data into a Euclidean space, without losing information that is necessary for ANOVA. To this end Anderson (2001) proposed a statistic that is similar to the F-statistic, but without the need of any central object. For real valued data $y_1, \dots, y_m \in \mathbb{R}$ the following equality always holds:

$$\sum_{j=1}^{m} (y_j - \bar{y})^2 = \frac{1}{2m} \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} (y_{j_1} - y_{j_2})^2 = \frac{1}{m} \sum_{j_1=1}^{m-1} \sum_{j_2=j_1+1}^{m} (y_{j_1} - y_{j_2})^2.$$

To transfer this identity to a general metric space, the squared differences between the observations and their respective group mean are substituted with squared distances between the observations. Thus, the sums of squares then become

$$\begin{aligned}
\mathrm{TSS} &= \frac{1}{n} \left(\sum_{i_1=1}^{k-1} \sum_{i_2=i_1+1}^{k} \sum_{j_1=1}^{n_{i_1}} \sum_{j_2=1}^{n_{i_2}} d^2(x_{i_1 j_1}, x_{i_2 j_2}) + \sum_{i=1}^{k} \sum_{j_1=1}^{n_i-1} \sum_{j_2=j_1+1}^{n_i} d^2(x_{ij_1}, x_{ij_2}) \right), \\
\mathrm{RSS} &= \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j_1=1}^{n_i-1} \sum_{j_2=j_1+1}^{n_i} d^2(x_{ij_1}, x_{ij_2}), \\
\mathrm{MSS} &= \mathrm{TSS} - \mathrm{RSS}.
\end{aligned}$$

In the $n \times n$-matrix of squared distances between all observations, TSS is the sum of the upper triangle divided by n. With these sums of squares the Anderson F-

statistic is

$$F_{\mathrm{A}} = \frac{n-k}{k-1}\frac{\mathrm{MSS}}{\mathrm{RSS}}.$$

Statistic F_{A}, like the original F-statistic, can be used to detect differences in the means of the distributions $P_1, \ldots, P_k$. To detect differences in the variances of the distributions on arbitrary metric spaces Müller et al. (2022b) developed a statistic similar to Levene's test, that relies only on the pairwise distances between observations. The construction of the test is based on ideas similar to the ones for the statistic F_A:

Suppose in group i are n_i observations. Then there are $N_i = \binom{n_i}{2}$ pairwise distances $d_{i,\{j_1,j_2\}} = \frac{1}{2}d(x_{i,j_1}, x_{i,j_2})$ between the observations of this group. The distances are multiplied by the factor $\frac{1}{2}$ because the sum of the distances between the observations is roughly twice the sum of distances between the observations and their barycenter. The distances are rescaled for convenience as will be visible further below. Since the factor appears in the numerator as well as in the denominator and therefore cancels out, it does not change properties of the statistic. Enumerate these distances $d_{i,\{j_1,j_2\}}$ as d_{ij} with $1 \leq j \leq N_i$. Analogously to the previous definitions define

$$\bar{d}_{i\cdot} = \frac{1}{N_i}\sum_{j=1}^{N_i} d_{ij} \quad \text{and} \quad \bar{d}_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} d_{ij},$$

the groupmeans and the overall mean of the pairwise distances, respectively. Like the z_{ij} in statistic (1.6) these distances now become the data in the k groups. The total number of observations, i.e. these pairwise distances, is therefore set to $N = \sum_{i=1}^{k} N_i$. The new Levene's statistic, based on statistic (1.6) and the ideas of Anderson (2001), is

$$L_M = \frac{N-k}{k-1} \cdot \frac{\frac{1}{n}\sum_{i_1=1}^{k-1}\sum_{i_2=i_1+1}^{k} n_{i_1}n_{i_2}(\bar{d}_{i_1\cdot} - \bar{d}_{i_2\cdot})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(d_{ij} - \bar{d}_{i\cdot})^2}. \tag{1.7}$$

Assuming $n_1 = \ldots = n_k$ statistic L_M simplifies to

$$L_M = \frac{N-k}{k-1} \cdot \frac{\sum_{i=1}^{k} n_i(\bar{d}_{i\cdot} - \bar{d}_{\cdot\cdot})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(d_{ij} - \bar{d}_{i\cdot})^2}. \tag{1.8}$$

These statistics can be calculated quickly, which is an advantage especially in permutation tests, where statistics are calculated a 1000 times, or even more often. In simulation the statistics perform well for point pattern data, cf. Müller et al. (2022b), Section 4, for more details. In Section 4.2 it is proven that the L_M-statistic is asymptotically χ^2 distributed up to a constant factor. A comparison of the new

L_M-statistic to previous methods can be found in Müller et al. (2022b), Section 6. Additionally, the main results from Müller et al. (2022b) regarding the L_M-statistics and further information on the different ANOVA statistics can be found in Chapter 3.

1.4 Barycenter of a cluster of points

Müller et al. (2020) present a heuristic algorithm for the computation of barycenters of point patterns. The algorithm alternates between two problems: First the optimal matching problems of the pseudo-barycenter to the data point patterns and second the calculation of a barycenter point for all the clusters that are defined through the matching (cf. Section 1.2). In the implementation of the algorithm, which can be found in the publicly available R package `ttbary`, Müller and Schuhmacher (2021), the barycenter of a cluster of points is computed with a computationally attractive heuristic. The exact solution to this problem and some related problems are the topic of Müller et al. (2022a). The main results are summarized in Chapter 4.

Apart from the barycenter problems from Sections 1.1-1.3, there is a variety of variations which are considered henceforth. To this end let the data be given by a finite set $\mathcal{A} \subseteq \mathbb{R}^k$, $|\mathcal{A}| = n$. Let d be a metric and $q \geq 1$. Denote by $\operatorname{diam}(\mathcal{A}) = \max_{a_1, a_2 \in \mathcal{A}} d(a_1, a_2)$ the *diameter* of $\mathcal{A}$. The *barycenter problem* is defined as

$$\mathcal{Z}^* = \min_{x \in \mathbb{R}^k} f(x, \mathcal{A}) = \min_{x \in \mathbb{R}^k} \sum_{a \in \mathcal{A}} d^q(x, a). \tag{Bar($\mathcal{A}$)}$$

and denote the set of its optimal solutions with $\mathcal{X}^*$. This is again the minimization of a Fréchet functional, cf. (1.4), over data in a Euclidean space. The barycenter problem (Bar($\mathcal{A}$)) is also known as *Weber* problem, *one median* problem or *minisum* problem. It dates back at least to Pierre de Fermat who formulated the problem of finding a point that minimizes its total distance to the three vertices of a triangle before 1640.

In this section three related problems are introduced:

1. the classical barycenter problem (Bar($\mathcal{A}$)),

2. the barycenter problem with a metric that is 'cut off',

3. a barycenter problem that allows the point $\aleph$ as a solution.

The relation between the solutions of these problems is discussed in Müller et al. (2022a). A summary of the main results can be found in Chapter 4.

Given a *cutoff value* $C > 0$, define the cutoff-distance

$$d_C^q(x, y) = \min\{d^q(x, y), C\}, \ x, y \in \mathbb{R}^k, \tag{1.9}$$

i.e., the distance is capped at the maximum value C. On the space $\mathbb{R}^k \cup \{\aleph\}$ define additionally

$$d_{C,\alpha}^q(\aleph, y) = \begin{cases} \alpha \cdot C & \text{if } y \neq \aleph \\ 0 & \text{if } y = \aleph \end{cases} \tag{1.10}$$

for a given $\alpha > 0$ and leave $d_{C,\alpha}^q(x, y) = d_C^q(x, y)$ for all $x, y \neq \aleph$. The idea for (1.9) and (1.10), and hence the for the problems 2 and 3, comes from (1.3).
Note that if d is a metric, then d_C is also a metric. And if $(\mathbb{R}^k, d_C)$ is a metric space, then $(\mathbb{R}^k \cup \{\aleph\}, d_{C,\alpha})$ is a metric space if $\alpha \geq \frac{1}{2}$, see Lemma 1 and Lemma 2 in Müller et al. (2022a).

The corresponding so called location problems are given as

- the *barycenter problem with cutoff*

  $$\mathcal{Z}_C^* = \min_{x \in \mathbb{R}^k} f_C(x, \mathcal{A}) = \min_{x \in \mathbb{R}^k} \sum_{a \in \mathcal{A}} \min\{d^q(x, a), C\}, \tag{$\mathrm{Bar}_C(\mathcal{A})$}$$

  with its set $\mathcal{X}_C^*$ of optimal solutions.

- the *barycenter problem with cutoff and* $\aleph$

  $$\mathcal{Z}_{C,\alpha}^* = \min_{x \in \mathbb{R}^k \cup \{\aleph\}} f_{C,\alpha}(x, \mathcal{A}) = \min_{x \in \mathbb{R}^k} \sum_{a \in \mathcal{A}} d_{C,\alpha}^q(x, a), \tag{$\mathrm{Bar}_{C,\alpha}(\mathcal{A})$}$$

  with its set $\mathcal{X}_{C,\alpha}^*$ of optimal solutions.

Problem ($\mathrm{Bar}_C(\mathcal{A})$) has been studied by Drezner et al. (1991) in a slightly different setting, where each point $a \in \mathcal{A}$ has its own cutoff C_a, however only the ℓ_1- and ℓ_2-norms are considered. They also present an algorithm for their problem, that computes the optimal solution in polynomial time.

($\mathrm{Bar}_C(\mathcal{A})$) considers only one cutoff C, but allows general metrics d and powers q. The main results regarding the connections between the three barycenter problems are presented in Chapter 4, see also Müller et al. (2022a) for more details. Additionally, improved versions of the algorithm of Drezner et al. (1991) are presented, which are fit to the problems ($\mathrm{Bar}_C(\mathcal{A})$) and ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$).

CHAPTER 2

Metrics and barycenters for point pattern data

This chapter evaluates Müller et al. (2020). It offers a literature review on recent metrics for point pattern data and on barycenter algorithms in Section 2.1 and a discussion of the main results of Müller et al. (2020) in Section 2.2. In Section 2.3 my own contributions to this article are pointed out.
The impact of the article on possible future research is presented in an outlook in Chapter 5.

2.1 Literature overview

This section provides a brief overview of the literature concerning metrics for probability measures, especially for discrete measures, and current methods for computing barycenters regarding these metrics. Müller et al. (2020) focuses mainly on the development of a new metric for point patterns, which are related to discrete probability measures, and a method for fast computations of pseudo-barycenters for point patterns.

The most common metric for probability measures is the *Wasserstein* distance. A point pattern can be interpreted as a discrete probability measure: assign positive mass to every point of the pattern, such that all the mass sums to one, cf. Section 1.1. The barycenter problem for (discrete) probability measures with respect to the Wasserstein distance is a well studied subject.
Borgwardt and Patterson (2021) proved that the computation of a sparse barycenter for three point patterns with respect to the Wasserstein distance is $\mathcal{NP}$-hard, even in two dimensions and when all the patterns have the same cardinality n and every point has weight $1/n$. It is hence safe to assume, that the calculation of

a barycenter with respect to the TT-distance is also $\mathcal{NP}$-hard, although this has yet to be proven. Still, there is a lot of on-going research about the computation of a (heuristical) Wasserstein barycenter, the most common approach being through linear programming. Examples include approaches by Cuturi and Doucet (2014), del Barrio et al. (2019), Borgwardt and Patterson (2020) and Borgwardt (2020).
A different approach can be found in Heinemann et al. (2021) where the barycenter is not computed for the whole data, but rather for random subsets of the data. If these subsets are small compared to the size of the original data, this method yields tremendous improvements in terms of runtime, while the solution is close to the population barycenter in expectation.
Altschuler and Boix-Adserà (2021) present a polynomial-time algorithm that approximates the barycenter up to a prespecified accuracy. In certain settings the algorithm computes an exact solution. This result is mainly theoretical and in practice the runtimes are high for larger inputs, rendering the previous methods still useful.
The (heuristic) barycenter computed by these methods is again a discrete probability measure, in general with unequal mass on its points. Hence the interpretation of the barycenter as a point pattern is not straightforward. The different masses on the points have to be taken into account, while the definition of a point pattern does not provide any mass, only locations.

When the focus is put on the locations of the points itself and no mass is assigned at all, the differences in the numbers of points per pattern have to be compensated. Previously known distances for point patterns in this setting are the *spike time distance*, see Victor and Purpura (1997) and the *Optimal SubPattern Assignment (OSPA) metric*, see Schuhmacher and Xia (2008) and Schuhmacher et al. (2008). Both distances penalize differences in the number of points by a constant factor. The spike time distance allows for different constants depending on whether points have to be added or removed. Those constants have to be equal to fulfill the symmetry property of a metric. The OSPA metric only allows penaltys that are sufficiently large compared to the diameter of the point patterns. The *TT-metric* and *RTT-metric* are generalizations of both these metrics. The metric properties of the (R)TT-metric are satisfied for all choices of non-negative penalty terms and arbitrary distance functions between single points. The main advantage of the (R)TT-metric compared to the spike time distance is the algorithm for computing these metrics. The Incremental Matching Algorithm for computing the spike time distance, cf. Diez et al. (2012), has a worst-time complexity of $\mathcal{O}(n^6)$, where n is the number of points in the larger pattern. Calculating the spike time distance through the (R)TT-metric allows the metric to be computed with a worst-time complexity of $\mathcal{O}(n^3)$.

Diez et al. (2012) and Mateu et al. (2015) work with barycenters with respect to the spike time distance. They present a heuristic that is based on a kernel density estimation. Due to the high computation cost of the spike time distance, the studied examples are rather small.

The algorithm of Müller et al. (2020) relies on an alternating procedure to find a local optimum. The idea is similar to the well known k-means algorithm of Lloyd (1982). Section 5 of Müller et al. (2020) compares the new barycenter heuristic to other algorithms that generate point patterns as the resulting barycenter.

2.2 Main results

The two main achievements in Müller et al. (2020) are the construction of the new metrics (1.1) and (1.2), defined in Section 1.1, to measure distance between point patterns, and the fast barycenter heuristic for the computation of a Fréchet mean. This section is mainly dedicated to the development of the algorithm.

Distances in the (R)TT-metric are raised to the power p, cf. Theorem 1, distances in the Fréchet functional $\mathcal{F}$ are raised to the power q, cf. (1.4). When setting $p = q$, the Fréchet functional (1.4) can be simplified as follows:
For point patterns $\xi_j = \sum_{i=1}^{n_j} \delta_{x_{ij}}$, $j \in [k]$, let $\tilde{n} = \left\lfloor \frac{2}{k+1} \sum_{j=1}^{k} n_j \right\rfloor$ and $n \geq \max\{\tilde{n}, n_j;\ 1 \leq j \leq k\}$. Set $x_{ij} = \aleph$ for $n_j + 1 \leq i \leq n$ and $\tilde{\xi}_j = \sum_{i=1}^{n} \delta_{x_{ij}}$ for any $j \in [k]$.
Then for any k permutations $\pi_1, \ldots, \pi_k \in S_n$ jointly minimizing the term

$$\sum_{i=1}^{n} \min_{z \in \mathcal{X}'} \sum_{j=1}^{k} d'(x_{\pi_j(i),j}, z)^p, \tag{2.1}$$

the point pattern $\zeta_*|_{\mathcal{X}}$ with

$$\zeta_* = \sum_{i=1}^{n} \delta_{z_i}, \text{ where } z_i \in \operatorname*{argmin}_{z \in \mathcal{X}'} \sum_{j=1}^{k} d'(x_{\pi_j(i),j}, z)^p$$

is a p-th order barycenter with respect to the TT-metric. Any barycenter of the data cannot contain more than the calculated n points in $\mathcal{X}$. All point patterns are *filled* with points in $\aleph$ until they have the cardinality n. Then every point of the barycenter ζ_* is matched with one point (possibly $\aleph$) of every point pattern according to the TT-distance. Now the barycenter problem (1.4) cuts down to finding the k permutations

$\pi_{*,1}, \dots, \pi_{*,k}$ which correspond to the minimum matchings that are calculated in the metric d' and finding the locations of the barycenter points $z_1, \dots, z_n$.

If the optimal ζ_* is known, the optimal permutations π_i can be calculated in polynomial time. If the optimal permutations π_i are known, the barycenter ζ_* can be calculated. The heuristic algorithm makes use of this duality. The algorithm works similar to Lloyds algorithm for the k-means clustering: Start with a random point pattern ζ as candidate for the barycenter pattern. Then compute $\tau(\zeta, \xi_i)$ for every data point pattern ξ_i, $1 \leq i \leq k$. The assignment problems that is solved when computing τ define clusters of points. Any single point of ζ is matched with exactly one point of every data point pattern. All the points that are matched to a single point of ζ form a cluster. For each of these clusters a barycenter point is calculated. This is now a single point either in $\mathcal{X}$ or $\aleph$. These new barycenter points combined to a point pattern define an improved barycenter candidate. These steps are then iterated until the objective function value only improves marginally, i.e. the relative improvement to the objective function value from the previous iteration does not exceed a specified value.

Compared to the (mainly) LP based algorithms, this algorithm is very fast and it still produces reliable results. The details to this algorithm can be found in Section 4 of Müller et al. (2020), a simulation study and comparison to similar algorithms can be found in Section 5 of Müller et al. (2020). Note that the other algorithms that are used in the comparison treat the point patters as discrete probability measures. Currently this is the only algorithm that is optimized for the TT-distance.

2.3 Own contributions

When I began my research, I was able to rely on a draft by Dominic Schuhmacher for substantial parts of the theory on the new metric. I worked out the details and reduced the worst-time complexity for the calculation of the metrics from $\mathcal{O}(n^6)$ to $\mathcal{O}(n^3)$ for a general cutoff C. Before that, the faster runtime was only possible for large cutoffs C, i.e. in the OSPA metric. The algorithms and their implementations are joint work with Dominic Schuhmacher.

CHAPTER 3

ANOVA for data in metric spaces

This chapter summarizes the results of Müller et al. (2022b). It provides a literature overview on the latest methods for the Analysis Of Variance for non-Euclidean data, with a focus on a point pattern setting, in Section 3.1. A discussion of the main results follows in Section 3.2. In Section 3.3 my own contributions to the article are pointed out.

The impact of the article on future research is presented in an outlook in Chapter 5.

3.1 Literature overview

Analysis Of Variance is a common method to determine if differences between observations are random or statistically significant. The theory dates back to the work of Fisher (1925). Most classic ANOVA methods require independent normal-distributed observations in a Euclidean space. Beyond that, many methods have been developed for non-Euclidean data, see for example Cuevas et al. (2004) for ANOVA for functional data, Huckemann et al. (2009) for MANOVA on Riemannian manifolds, or Ramón et al. (2016) for point pattern data. The latter is using K-functions to measure distances between point patterns. When working with point patterns, using K-functions was a common method, see for example Diggle et al. (1991), Baddeley et al. (1993), Diggle et al. (2000), Landau and Everall (2008), Hahn (2012). A similar approach can be found in González et al. (2021), who developed a two-factor design for ANOVA for point pattern data.

The metric developed in Müller et al. (2020) makes it possible to treat the space of point patterns as a metric space, whereas e.g. K-functions transform the point patterns into data of a different metric space. Since this transformation is not a bijection, information of the data can be lost in this process. Dubey and Müller (2019) developed an ANOVA statistic for arbitrary metric spaces. However, this requires a central object, i.e. barycenter of the data. This is applicable for point pattern data, but calculating the exact barycenter of point patterns is computationally infeasible. Anderson (2001) developed an ANOVA statistic for non-Euclidean data that relies on distances between observations. Müller et al. (2022b) follow that approach and develop a new statistic that also does not require the calculation of a barycenter. The former statistic tests for differences in group means, i.e. it detects differences in the spatial distribution of the points, while the latter aims on detecting differences in dispersion, e.g. different interactions between points.

3.2 Main results

Müller et al. (2022b) present new statistics, based on Levene's test, for detecting differences in dispersion. The statistics can be applied to i.i.d. data in any metric space and, in contrast to most ANOVA statistics, do not require a barycenter or other kind of central object in the data space. They solely require pairwise distances between the observations. Müller et al. (2022b) also give an overview of procedures for testing group differences of data in metric spaces and compare the new statistics with previous methods in a simulation study. This section is mainly dedicated to the statistics and their asymptotic properties. A brief overview of the results of the simulation study is presented as well.

The new Levene's statistic L_M in (1.7) is already introduced in Section 1.3. This statistic is similar to the usual Levene's statistic, but includes the pairwise distances between observations as data, instead of distances to a Fréchet mean.

In the L_M-statistic, the numerator is divided by a variance estimator of the pairwise distances of the data, i.e. $\frac{1}{N-k}$RSS. The observations themselves are independent, but the pairwise distances used for the computations are not. In Section 1.3 it is mentioned, that the L_M-statistic is asymptotically χ^2 distributed up to a constant factor.

In the remainder of this section, the statistic L_M is slightly modified, such that it is asymptotically χ^2 distributed without any additional factors. To compensate for the dependence of the pairwise distances, the denominator of (1.7) is replaced by a *covariance* estimator. Leading to:

$$\widetilde{L} = \frac{N^* - k}{k-1} \frac{\frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_i n_j (\bar{d}_{i\cdot} - \bar{d}_{j\cdot})^2}{4\,T_n}, \tag{3.1}$$

where $N^* = \sum_{i=1}^{k} n_i(n_i - 1)^2$ and

$$T_n = \sum_{i=1}^{k} \sum_{\substack{j_1,j_2,j_3=1 \\ j_1 \notin \{j_2,j_3\}}}^{n_i} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)\big(d_{i,\{j_1,j_3\}} - \bar{d}_{i\cdot}\big). \tag{3.2}$$

This statistic $\tilde{L}$ is asymptotically χ^2 distributed:

Theorem 2. *Assume that the Borel σ-algebra for $(\mathcal{X}, d)$ is countably generated. In the usual 1-way setup of Subsection* 4.1 *of Müller et al. (2022b) assume that $P_1 = \ldots = P_k = P$ for a distribution P that is not a Dirac distribution and satisfies $\int_{\mathcal{X}} \int_{\mathcal{X}} d^2(x, y)\, P(dx)\, P(dy) < \infty$. Suppose that there are $\lambda_i > 0$ such that $n_i/n \to \lambda_i$ for every i as $n \to \infty$. Then*

$$(k-1)\, \widetilde{L} \xrightarrow{\mathcal{D}} \chi^2_{k-1} \quad \text{as } n \to \infty.$$

With this result it follows that the statistic L_M is asymptotically χ^2 distributed up to a factor. This factor only depends on the variance and covariance of the data. So for any fixed distribution P this factor is constant.

Corollary 3. *Under the conditions of Theorem 2*

$$(k-1)\, L_M \xrightarrow{\mathcal{D}} \frac{4\gamma^2}{\sigma^2} \chi^2_{k-1} \quad \text{as } n \to \infty,$$

where

$$\gamma^2 = \mathrm{Cov}(d(X, Y), d(X, Z)),$$
$$\sigma^2 = \mathrm{Var}(d(X, Y))$$

with independent $X, Y, Z \sim P$.

More details on these statistics and the proofs of the statements can be found in Section 4 of Müller et al. (2022b). In Section 6.3 of Müller et al. (2022b) a simulation study examines the convergence rate of the statistic $\widetilde{L}$ for point pattern data.

These statistics do not rely on a specific structure of the data or the underlying space. They can be applied to independent data in any metric space. However, in Müller et al. (2022b) they are primarily tested for point pattern data, as for point pattern data the calculation of a barycenter is practically infeasible, cf. Chapter 2. A simulation study reveals that for groups of observed point patterns the L_M-statistic works better than comparable methods when it comes to detecting differences in dispersion of the data. An extensive study of different point pattern statistics with different kind of data is provided in Section 6 of Müller et al. (2022b). The main results are summarized here.

Point pattern distributions can deviate in infinitely ways from *complete spatial randomness* (*CSR*), i.e. all the points are drawn independently from a uniform distribution on a compact space. In the simulations this compact space is assumed to be the unit square in $\mathbb{R}^2$.

The simulation study concentrates on two ways in which the the point pattern distribution (*point process*) deviates from CSR:

- inhomogeneity: the probability of a point $(x, y) \in [0, 1]^2$ belonging to a point pattern drawn from this distribution is not equal for all $(x, y) \in [0, 1]^2$, some points or areas are more likely than others.
- point interaction: points of single pattern attract or repel each other.

In the simulations the point interaction is focused on point repulsion. This is realized through a Strauss point process. Further details on the Strauss point process can be found in Kelly and Ripley (1976).

The different levels of interaction are marked by the interaction parameter γ, where $\gamma = 1$ means CSR, there is no interaction at all, and $\gamma = 0$ (*Strauss hard core process*) means no two points of a pattern can lie closer than a specified distance R from each other. These explanations are only meant to give a general impression of how the point processes deviate from each other. For the complete definitions of the scenarios, see Section 6 in Müller et al. (2022b).

In the study $k = 2$ groups of $\tilde{n} = 20$ point patterns are compared. All parameters are set in such a way that the expected number of points per pattern is 35. The interaction radius for the Strauss processes is $R = 0.1$. All tests are performed at a significance level $\alpha = 5\%$.

In the first study a group of point patterns that are drawn from a CSR distribution is tested against a group of point patterns drawn from an inhomogeneous distribution. In total there are six inhomogeneous scenarios. Example point patterns of scenarios 1 to 6 are visible in Figure 3.1, Scenario 0 is CSR. Details about the spatial distribution λ of the points in the patterns is listed in Table 3.1.

The results of the simulation study are listed in Table 3.2. The tests Fréchet T_F and Fréchet T_L denote the ANOVA statistic and Levene's statistic from Dubey and Müller (2019), respectively. In these scenarios the F-tests F_A and T_F are expected to perform better than the Levene's tests L_M and T_L.

Scenario	$\lambda(x,y)$ proportional to	Scenario	$\lambda(x,y)$ proportional to
1	$\sum_{i=1}^{3} \varphi_{\mu_i,0.075}(x,y)$	4	$\exp(-2x)$
2	$\sum_{i=1}^{3} \varphi_{\mu_i,0.1}(x,y)$	5	$\exp(-1x)$
3	$\sum_{i=1}^{4} \varphi_{\mu_i,0.1}(x,y)$	6	$\exp(-0.02x)$

Table 3.1: Overview of the Poisson process intensities for the six scenarios. The proportionality constant is chosen such that the expected number of points in each scenario is 35. φ_{μ,σ^2} denotes the density of the bivariate normal distribution with mean $\mu \in \mathbb{R}^2$ and covariance matrix $\sigma^2 I$. The different μ_i used are visible in Figure 3.1.

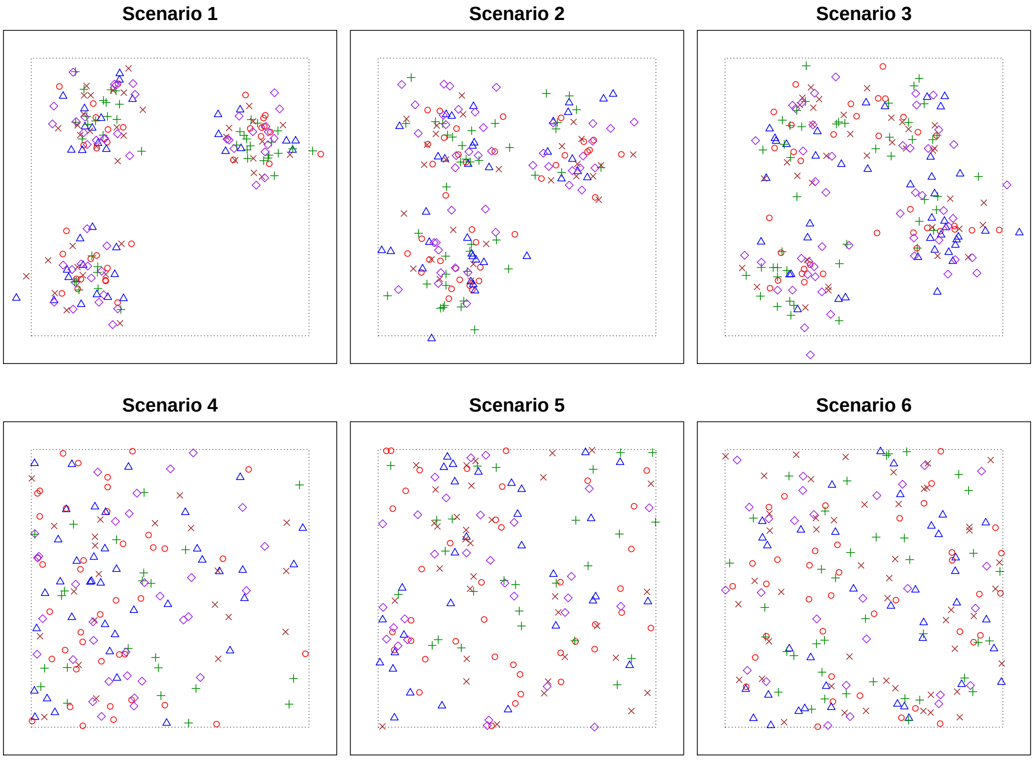


Figure 3.1: Six scenarios of sample point patterns in the unit square (dotted line) in which the second group (CSR) is sampled. Points outside the unit square are not rejected, since this rarely occurs. In each scenario five sample point patterns are plotted, each point pattern is marked by one distinct symbol and color.

It is visible that the F-tests, Anderson F_A and Fréchet T_F, do perform better than L_M and T_L. However for both the F-statistics and the Levene's statistics, the distance based tests work better than the statistics that use a Fréchet mean of the point patterns.

In the second study a group of point patterns that are drawn from CSR is tested against a group of point patterns drawn from a Strauss distribution with six different

Scenario	1	2	3	4	5	6	0
Anderson F_A	100	100	100	100	99	39	2
L_M	93	76	77	14	7	9	3
Fréchet T_F	100	100	100	99	11	0	4
Fréchet T_L	59	24	47	13	9	12	4

Table 3.2: Numbers of rejections of the null hypothesis "equal distribution in both groups" based on 100 data sets per column. In each data set the first group is sampled from the scenario indicated in the column and the second group is sampled from CSR.

interaction parameters $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$. The six different γ define six different scenarios. First, groups of point patterns that are drawn with $\gamma = 1$ (i.e. CSR) are tested against groups drawn from all the six scenarios. And second, groups of point patterns drawn from a Strauss hard core process, i.e. $\gamma = 0$, are tested against the six scenarios. In Figure 3.2 it is visible how the interaction parameters influences the spatial distribution of the points in a single pattern. The first image shows a realization of a hard core process. No two points lie closer than the interaction radius $R = 0.1$ in Euclidean distance together. The sixth image, on the bottom right, shows a point pattern drawn from CSR.

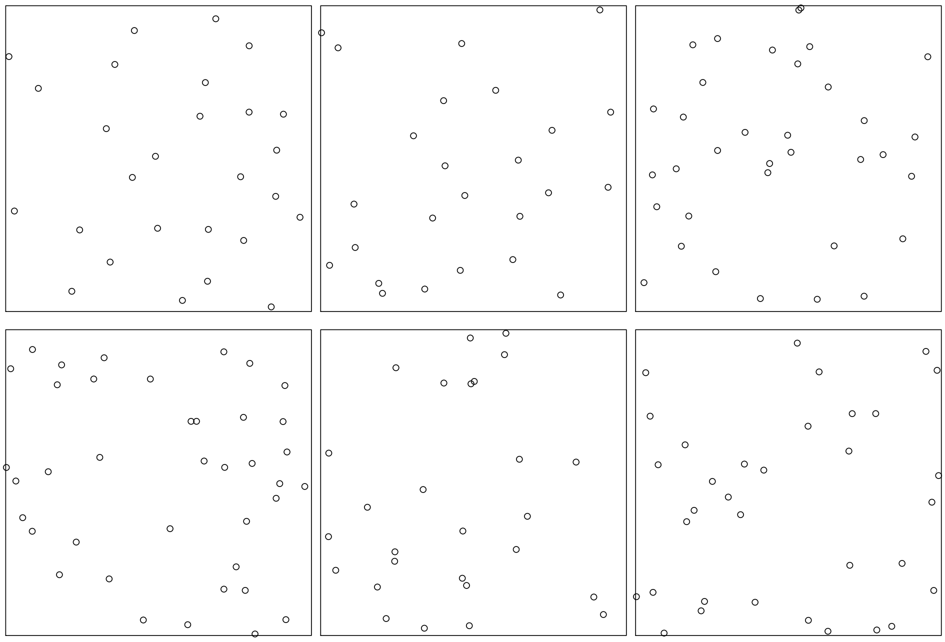


Figure 3.2: Simulations from Strauss distributions, where rowwise from left to right $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$. For $\gamma = 0$ this is a realization of a hard core process, for $\gamma = 1$ a realization from CSR.

The interaction between points of a single pattern is not necessarily visible in their Fréchet mean. Hence in this second study the Levene's statistics L_M and T_L are expected to perform better than the F statistics F_A and T_F. The results of the simulation study are listed in Table 3.3.

$\gamma = 1$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	98	41	13	8	4	5
L_M	100	100	95	67	20	3
Fréchet T_F	100	98	78	45	20	4
Fréchet T_L	100	99	88	45	20	4

$\gamma = 0$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	3	55	90	98	97	99
L_M	11	60	96	100	100	100
Fréchet T_F	6	57	91	97	100	100
Fréchet T_L	9	33	82	95	99	100

Table 3.3: Numbers of rejections of the null hypothesis “equal distribution in both groups” based on 100 data sets per column. In each data set the point patterns in both groups are sampled from a Strauss distribution. The first group is sampled using $\gamma = 1$ or $\gamma = 0$ as indicated on the top left of the table and the second group uses γ as indicated in the column. The other parameters are set in such a way that the expected number of points per pattern is 35.

As expected, the distance based Levene's test L_M performs very well. However, in the second case, the hard core process ($\gamma = 0$) against the six scenarios, the F-tests have a good performance as well.

The conclusion based on these results is that the combination of the distance based statistics Anderson F_A and the L statistic performs similar to the combined statistic of Dubey and Müller (2019), which is based on Fréchet means. In these scenarios the computation of the distance based statistics takes a few seconds, while the computation of the Fréchet statistics takes several minutes.
Currently the computation of a Fréchet mean of point patterns is computationally infeasible in general. Thus, the statistics that rely solely on distances between observations are a good alternative in settings where the computation of a central object of the data is costly.

3.3 Own contributions

The statistics were developed in collaboration between all three co-authors. The theoretical results were obtained jointly with Dominic Schuhmacher. Testing out different versions of statistics and all the programming was done by me. The interpretation of results was then again a collaborative part of all three co-authors.

CHAPTER 4

Location problems with cut off distances

This chapter summarizes the results of Müller et al. (2022a). It starts in Section 4.1 with a literature overview on problems related to the ones treated in Müller et al. (2022a). A discussion of the main results follows in Section 4.2. In Section 4.3 my own contributions to the article are pointed out.
The impact of the article on future research is presented in an outlook in Chapter 5.

4.1 Literature overview

The location problems of Müller et al. (2022a) are based on the *Weber* problem, Weber (1909). This problem itself is based on the *Fermat* problem, which was formulated by *Pierre de Fermat* before 1640. The Fermat problem is about finding a point that minimizes its total distance to the three vertices of the triangle. In tribute to Pierre de Fermat this point is called *Fermat Point*. The Fermat problem with weighted distances dates back to Simpson (1750) and was made famous by Weber (1909). Nowadays the term Weber problem refers to the problem of finding a point that minimizes its total weighted distances to a set of n fixed points, with respect to some metric d. This problem is also called *One-Median* problem or *Minisum* problem.
The Weber problem, (Bar($\mathcal{A}$)), is a well studied subject. For the $d = \ell_2$ metric the optimal solution can be computed with the Weiszfeld algorithm, Weiszfeld (1937), for $d = \ell_1$ an optimal solution is the coordinate wise median of the n points and for $d = \ell_2^2$ the unique optimal solution is the arithmetic mean of the n points.
A more general version of this problem is the *multisource* Weber problem. The objective is to place m points (sources) in the metric space to minimize the sum of distances between the n fixed points (locations) and their respective closest source. This problem is relevant for example in logistics, where the m sources are warehouses,

or the m sources are store branches that try to attract n customers. For an overview of different formulations of location problems in operations research and methods to solve them see for example Love et al. (1988).

Müller et al. (2022a) study the Weber problem, where the underlying metric is cut off at a certain value $C > 0$, meaning the maximal distance between two points is bounded by C. The idea to the Weber problem with cut off distances originates from Müller et al. (2020), where the distances between point patterns are cut off at a given value. However, it has first been studied by Drezner et al. (1991), who looks at instances in the plane with the Euclidean metric and the Manhattan metric and allows for different cutoffs for the different given points. They also present an algorithm that solves this problem with $\mathcal{O}(n^2)$ calls of an oracle which solves the Weber problem for a subset of the n points. In the plane the number of these subsets is bounded by $2n^2$. Drezner's algorithm works by intelligently enumerating said subsets. Later Aloise et al. (2012) and Venkateshan (2020) found out, that there are instances where the algorithm of Drezner et al. (1991) fails to find the optimal solution and present improved algorithms.

4.2 Main results

Müller et al. (2022a) studies the connection between the Weber problem and the Weber problem with a cut off distance. It also presents a new problem, where the option to choose an *empty* solution at fixed cost is allowed.
Recall that this section considers a finite set $\mathcal{A} \subseteq \mathbb{R}^k$, $|\mathcal{A}| = n$, a metric d and $q \geq 1$. Let $\operatorname{diam}(\mathcal{A}) = \max\limits_{a_1,a_2 \in \mathcal{A}} d(a_1, a_2)$ the diameter of $\mathcal{A}$.
The Weber problem, $(\mathrm{Bar}(\mathcal{A}))$, is already well studied for many different metrics. The first extension of the problem, $(\mathrm{Bar}_C(\mathcal{A}))$, is closely related to the problem presented by Drezner et al. (1991), while the second extension, $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$, has not been considered in the literature so far. In this section the relations between the classical problem and the two extensions are presented. This includes settings in which solving the easier problem $(\mathrm{Bar}(\mathcal{A}))$ is guaranteed to yield an optimal solution to the extended problems, and settings in which solving $(\mathrm{Bar}(\mathcal{A}))$ in general does not suffice. The algorithm of Drezner et al. (1991) is tailored to the extended problems and for each of the extended problems improved versions of the algorithm are introduced. Additionally an algorithm that solves $(\mathrm{Bar}_C(\mathcal{A}))$ for all cutoff values C is presented.

4.2.1 Relation between Bar and Bar$_C$

From the algorithms of Drezner et al. (1991) and Müller et al. (2022a) it is known that can be solved (Bar$_C(\mathcal{A})$) efficiently as long as (Bar($\mathcal{A}$)) can be solved. Section 4 of Müller et al. (2022a) studies the connection between those two problems and under which circumstances solving (the easier problem) (Bar($\mathcal{A}$)) is sufficient. The meaning of sufficiency is twofold:

(a) (Bar) has the same objective function value as (Bar$_C$), i.e. $\mathcal{Z}^* = \mathcal{Z}^*_C$.

(b) Any solution to (Bar) is a solution to (Bar$_C$), i.e., $\mathcal{X}^* \subseteq \mathcal{X}^*_C$.

Under condition (a) solving (Bar($\mathcal{A}$)) will yield the objective function value of (Bar$_C(\mathcal{A})$), and under condition (b) an optimal solution is calculated. These conditions are connected.

Lemma 4. *If condition* (*a*) *holds then* (*b*) *holds as well.*

The proof of this statement can be found in Müller et al. (2022a). The implication $(b) \Rightarrow (a)$ is in general not true:

Example 1 (Counterexample to (b) ⇒ (a))**.**
Let 4 points in $\mathbb{R}$ *be given,* $a_1 = a_2 = 0$, $a_3 = C + \varepsilon$ *and* $a_4 = -C - \varepsilon$ *and let* $q = 1$. *The solution to both problems,* (Bar($\mathcal{A}$)) *and* (Bar$_C(\mathcal{A})$) *is* $\mathcal{X}^* = \mathcal{X}^*_C = \{0\}$, *but* $\mathcal{Z}^* = 2(C + \varepsilon) > 2C = \mathcal{Z}^*_C$.

To investigate further under which circumstances it is sufficient to solve (Bar) instead of (Bar$_C$) or where it will in general never be sufficient, define the following property.

$$\text{For any subset } A \subseteq \mathcal{A} : \mathcal{X}^*(A) \cap \operatorname{conv}(\mathcal{A}) \neq \emptyset, \tag{conv}$$

where conv(A) denotes the *convex hull* of the set A and $\mathcal{X}^*(A)$ denotes the set of optimal solutions to (Bar(A)). The property (conv) is a property of the space $\mathbb{R}^k$ in combination with the metric d and the power $q \geq 1$. For example in the space $\mathbb{R}^k$ equipped with the squared Euclidean norm, i.e. $q = 2$, (conv) is true, since the (unique) optimal solution to (Bar) is the arithmetic mean of the points in $\mathcal{A}$. Therefore, the optimal solution is a convex combination of the points. In $(\mathbb{R}^k, d), k \in \{1, 2\}$ (conv) is also true for any distance d that is derived from a norm. For $k > 2$ on the other hand it is in general *only* true for distances d that are linearly equivalent to the ℓ_2-norm, see Plastria (1984). An example where (conv) is not true is $(\mathbb{R}^3, \ell_1)$:

Example 2 (Example where (conv) is not true)**.**
Consider the set $\mathcal{A} = \{a_1, a_2, a_3\} \subseteq \mathbb{R}^3$ *with* $a_1 = (1,0,0), a_2 = (0,1,0), a_3 = (0,0,1)$*. Let* $d = \ell_1$ *and* $q = 1$*. Then* $\mathcal{X}^* = \{(0,0,0)\}$*, but* $(0,0,0) \notin \operatorname{conv}(\mathcal{A})$*.*

Depending on the diameter $\operatorname{diam}(\mathcal{A})$ condition (a) (or (b)) is always true, sometimes true or never true, as follows from the following table

	$\operatorname{diam}(\mathcal{A}) \leq \sqrt[q]{C}$	$\sqrt[q]{C} < \operatorname{diam}(\mathcal{A}) \leq 2\sqrt[q]{C}$	$\operatorname{diam}(\mathcal{A}) > 2\sqrt[q]{C}$
(a) $\mathcal{Z}^* = \mathcal{Z}^*_C$	holds if (*conv*) see Thm 17	may or may not hold, see Examples 2 to 5	never for $q = 1$, never for $q > 1$ for ℓ_p-norms, see Thm 16
(b) $\mathcal{X}^* \subseteq \mathcal{X}^*_C$	holds if (*conv*) follows from Lem 14	may or may not hold, see Examples 2 to 5	may or may not hold, see Examples 1 and 3

Table 4.1: Overview of sufficiency under different assumptions. Taken from Müller et al. (2022a), Section 4.

4.2.2 Relation between Bar_C and $\text{Bar}_{C,\alpha}$

Similar to the previous section, the relation between $\text{Bar}_{C,\alpha}$ and Bar_C can be studied. For a finite set $\mathcal{A}$, with $n = |\mathcal{A}|$, the objective function value of $\emptyset$ is always $n \cdot \alpha \cdot C$. That means to solve ($\text{Bar}_{C,\alpha}(\mathcal{A})$) it is sufficient to solve ($\text{Bar}_C(\mathcal{A})$) and compare the objective function value of the optimal solution to $n \cdot \alpha \cdot C$.
However, there are criteria under which the empty set must be an optimal solution, or can never be an optimal solution to ($\text{Bar}_{C,\alpha}(\mathcal{A})$).

All references hereafter refer to the respective results in Müller et al. (2022a).
The empty barycenter is always an optimal solution to ($\text{Bar}_{C,\alpha}(\mathcal{A})$) if

- there is no ball B with radius $\sqrt[q]{C}$ that contains at least $(1-\alpha)\cdot n$ points, cf. Lemma 19
- $\alpha \leq \min\left\{\frac{1}{2^q}\frac{\operatorname{diam}(\mathcal{A})^q}{n \cdot C}, \frac{1}{n}\right\}$, cf. Lemma 26.

In contrast, the empty barycenter cannot be an optimal solution in the following cases:

- $\alpha > \frac{n-1}{n}$, see Lemma 21,
- $\alpha > \frac{\operatorname{diam}(\mathcal{A})^q}{C} \cdot \frac{n-1}{n}$, see Corollary 25.

The improvements to the algorithm of Drezner et al. (1991) in this setting are based on these results. The details to the algorithms can be found in Müller et al. (2022a), Algorithm 2 and Algorithm 3.

For the idea of the algorithms one observation is crucial. The optimal solution to ($\text{Bar}_C(\mathcal{A})$) is the optimal solution to $\text{Bar}(A)$ for some subset $A \subseteq \mathcal{A}$. In general there are $2^n - 1$ possible subsets A. However, Drezner et al. (1991) proved that the optimal solution can be calculated by looking at only polynomially many subsets. The improved versions of the algorithm further reduce the number of these subsets by computing lower bounds on the objective function value and skipping parts of the computation, without losing information w.r.t. the optimal solution. Section 8 of Müller et al. (2022a) provides a study for different scenarios and states results for the performance of the improvements.
The theory for Algorithm 4, that solves ($\text{Bar}_C(\mathcal{A})$) for all values of the cutoff C can be found in Section 6 of Müller et al. (2022a). The algorithm starts with two solutions that can be computed easily. After that more solutions are calculated by a routine similar to a bisection method. By solving ($\text{Bar}_C(\mathcal{A})$) n times it is solved for all C.

4.3 Own contributions

Most parts are the collaboration of all authors. Chapters 4, 5 and 8 are largely my own work.

CHAPTER 5

Discussion and outlook

Müller et al. (2020) introduced the (R)TT-metric and a barycenter heuristic. The metric fits well into pre-existing metrics and can be computed quickly. Although the results of the barycenter heuristic are reliable, it might be interesting for future research to find an algorithm with an approximation guarantee. The algorithm of Lloyd (1982) allows for no approximation guarantees. Hence it might be difficult to prove upper bounds on the approximation for the barycenter heuristic, which uses the same principle as the algorithm of Lloyd (1982).

Most algorithms for similar data are based on linear programming. A new algorithm that is also based on linear programming and approximates the barycenter with respect to the (R)TT-metric could improve the computed pseudo-barycenters.

It requires further research, whether the barycenter computation with respect to the (R)TT-metric is $\mathcal{NP}$-hard. The proof would render the heuristic even more notable. A proof for the barycenter problem with respect to the (R)TT-metric is not straightforward, although proofs for similar problems exist. However, the result does not help improving the calculations of the barycenter.

Barycenters open the door to more advanced studies like the ANOVA that is presented in Müller et al. (2022b). But also methods like time series analyses or Fréchet regression are possible research topics. However, further research in these topics is beyond the scope of this work.

The Levene's L_M-statistics of Müller et al. (2022b) is an improvement to previous point pattern statistics, in terms of runtime as well as in terms of performance. In combination with the F_A-statistic of Anderson (2001) it is able to detect differences in spatial distribution and interaction of the points in two or more groups of point patterns. The Levene's statistic might be improved by an unbiased estimator of the true covariance γ of the data. The presented statistic uses a straightforward, but biased estimator. A major bottleneck at this end is the structure of the covariance of two pairwise distances. Three random variables are factored in, hence the structure of the covariance estimator is complicated, and pruning it to become unbiased is not straightforward.

Additionally, the concept of doing statistics based on the pairwise distances between the observations can be adapted to other statistics. Müller et al. (2022b) present a two factor Levene's test. This can be generalized to more than two factors. The design of such a statistic follows the formulation of a k-factor F-test, but is beyond the scope of this thesis.

Another possibility is to use different sums of squares for the statistics. But also completely different statistics might be the basis for a new distance based statistic, e.g. the F-test for variances or the χ^2-test for the equality of distributions. For the presented statistics the substitution of the central object is straightforward. For more complex designs and statistics replacing the central object might require a lot more effort and research.

In Müller et al. (2022a) the Weber problem, ($\mathrm{Bar}(\mathcal{A})$), and the two extensions ($\mathrm{Bar}_C(\mathcal{A})$) and ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) are investigated. The connection between these three problems is studied, especially under which circumstances the extensions can be reduced to ($\mathrm{Bar}(\mathcal{A})$). For each of the two extensions, an algorithm is introduced which reduces the computation times, compared to the algorithm presented by Drezner et al. (1991). It can be investigated if theses algorithms can be generalized to the setting of Drezner et al. (1991), where each point of $\mathcal{A}$ has its own cutoff. The improvements of the algorithm of Drezner et al. (1991) are based on lower bounds on the objective function value, which can be computed early in each iteration. When every point has its own cutoff value, these lower bounds can still be computed with more effort. It yields an interesting perspective if this still significantly improves the runtime of the algorithm.

Although there are many criteria presented in which the extensions *can* be reduced to ($\mathrm{Bar}(\mathcal{A})$), and also in which it is certain, that solving ($\mathrm{Bar}(\mathcal{A})$) does not suffice, there are scenarios in which it is not certain, if solving ($\mathrm{Bar}(\mathcal{A})$) calculates an optimal solution to ($\mathrm{Bar}_C(\mathcal{A})$) or ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$). Finding new criteria which guarantee that a solution to ($\mathrm{Bar}(\mathcal{A})$) is also a solution to the extended problems is an interesting topic for future research. For the problem ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) it is further worthwile finding more criteria under which the empty barycenter is an optimal solution. Criteria which solely use the geometric structure of the data are investigated in this thesis, further criteria are beyond the scope of this work.

When the investigated metric d is the ℓ_1 metric and $q = 1$, then problem ($\mathrm{Bar}(\mathcal{A})$) can be solved quickly, by optimizing separately over the k dimensions. It is still

an open question whether ($\mathrm{Bar}_C(\mathcal{A})$) can be simplified in this, or a comparable, setting. Due to the structure of the ℓ_1 metric one optimal solution to ($\mathrm{Bar}_C(\mathcal{A})$) lies on a grid of n^k points in a k dimensional space. For each point of the grid, lower bounds on the objective function value can be computed quickly. This can be used to further improve the algorithms for solving ($\mathrm{Bar}_C(\mathcal{A})$).
On the algorithmic side, it might be time saving to split the problem ($\mathrm{Bar}_C(\mathcal{A})$) into smaller subproblems, where the solution of one of the subproblems is an optimal solution to the whole problem. Criteria to determine if this is possible for a given instance have to be investigated.

A simulation study in Section 8.1 of Müller et al. (2022a) has shown, that the barycenter algorithm of Müller et al. (2020) is not improved when the subproblem ($\mathrm{Bar}_C(\mathcal{A})$) is solved exactly.
The results in Müller et al. (2022b) suggest that the distance based statistics perform as well as the Fréchet based statistics, while their computation times are a lot faster. E.g. the distance based permutation tests in Table 3.3 are calculated in a few seconds, while the Fréchet based tests take several minutes, even with the fast barycenter heuristic of Müller et al. (2020) instead of the exact barycenters.
Overall, the results of this thesis indicate that the exact computation of barycenters of point patterns is of little avail, since the computation times are large and the barycenter heuristic produces consistent results. For statistical purposes the ANOVA procedures that are based on pairwise distances perform similar to the methods that rely on barycenters, while the distance based statistics allow for considerably faster computation times.

Bibliography

Aloise, D., Hansen, P., and Liberti, L. (2012). An improved column generation algorithm for minimum sum-of-squares clustering. *Mathematical Programming*, 131(1-2):195–220.

Altschuler, J. M. and Boix-Adserà, E. (2021). Wasserstein barycenters can be computed in polynomial time in fixed dimension. *Journal of Machine Learning Research*, 22(44):1–19.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral ecology*, 26(1):32–46.

Anderson, M. J. (2017). Permutational multivariate analysis of variance (PERMANOVA). *Wiley statsref: statistics reference online*, pages 1–15.

Baddeley, A., Moyeed, R., Howard, C., and Boyde, A. (1993). Analysis of a three-dimensional point pattern with replication. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 42(4):641–668.

Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. Chapman and Hall/CRC.

Bertsekas, D. P. (1988). The auction algorithm: A distributed relaxation method for the assignment problem. *Annals of operations research*, 14(1):105–123.

Błaszczyszyn, B., Haenggi, M., Keeler, P., and Mukherjee, S. (2018). *Stochastic geometry analysis of cellular networks*. Cambridge University Press.

Borgwardt, S. (2020). An LP-based, strongly-polynomial 2-approximation algorithm for sparse Wasserstein barycenters. *Operational Research, online.*

Borgwardt, S. and Patterson, S. (2020). A column generation approach to the discrete barycenter problem. *Preprint*. Available at `https://arxiv.org/abs/1907.01541`.

Borgwardt, S. and Patterson, S. (2021). On the computational complexity of finding a sparse Wasserstein barycenter. *Journal of Combinatorial Optimization*, 41(3):736–761.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. (2018). Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609.

Cuevas, A., Febrero, M., and Fraiman, R. (2004). An ANOVA test for functional data. *Computational statistics & data analysis*, 47(1):111–122.

Cuturi, M. and Doucet, A. (2014). Fast computation of Wasserstein barycenters. In Xing, E. P. and Jebara, T., editors, *Proceedings of the 31st International Conference on Machine Learning*, pages 685–693.

del Barrio, E., Cuesta-Albertos, J. A., Matrán, C., and Mayo-Íscar, A. (2019). Robust clustering tools based on optimal transportation. *Statistics and Computing*, 29(1):139–160.

Diez, D. M., Schoenberg, F. P., and Woody, C. D. (2012). Algorithms for computing spike time distance and point process prototypes with application to feline neuronal responses to acoustic stimuli. *Journal of Neuroscience Methods*, 203(1):186–192.

Diggle, P. J. (2013). *Statistical analysis of spatial and spatio-temporal point patterns.* Chapman and Hall/CRC.

Diggle, P. J., Lange, N., and Beneš, F. M. (1991). Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. *Journal of the American Statistical Association*, 86(415):618–625.

Diggle, P. J., Mateu, J., and Clough, H. E. (2000). A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Advances in Applied Probability*, 32(2):331–343.

Drezner, Z., Mehrez, A., and Wesolowsky, G. O. (1991). The facility location problem with limited distances. *Transportation Science*, 25(3):183–187.

Dubey, P. and Müller, H.-G. (2019). Fréchet analysis of variance for random objects. *Biometrika*, 106(4):803–821.

Fisher, R. (1925). *Statistical Methods for Research Workers.* Oliver & Boyd.

Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré*, 10(4):215–310.

González, J. A., Lagos-Álvarez, B. M., and Mateu, J. (2021). Two-way layout factorial experiments of spatial point pattern responses in mineral flotation. *TEST*, 30(4):1046–1075.

Hahn, U. (2012). A studentized permutation test for the comparison of spatial point patterns. *Journal of the American Statistical Association*, 107(498):754–764.

Hamidi, B., Wallace, K., Vasu, C., and Alekseyenko, A. V. (2019). W_d^*-test: robust distance-based multivariate analysis of variance. *Microbiome*, 7(1):1–9.

Heinemann, F., Munk, A., and Zemel, Y. (2021). Randomised Wasserstein barycenter computation: Resampling with statistical guarantees. *Preprint.* Available at `https://arxiv.org/abs/2012.06397`.

Huckemann, S., Hotz, T., and Munk, A. (2009). Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):593–603.

Kelly, F. P. and Ripley, B. D. (1976). A note on Strauss's model for clustering. *Biometrika*, 63(2):357–360.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97.

Landau, S. and Everall, I. P. (2008). Nonparametric bootstrap for k-functions arising from mixed-effects models with applications in neuropathology. *Statistica Sinica*, 18(4):1375–1393.

Lehmann, E. L., Romano, J. P., and Casella, G. (2005). *Testing statistical hypotheses*, volume 3. Springer.

Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137.

Love, R. F., Morris, J. G., and Wesolowsky, G. O. (1988). Facilities location: Models & methods. *Publications in Operations Research*, 7.

Mateu, J., Schoenberg, F. P., Diez, D. M., Gonzáles, J. A., and Lu, W. (2015). On measures of dissimilarity between point patterns: Classification based on prototypes and multidimensional scaling. *Biometrical Journal*, 57(2):340–358.

Müller, R., Schöbel, A., and Schuhmacher, D. (2022a). Location problems with cutoff. *Preprint.* Available at https://arxiv.org/abs/2203.00910.

Müller, R. and Schuhmacher, D. (2021). *ttbary: Barycenter Methods for Spatial Point Patterns.* R package version 0.2-0. https://CRAN.R-project.org/package=ttbary.

Müller, R., Schuhmacher, D., and Mateu, J. (2020). Metrics and barycenters for point pattern data. *Statistics and Computing*, 30(4):953–972.

Müller, R., Schuhmacher, D., and Mateu, J. (2022b). ANOVA for data in metric spaces, with applications to spatial point patterns. *Preprint.* Available at https://arxiv.org/abs/2201.08664.

Plastria, F. (1984). Localization in single facility location. *European Journal of Operational Research*, 18(2):215–219.

Ramón, P., de la Cruz, M., Chacón-Labella, J., and Escudero, A. (2016). A new non-parametric method for analyzing replicated point patterns in ecology. *Ecography*, 39(11):1109–1117.

Rice, J. A. (2006). *Mathematical statistics and data analysis.* Brooks/Cole, 3 edition.

Schoenberg, F. P. and Tranbarger, K. E. (2008). Description of earthquake aftershock sequences using prototype point patterns. *Environmetrics*, 19(3):271–286.

Schuhmacher, D., Vo, B.-T., and Vo, B.-N. (2008). A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Processing*, 56(8, part 1):3447–3457.

Schuhmacher, D. and Xia, A. (2008). A new metric between distributions of point processes. *Advances in Applied Probability*, 40(3):651–672.

Simpson, T. (1750). *The Doctrine and Application of Fluxions.* John Nourse.

Venkateshan, P. (2020). A note on "The facility location problem with limited distances". *Transportation Science*, 54(6):1439–1445.

Victor, J. D. and Purpura, K. P. (1997). Metric-space analysis of spike trains: Theory, algorithms and application. *Network: Computation in Neural Systems*, 8(2):127–164.

Weber, A. (1909). *Über den Standort der Industrien. Erster Teil. Reine Theorie des Standorts.* Mohr, Tübingen.

Weiszfeld, E. (1937). Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematical Journal*, 43:355–386.

Addenda

APPENDIX A

Metrics and barycenters for point pattern data

Statistics and Computing (2020) 30:953–972
https://doi.org/10.1007/s11222-020-09932-y

Metrics and barycenters for point pattern data

Raoul Müller[1] · Dominic Schuhmacher[1] · Jorge Mateu[2]

Received: 19 September 2019 / Accepted: 11 February 2020 / Published online: 24 February 2020

Abstract
We introduce the transport–transform and the relative transport–transform metrics between finite point patterns on a general space, which provide a unified framework for earlier point pattern metrics, in particular the generalized spike time and the normalized and unnormalized optimal subpattern assignment metrics. Our main focus is on barycenters, i.e., minimizers of a q-th-order Fréchet functional with respect to these metrics. We present a heuristic algorithm that terminates in a local minimum and is shown to be fast and reliable in a simulation study. The algorithm serves as a general plug-in method that can be applied to point patterns on any state space where an appropriate algorithm for solving the location problem for individual points is available. We present applications to geocoded data of crimes in Euclidean space and on a street network, illustrating that barycenters serve as informative summary statistics. Our work is a first step toward statistical inference in covariate-based models of repeated point pattern observations.

Keywords Fréchet mean · Fréchet median · Network · Optimal transport · Point process · Unbalanced · Wasserstein

1 Introduction

Point pattern data are abundant in modern scientific studies. From biomedical imagery over geo-referenced disease cases and positions of mobile phone users to climate change-related space–time events, such as landslides, we have more and more complicated data available. See Chiaraviglio et al. (2016), Lombardo et al. (2018), Konstantinoudis et al. (2019), Samartsidis et al. (2019) for individual examples and the textbooks Diggle (2013), Baddeley et al. (2015), Błaszczyszyn et al. (2018) for a broad overview of further applications. While a few decades ago, data consisted typically of a single point pattern in a low-dimensional Euclidean space, maybe with some low-dimensional mark information, we have nowadays often multiple observations of point patterns available that may live on more complicated spaces, e.g., manifolds (including shape spaces), spaces of convex sets or function spaces. A setting that has received a particularly large amount of attention recently is point patterns on graphs, such as street networks, see Moradi et al. (2018), Moradi and Mateu (2019) and Rakshit et al. (2019) among others.

Multiple point pattern observations may occur by i.i.d. replication (e.g., of a biological experiment), but may also be governed by one or several covariates or form a time series of possibly dependent patterns. Additional mark information can easily be high-dimensional. Methodology for treating such point pattern data in all these situations is the subject of ongoing statistical research, see, e.g., Baddeley et al. (2015).

From a more abstract point of view, consider the set $\mathfrak{N}_{\text{fin}}$ of finite counting measures on some metric space $(\mathcal{X}, d)$. If we manage to equip $\mathfrak{N}_{\text{fin}}$ with a metric τ that reflects the concept of distance between point patterns in an appropriate problem-related way, there are a number of standard methods which can be applied, including multidimensional scaling, discriminant and cluster analysis techniques. This is a stance already taken in Schuhmacher (2014), Section 1.4, and Mateu et al. (2015). In the metric space $(\mathfrak{N}_{\text{fin}}, \tau)$, we can furthermore define a Fréchet mean of order $q \geq 1$; that is, for data $\xi_1, \ldots, \xi_k \in \mathfrak{N}_{\text{fin}}$ any $\zeta \in \mathfrak{N}_{\text{fin}}$ minimizing

Raoul Müller is supported by *Deutsche Forschungsgemeinschaft* GRK 2088. Jorge Mateu is supported by MTM2016-78917-R from the *Spanish Ministry of Economy and Competitiveness*.

✉ Raoul Müller
raoul.mueller@uni-goettingen.de

[1] Institute for Mathematical Stochastics, University of Göttingen, 37077 Göttingen, Germany

[2] Department of Mathematics, University Jaume I, 12071 Castellón, Spain

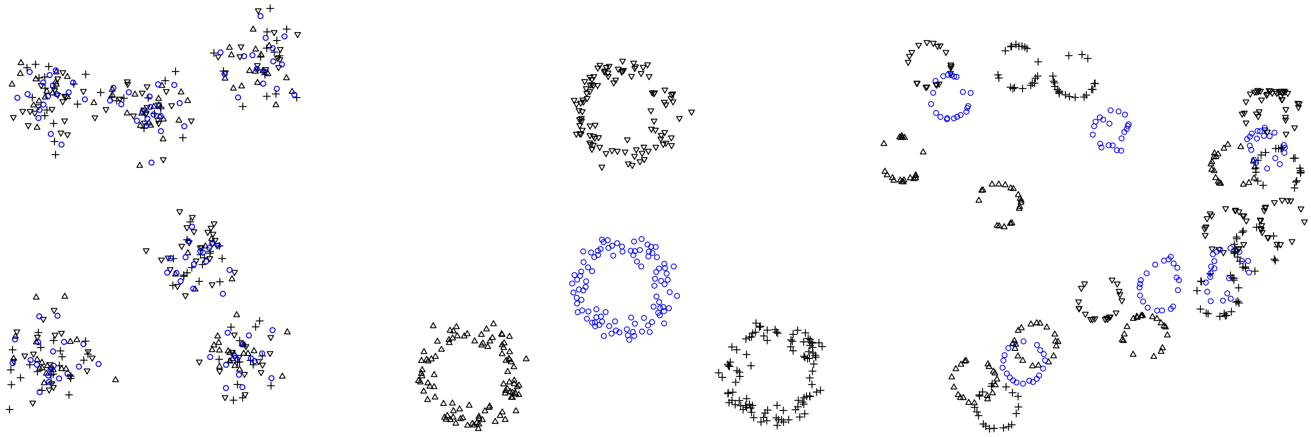


Fig. 1 An example of barycenters computed by our algorithm for three different data sets. In each panel, there are three data point patterns indicated by different symbols (black). The resulting (pseudo-)barycenter pattern with respect to Euclidean distance is given by the blue circles ($p = q = 2$). (Color figure online)

$$\sum_{j=1}^{k} \tau(\xi_j, \zeta)^q. \tag{1}$$

Such a q-th-order mean may serve as a "typical" element of $\mathfrak{N}_{\text{fin}}$ to represent the data and gives rise to more complex statistical analyses, such as Fréchet regression; see Lin and Müller (2019) and Petersen and Müller (2019).

Two metrics on the space of point patterns that have been widely used are the spike time metric, see Victor and Purpura (1997) for one dimension and Diez et al. (2012) for higher dimension, and the optimal subpattern assignment (OSPA) metric, see Schuhmacher and Xia (2008) and Schuhmacher et al. (2008). In the present paper, we introduce the *transport–transform (TT) metric* and its normalized version, the *relative transport–transform (RTT) metric*, which provide a unified framework for the earlier metrics. Both the TT- and the RTT-metrics are based on matching the points between two point patterns on $\mathcal{X}$ optimally in terms of some power p of d and penalizing points that cannot be reasonably matched. We may interpret these metrics as unbalanced p-th-order Wasserstein metrics, see Remark 3 below. In the present paper, we always set $p = q$.

Among others Schoenberg and Tranbarger (2008), Diez et al. (2012) and Mateu et al. (2015) have treated Fréchet means of order 1 (medians) for the spike time metric under the name of prototypes. However, computations in 2d and higher were only possible for very small data sets due to a prohibitive computational cost of $O(n^6)$ for the distance between two point patterns with n points each. In the present work, we use an adapted auction algorithm that is able to compute TT- and RTT-distances between point patterns in $O(n^3)$. We further provide a heuristic algorithm that bears some resemblance to a k-means cluster algorithm and is able to compute local minima of the barycenter problem very efficiently. This makes it possible to compute "quasi-barycenters" for 100 patterns of 100 points in $\mathbb{R}^2$ in a few seconds when basing the TT-distance on the Euclidean distance between points and choosing $p = q = 2$.

In Fig. 1, we show some typical barycenters obtained by our algorithm in this setting. We use smaller data sets for better visibility. In each scenario, there are three different point patterns distinguished by the different symbols in black. The (pseudo-)barycenter represented by the blue circles captures the characteristics of each data set rather well. Some minor irregularities, especially in the third panel, may be due to the fact that only a (good) local optimum is computed.

More important than being fast for point pattern data on $\mathbb{R}^D$ when using squared Euclidean distances is the fact that our algorithm provides a general plug-in method that can in principle be used for point patterns on *any* underlying space $\mathcal{X}$ where an appropriate "cost function" between objects is specified as p-th power of a metric d. All that is required is an algorithm that finds (maybe heuristically) a p-th-order Fréchet mean for individual points in $\mathcal{X}$, i.e., finds $z \in \mathcal{X}$ minimizing $\sum_{j=1}^{k} d(x_j, z)^p$ for any given $x_1, \dots, x_k \in \mathcal{X}$. We refer to this in what follows as the underlying *location problem*. The reduction to the underlying location problem allows us to treat the case of point patterns on a network equipped with the shortest-path metric and $p = 1$. Figure 2 gives an example for crime data in Valencia, Spain, which we study in more detail in Sect. 6.

The barycenter problem we consider in this paper is closely related to the problem of computing an unbalanced Wasserstein barycenter, see, e.g., Chizat et al. (2018). However, rather than minimizing a Fréchet functional on the space of all measures, we minimize on the space $\mathfrak{N}_{\text{fin}}$ of $\mathbb{Z}_+$-valued measures, see Remark 5.

The plan of the paper is as follows. In Sect. 2, we introduce the TT- and RTT-metrics and discuss their relations to spike time, OSPA, and incomplete Wasserstein metrics. Sec-

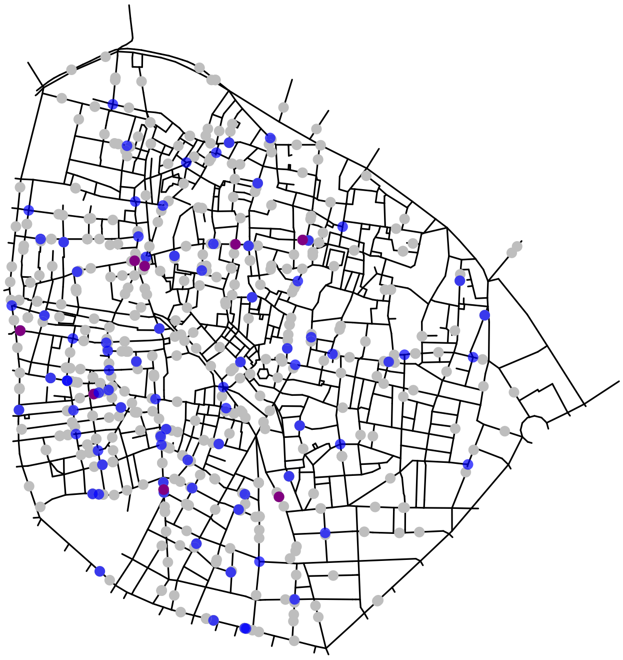


Fig. 2 An example of a barycenter on a street network. Shown are 8 patterns of assault crimes during the summer months of 2010–2017 in the old town of Valencia (all in gray for better overall visibility). The resulting barycenter with respect to shortest-path distance along the streets is given in blue, with multipoints in purple ($p = q = 1$). (Color figure online)

tion 3 specifies what we mean by a barycenter (or Fréchet mean) with respect to these metrics and gives an important result that forms the basis for our heuristic algorithm. Two versions of this algorithm, a more direct one and an improved one, which saves computation steps that are unlikely to substantially influence the final result, are discussed in detail in Sect. 4, along with some practical aspects. Section 5 contains a larger simulation study, which investigates robustness and runtime performances of the two algorithms for the case of Euclidean distance and $p = 2$. Finally, we give two applications to data of crime events on a city map for real data in Sect. 6. The first one concerns street thefts in Bogotá, Colombia. We treat this again as data in Euclidean space, using $p = 2$. The second one deals with assault cases in the streets of Valencia, Spain. Here, we compute barycenters based on the actual shortest-path distance on the street network and use $p = 1$.

2 The transport–transform metric

Denote by $\mathfrak{N}_{\text{fin}}$ the space of finite point patterns (counting measures) on a complete separable metric space $(\mathcal{X}, d)$, equipped with the usual σ-algebra $\mathcal{N}_{\text{fin}}$ generated by the point count maps $\Psi_A : \mathfrak{N}_{\text{fin}} \to \mathbb{R}$, $\xi \mapsto \xi(A)$ for $A \subset \mathcal{X}$ Borel measurable. Elements of $\mathfrak{N}_{\text{fin}}$ are typically denoted by ξ, η, ζ here. As usual, we write δ_x for the Dirac measure with unit mass at $x \in \mathcal{X}$. In the present section, we mostly use measure notation such as $\xi = \sum_{i=1}^{n} \delta_{x_i}$, $\xi(\{x\}) \geq 1$ or $\xi + \eta$, but in later sections we also use corresponding (multi)set notation such as $\xi = \{x_1, \ldots, x_n\}$, $x \in \xi$ or $\xi \cup \eta$ where this is unambiguous.

We use $|\xi| = \xi(\mathcal{X})$ to denote the total number of points in the pattern ξ. For $n \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ write $[n] = \{1, 2, \ldots, n\}$ (including $[0] = \emptyset$) and denote by $\mathfrak{N}_n$ the set of point patterns with exactly n points. We first introduce the metrics we use on $\mathfrak{N}_{\text{fin}}$, which unify and generalize two of the main metrics used previously in the literature.

Definition 1 Let $C > 0$ and $p \geq 1$ be two parameters, referred to as *penalty* and *order*, respectively.

(a) For $\xi = \sum_{i=1}^{m} \delta_{x_i}$, $\eta = \sum_{j=1}^{n} \delta_{y_j} \in \mathfrak{N}_{\text{fin}}$, define the *transport–transform (TT) metric* by

$$\tau(\xi, \eta) = \tau_{C,p}(\xi, \eta) = \left(\min \left((m + n - 2l) C^p + \sum_{r=1}^{l} d(x_{i_r}, y_{j_r})^p \right) \right)^{1/p}, \tag{2}$$

where the minimum is taken over equal numbers of pairwise different indices $i_1, \ldots, i_l$ in $[m]$ and $j_1, \ldots, j_l$ in $[n]$, i.e. over the set

$$\begin{aligned} S(m, n) = \big\{ &(i_1, \ldots, i_l;\ j_1, \ldots, j_l)\,; \\ &l \in \{0, 1, \ldots, \min\{m, n\}\}, \\ &i_1, \ldots, i_l \in [m] \text{ pairwise different}, \\ &j_1, \ldots, j_l \in [n] \text{ pairwise different} \big\}. \end{aligned}$$

(b) For $\xi, \eta \in \mathfrak{N}_{\text{fin}}$, define the *relative transport–transform (RTT) metric* by

$$\bar{\tau}(\xi, \eta) = \bar{\tau}_{C,p}(\xi, \eta) = \frac{1}{\max\{|\xi|, |\eta|\}^{1/p}} \tau_{C,p}(\xi, \eta). \tag{3}$$

We state and prove below that τ and $\bar{\tau}$ are indeed metrics.

The following result simplifies proofs of statements about these metrics and is furthermore invaluable for their computation. The idea is to extend the metric space $(\mathcal{X}, d \wedge (2^{1/p} C))$, where $[d \wedge (2^{1/p} C)](x, y) = \min\{d(x, y), 2^{1/p} C\}$, by setting $\mathcal{X}' = \mathcal{X} \cup \{\aleph\}$ for an auxiliary element $\aleph \notin \mathcal{X}$ and

$$d'(x, y) = \begin{cases} \min\{d(x, y), 2^{1/p} C\} & \text{if } x, y \in \mathcal{X}; \\ C & \text{if } \aleph \in \{x, y\},\ x \neq y; \\ 0 & \text{if } x = y = \aleph. \end{cases}$$

It is shown in Lemma A.1 that $(\mathcal{X}', d')$ is a metric space again. We may then compute distances in the τ and $\bar{\tau}$ metrics by solving an optimal matching problem between point patterns with the same cardinality. For $n \in \mathbb{N}$ denote by S_n, the set of permutations on $[n]$.

Theorem 1 *Let $\xi = \sum_{i=1}^{m} \delta_{x_i}$, $\eta = \sum_{j=1}^{n} \delta_{y_j} \in \mathfrak{N}_{\text{fin}}$, where w.l.o.g. $m \le n$ (otherwise swap ξ and η). Set $x_i = \aleph$ for $m+1 \le i \le n$ and $\tilde{\xi} = \sum_{i=1}^{n} \delta_{x_i}$. Then,*

$$\tau(\xi, \eta) = \left(\min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \right)^{1/p} \quad \textit{and}$$

$$\bar{\tau}(\xi, \eta) = \left(\frac{1}{n} \min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \right)^{1/p}.$$

The proof of this and the other theorems in this section can be found in the appendix.

Remark 1 (Computation of TT- and RTT-metrics) Writing n for the maximum cardinality as in Theorem 1, this result shows that we can compute both $\tau(\xi, \eta)$ and $\bar{\tau}(\xi, \eta)$ in worst-time complexity of $O(n^3)$ by using the classic Hungarian method for the assignment problem; see Kuhn (1955). In practice, we use the auction algorithm proposed in Bertsekas (1988), because it has usually much better runtime in our experience, although the default version has a somewhat worse worst-case performance of $O(n^3 \log(n))$.[1]

Theorem 2 *The maps τ and $\bar{\tau}$ are metrics on $\mathfrak{N}_{\text{fin}}$.*

The next result establishes our previous claim that the new transport–transform construction generalizes two metrics on $\mathfrak{N}_{\text{fin}}$ previously used in the literature.

Theorem 3

(a) If $p = 1$, then for any $\xi, \eta \in \mathfrak{N}_{\text{fin}}$

$$\tau(\xi, \eta) = \min_{(\xi_0, \ldots, \xi_N)} \sum_{i=0}^{N-1} c_{elem}(\xi_i, \xi_{i+1}), \tag{4}$$

where the minimum is taken over all $N \in \mathbb{N}$ and all paths $(\xi_0, \ldots, \xi_N) \in \mathfrak{N}_{\text{fin}}^{N+1}$ such that $\xi_0 = \xi$, $\xi_N = \eta$, and from ξ_i to ξ_{i+1} either a single point is added or deleted at cost $c_{elem}(\xi_i, \xi_{i+1}) = C$ or a single point is moved from x to y at cost $c_{elem}(\xi_i, \xi_{i+1}) = d(x, y)$.

(b) If $\text{diam}(\mathcal{X}) = \sup_{x,y \in \mathcal{X}} d(x, y) \le 2^{1/p} C$, then for any $\xi = \sum_{i=1}^{m} \delta_{x_i}$, $\eta = \sum_{j=1}^{n} \delta_{y_j} \in \mathfrak{N}_{\text{fin}}$, assuming w.l.o.g. $m \le n$

$$\bar{\tau}(\xi, \eta)^p = \frac{1}{n} \left((n - m) C^p + \min_{\pi \in S_n} \sum_{i=1}^{m} d(x_i, y_{\pi(i)})^p \right). \tag{5}$$

Theorem 3(a) implies that the TT-metric is the same as the spike time metric (using add and delete penalties $P_a = P_d = C$ and a move penalty $P_m = 1$), which was originally introduced on $\mathbb{R}_+$ by Victor and Purpura (1997) and generalized to metric spaces by Diez et al. (2012). It can be seen from the proof in the appendix that the right hand side of (4) is not a metric in general if $p > 1$.

Theorem 3(b) implies that the RTT-metric is the same as the OSPA metric, introduced in Schuhmacher and Xia (2008) and Schuhmacher et al. (2008). Note that in the definition of the OSPA metric $\text{diam}(\mathcal{X}) \le C \le 2^{1/p} C$ was either required or enforced by taking the minimum of d with C. Here, it can be seen that the right hand side of (5) is not a metric in general if $\text{diam}(\mathcal{X}) > 2C$.

Remark 2 (Computation of spike time distances) The spike time distances in Victor and Purpura (1997) and Diez et al. (2012) allowed for separate add and delete penalties P_a and P_d, as well as a move penalty P_m (factor in front of $d(x, y)$). We set here $P_a = P_d = C$ to obtain a proper metric and divide distances by P_m, which is just a scaling. Thus, the parameter $C = P_a/P_m = P_d/P_m$ is all that remains.
As noted at the end of Section 4 in Diez et al. (2012), having different add and delete penalties may be useful for controlling the total number of points in a barycenter point pattern. Let us point out therefore that Theorem 1 is easily adapted to this more general situation by setting $d'(x, y) = \min\{d(x, y), 2^{1/p}(P_a + P_d)\}$, $d'(\aleph, y) = P_a$ and $d'(x, \aleph) = P_d$ for all $x, y \in \mathcal{X}$.
In particular, this yields a worst-time complexity of $O(n^3)$ for general (maybe asymmetric) spike time distances in general metric spaces, which is a substantial improvement over the $O(n^6)$ complexity of the incremental matching algorithm presented in Diez et al. (2012).

Remark 3 (Unbalanced Wasserstein metrics) The TT- and RTT-metrics can be seen as unbalanced Wasserstein metrics, see, e.g., Chizat et al. (2018), Liero et al. (2018) and the references therein. Minimizing over the space $\mathfrak{M}_{\text{fin}}$ of all finite measures on $\mathcal{X} \times \mathcal{X}$, we obtain the TT-distance as a solution to a particular instance of the unbalanced optimal transport problem in Chizat et al. (2018), Definition 2.11, namely

$$\tau(\xi, \eta)^p = \inf_{\gamma \in \mathfrak{M}_{\text{fin}}} \biggl(\int_{\mathcal{X} \times \mathcal{X}} d(x, y)^p \, \gamma(dx, dy) \\ + C^p \|\xi - \gamma_1\|_{\text{TV}} + C^p \|\eta - \gamma_2\|_{\text{TV}} \biggr), \tag{6}$$

[1] There is a modified auction algorithm that can improve the worst-case performance to $O(n^{5/2} \log(n))$; for the performance discussion see Bertsekas (1988), page 109. Actually, both orders include a factor c in the log which measures the numerical precision, assumed to be bounded here.

where $\gamma_1 = \gamma(\cdot \times \mathcal{X})$ and $\gamma_2 = \gamma(\mathcal{X} \times \cdot)$ denote the marginals of γ, and $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm of signed measures; specifically $\|\mu-\nu\|_{\mathrm{TV}} = \sup_A(\mu(A)-\nu(A))+\sup_A(\nu(A)-\mu(A))$ for $\mu, \nu \in \mathfrak{M}_{\mathrm{fin}}$, where the suprema are taken over all measurable subsets of $\mathcal{X}$.
Equation (6) can be shown as follows. It is straightforward to see that we may take the infimum on the right hand side only over $\gamma \in \mathfrak{M}_{\mathrm{fin}}$ with marginals $\gamma_1 \leq \xi$ and $\gamma_2 \leq \eta$, because any additional mass in γ may be removed without increasing the total cost of γ. Writing $\xi = \sum_{i=1}^n \delta_{x_i}$ and $\eta = \sum_{i=1}^n \delta_{y_i}$ with the help of additional points at $\aleph$ (if necessary), we obtain by similar arguments as in the proof of Theorem 1 that the latter problem is equivalent to the discrete transportation problem

$$\begin{aligned}&\min_{(\gamma_{ij})_{1\leq i,j\leq n}} \sum_{i,j=1}^n d'(x_i,y_j)^p \cdot \gamma_{ij}\\ &\qquad \text{s.t.} \sum_{j=1}^n \gamma_{ij} = 1 \text{ for all } i, \ \sum_{i=1}^n \gamma_{ij} = 1 \text{ for all } j,\\ &\qquad \gamma_{ij} \geq 0 \text{ for all } i,j.\end{aligned}$$

It is a standard result in linear programming that this problem always has a solution $\gamma_{ij} \in \{0,1\}$, $1 \leq i,j \leq n$; see, e.g., the theorem in Section 6.5 of Luenberger and Ye (2008), which is essentially due to the fact that the structure of the constraint allows for a back substitution approach involving only additions and subtractions. We may therefore conclude from Theorem 1 that Equation (6) holds and that the infimum on the right hand side is attained for $\gamma = \sum_{i,j=1}^n \mathbb{1}\{x_i, y_j \neq \aleph\}\gamma_{ij}\delta_{(x_i,y_j)}$.

In principle, Remark 3 allows us to specialize results and algorithms for unbalanced Wasserstein metrics to TT- and RTT-metrics. However, the discrete setting we consider here is sometimes not included in the general theorems or requires a more specialized treatment. Algorithms for computing unbalanced transport plans are typically derived from balanced optimal transport algorithms; a selection can be found in Chizat (2017). The auction algorithm we use in this paper is derived from the auction algorithm used for balanced assignment problems in a similar way.

3 Barycenters with respect to the TT-metric

For data on quite general metric spaces, barycenters can formalize the idea of a center element representing the data. In the case of $\mathfrak{N}_{\mathrm{fin}}$, we are thus looking for a center point pattern that gives a good first-order representation of a set of data point patterns $\xi_1, \ldots, \xi_k$. More formally, we may define a barycenter as the (weighted) q-th-order Fréchet mean with respect to τ; see Fréchet (1948).

Definition 2 For $k \in \mathbb{N}$, let $\xi_1, \ldots, \xi_k \in \mathfrak{N}_{\mathrm{fin}}$ be data point patterns and $\lambda_1, \ldots, \lambda_k > 0$ with $\sum_{j=1}^k \lambda_j = 1$ be weights. Let furthermore $q \geq 1$. Then, we call any

$$\zeta_* \in \arg\min_{\zeta \in \mathfrak{N}_{\mathrm{fin}}} \sum_{j=1}^k \lambda_j \tau(\xi_j, \zeta)^q \tag{7}$$

a *(weighted) barycenter of order* q. If no weights are specified, we tacitly assume that $\lambda_j = 1/k$ for $1 \leq j \leq k$, leading to an "unweighted" barycenter.

Remark 4 For $q = 2$, barycenters on general metric spaces are simply known as (empirical) *Fréchet means*. For $q = 1$, they are sometimes known as *Fréchet medians*. This comes from the fact that given $x_1, \ldots, x_k \in \mathbb{R}^D$, we have

$$\arg\min_{z \in \mathbb{R}^D} \sum_{j=1}^k \|x_j - z\|^2 = \frac{1}{k}\sum_{j=1}^k x_j \tag{8}$$

(the arg min is unique here), and that given $x_1, \ldots, x_k \in \mathbb{R}$, we have

$$\arg\min_{z \in \mathbb{R}} \sum_{j=1}^k \|x_k - z\| = \mathrm{median}\{x_1, \ldots, x_k\}, \tag{9}$$

where the right hand side denotes the set of medians $\big\{z \in \mathbb{R};\ \#\{j;\ x_j \leq z\} = \#\{j;\ x_j \geq z\}\big\}$.

Remark 5 As seen in Remark 3, we may interpret τ as an unbalanced Wasserstein metric. There has been a great deal of research on Wasserstein barycenters (in the Fréchet mean sense as above, see, e.g., Agueh and Carlier (2011) or Cuturi and Doucet (2014)), which more recently also extends to unbalanced Wasserstein metrics, see, e.g., Chizat et al. (2018) or Schmitz et al. (2018). In addition to the fact that much of the corresponding theory is not well adapted to the case of discrete input measures, with the notable exception of Anderes et al. (2016), we point out that a fundamental difference of (7) lies in the fact that we minimize over the space $\mathfrak{N}_{\mathrm{fin}}$ of $\mathbb{Z}_+$-valued measures. This space is smaller than the space $\mathfrak{M}_{\mathrm{fin}}$ of general finite measures, but has a more complicated structure because it decays into connected components $\mathfrak{N}_n = \{\xi \in \mathfrak{N}_{\mathrm{fin}};\ |\xi| = n\}$ (under the TT-metric), implying, e.g., that continuous optimization procedures will not work directly.

In what follows, we always set $p = q$ and choose this number mostly $\in \{1, 2\}$. We refer to the resulting barycenters simply as 1- and 2-*barycenter* or as *point pattern median* and *point pattern mean*, respectively. Point pattern medians have been introduced under the name of *prototypes* in Schoenberg and Tranbarger (2008) on $\mathbb{R}$ and studied in higher dimensions in Diez et al. (2012) and Mateu et al. (2015). However, in these papers the applicability was limited to rather small data sets due to the large computation cost of $O(n^6)$ mentioned in Remark 2.

Using the construction from Theorem 1, we may reformulate the barycenter problem as a multidimensional assignment problem, generalizing Lemma 16 in Koliander et al. (2018). Note that for the TT-metric we can add an arbitrary number of points at $\aleph$ to both point patterns without changing the minimum in Theorem 1.

Theorem 4 *For point patterns $\xi_j = \sum_{i=1}^{n_j} \delta_{x_{ij}}$, $j \in [k]$, let $\tilde{n} := \lfloor \frac{2}{k+1} \sum_{j=1}^{k} n_j \rfloor$ and $n \geq \max\{\tilde{n}, n_j;\ 1 \leq j \leq k\}$. Set $x_{ij} = \aleph$ for $n_j + 1 \leq i \leq n$ and $\tilde{\xi}_j = \sum_{i=1}^{n} \delta_{x_{ij}}$ for any $j \in [k]$.*

Then, for any $\pi_{*,1}, \ldots, \pi_{*,k} \in S_n$ jointly minimizing

$$\sum_{i=1}^{n} \min_{z \in \mathcal{X}'} \sum_{j=1}^{k} d'(x_{\pi_j(i),j}, z)^p \tag{10}$$

the point pattern $\zeta_*|_{\mathcal{X}}$ with $\zeta_* = \sum_{i=1}^{n} \delta_{z_i}$, where $z_i \in \arg\min_{z \in \mathcal{X}'} \sum_{j=1}^{k} d'(x_{\pi_{*,j}(i),j}, z)^p$ is a p-th-order barycenter with respect to the TT-metric.

The $\pi_{*,1}, \ldots, \pi_{*,k} \in S_n$ above define n disjoint “clusters” $\mathcal{C}_i = \{x_{\pi_{*,j}(i),j};\ 1 \leq j \leq k\}$, where each contains exactly one (maybe virtual) point of each point pattern. The minimization of (10) may thus be interpreted as a multidimensional assignment problem with cluster cost

$$\text{cost}_*(\mathcal{C}) = \min_{z \in \mathcal{X}'} \sum_{x \in \mathcal{C}} d'(x, z)^p. \tag{11}$$

Proof Let us first give an upper bound on the cardinality of the barycenter. A single barycenter point can be matched with up to k points (one from each point pattern). If said point is matched with only $\frac{k}{2}$ points or fewer, it cannot be worse to delete it. The contribution for this point in the objective function is at least $\frac{k}{2}C$, while deleting it adds at most $\frac{k}{2}C$ to the objective function.

So, every barycenter point should be matched with at least $\lceil \frac{k+1}{2} \rceil$ points. The total number of points is $\sum_{j=1}^{k} n_j$. Therefore, the number of barycenter points is bounded above by $\tilde{n} = \lfloor \frac{2}{k+1} \sum_{j=1}^{k} n_j \rfloor$.

It is thus sufficient to fill up all the point patterns ξ_j to n points and work also with an ansatz of n points for ζ. Theorem 1 yields

$$\begin{aligned} &\min_{\zeta \in \mathfrak{N}_{\text{fin}}} \sum_{j=1}^{k} \tau(\xi_j, \zeta)^p \\ &\quad = \min_{z_1,\ldots,z_n \in \mathcal{X}'} \sum_{j=1}^{k} \min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_{\pi(i),j}, z_i)^p \\ &\quad = \min_{z_1,\ldots,z_n \in \mathcal{X}'} \min_{\pi_1,\ldots,\pi_k \in S_n} \sum_{j=1}^{k} \sum_{i=1}^{n} d'(x_{\pi_j(i),j}, z_i)^p \\ &\quad = \min_{\pi_1,\ldots,\pi_k \in S_n} \sum_{i=1}^{n} \min_{z_i \in \mathcal{X}'} \sum_{j=1}^{k} d'(x_{\pi_j(i),j}, z_i)^p \end{aligned} \tag{12}$$

and that any minimizer $\zeta_*|_{\mathcal{X}} = \sum_{i=1}^{n} \delta_{z_i}|_{\mathcal{X}}$ on the left hand side is obtained from jointly minimizing in $\pi_1, \ldots, \pi_k$ and $z_1, \ldots, z_n$ on the right hand side. □

4 Alternating clustering algorithms

Based on Theorem 4, we propose an algorithm that alternates between minimizing

$$\sum_{j=1}^{k} \sum_{i=1}^{n} d'(x_{\pi_j(i),j}, z_i)^p \tag{13}$$

in $\pi_1, \ldots, \pi_k \in S_n$ and in $z_1, \ldots, z_n \in \mathcal{X}'$ until convergence. Such an algorithm terminates in a local minimum of (13) after a finite number of steps, because (13) can never increase and the minimization in the permutations is over a finite space.

Since this underlying idea is close to the popular k-means clustering algorithm, we named the main function in the pseudocode and in the actual implementation `kMeansBary` (note, however, that n plays the role of k in our notation). Similar alternating algorithms in the context of Wasserstein-2 barycenters for finitely supported probability measures have been proposed in Cuturi and Doucet (2014), Borgwardt (2019) and del Barrio et al. (2019). See Sect. 5, where we compare results between `kMeansBary` and Algorithm 2 in Cuturi–Doucet.

In what follows, we present pseudocode along with the underlying ideas and explanations for two versions of the `kMeansBary`-algorithm that we dub *original* and *improved*. Here, “improved” refers to the fact that we cut down on certain computation steps in order to save runtime. We will see in Sect. 5 that this comes essentially without any performance loss.

User-friendly implementations of both algorithms are publicly available in the R-package ttbary; see Müller and Schuhmacher (2019).

4.1 Our original `kMeansBary` algorithm

The pseudocode for the basic alternating strategy described above is given in Algorithm 1. We have introduced a stopping parameter δ to allow termination before the local optimum is reached. Since we are not interested in the actual clustering, but only in the position of the centers $z_1, \ldots, z_n$, it seems very unlikely (though possible) that the solution changes substantially once the cost decrease has become very small. What is more, such a change might be spurious due to rounding errors in the data or when we use an approximation method for optimizing in the centers. Note also that we can always set δ to the smallest representable positive floating-point number to ensure convergence to the local optimum.

Algorithm 1: `kMeansBary`. Dependence on data pplist suppressed for simplicity.

```
Input  : center an initial pseudo-barycenter;
         pplist the list of data point patterns;
         $\delta > 0$ a constant for the termination;
         $N$ the maximum number of iterations.
Output : Locally optimal pseudo-barycenter center.
perm, cost ← optimPerm(center);
for it ← 1 to $N$ do
    costold ← cost;
    center ← optimBary(perm, center);
    center ← optimDelete(perm, center);
    center ← optimAdd(perm, center);
    perm, cost ← optimPerm(center);
    if costold − cost < $\delta$ then break;    // difference always nonnegative
end
return center;  // warn if the loop has run out
```

The minimization with respect to $\pi_1, \ldots, \pi_k$ is performed by `optimPerm`. This function computes an optimal matching between the current center and each data point pattern in `pplist`, using an alternating version of the auction algorithm with ε-scaling; see Remark 1 and Bertsekas (1988) for more details. We output the `cost` of the current matching and an $n \times k$ matrix `perm`, whose j-th column specifies the order in which the points of the j-th data pattern are matched to $z_1, \ldots, z_n$. For greater efficiency, we save auxiliary information (price and profit vectors) and use it for initializing the auction algorithm when calling it again with the same data point pattern.

For practical purposes, we have split up the minimization with respect to $z_1, \ldots, z_n \in \mathcal{X}'$ into a function `optimBary` that optimizes the positions within $\mathcal{X}$ and functions `optimDelete` and `optimAdd` that optimize which of the z_i to move from $\mathcal{X}$ to $\aleph$ and from $\aleph$ to $\mathcal{X}$, respectively. We discuss details of these functions under the separate headings below.

In addition to the outputs of the various functions shown in Algorithm 1, we also keep information on the quality of each match of points up to date. We call the match of a z_i with a data point $x_{i'j}$

$$\begin{aligned} \textit{happy} \text{ if } & z_i, x_{i'j} \in \mathcal{X} \text{ and } d'(z_i, x_{i'j}) < 2^{1/p}C \\ \textit{miserable} \text{ if } & z_i, x_{i'j} \in \mathcal{X} \text{ and } d'(z_i, x_{i'}) = 2^{1/p}C \\ & \text{or if } z_i = \aleph, x_{i'j} \in \mathcal{X} \\ \textit{to } \aleph \text{ if } & x_{i'j} = \aleph. \end{aligned}$$

Note that a miserable match is worst possible in the sense that $\text{cost}(\mathcal{C}_i) = \sum_{x \in \mathcal{C}_i} d'(x, z_i)^p$ for center z_i cannot increase if $x_{i'j}$ is replaced by *any* other $x \in \mathcal{X}'$.

Details on `optimBary`

The purpose of this function is to find for each $z_i \in \mathcal{X}$ (i.e., not currently at $\aleph$) a location in $\mathcal{X}$ that minimizes $\text{cost}(\mathcal{C}_i)$ for its current cluster $\mathcal{C}_i = \{x_{\pi_j(i),j};\ 1 \leq j \leq k\}$. This amounts to a more traditional location problem in $\mathcal{X}$, except that it is typically made (much) more difficult by the fact that we have to truncate distances at $2^{1/p}C$.

Note that any cluster points at $\aleph$ can be ignored because they always contribute the same amount to the cluster cost, no matter where the center lies. The same is true for individual points that have a much larger distance than $2^{1/p}C$ from the bulk of the points. However, there are countless scenarios with (groups of) points being around distance $2^{1/p}C$ apart from one another for which the optimization of the cluster cost becomes a difficult optimization problem (piecewise smooth on a space that is fragmented in complicated ways).

As a simple heuristic that works well in cases where we do not have to cut too many distances (i.e., C is not too small), we suggest to ignore all points that are at the maximal d'-distance $2^{1/p}C$ from the current z_i when computing the new z_i. Note that in this way the cluster cost can never increase.

Algorithm 2: `optimBary`: find optimal center pattern *within* $\mathcal{X}$ for given clusters $\mathcal{C}_i$.

```
for $i$ ← 1 to $n$ do
    if $z_i \in \mathcal{X}$ then    // $z_i$ is the current center of the $i$th cluster $\mathcal{C}_i$
        happypoints ←
          {points in $\mathcal{C}_i$ that are happily matched to $z_i$};
        if happypoints != ∅ then
            $z_i$ ← optimClusterCenter(happypoints);
        end    // otherwise $z_i$ is deleted in next call to optimDelete
    end
end
return $\{z_1, \ldots, z_n\}$;
```

Algorithm 2 gives corresponding pseudocode. The function `optimClusterCenter` handles the location problem for the untruncated metric d on $\mathcal{X}$. If for example $\mathcal{X} = \mathbb{R}^D$ equipped with the Euclidean metric and $p = 2$, Equation (8) implies that `optimClusterCenter` simply has to take the (coordinatewise) average of all happy points. The case $p = 1$ can be tackled with higher computational effort by approximation via the popular Weiszfeld algorithm; see Weiszfeld (1937).

As a further instance, which we will take up in Sect. 6, we consider the situation where $\mathcal{X}$ is a simple graph (V, E) equipped with the shortest-path distance and $p = 1$. It can be shown that in this case the location problem in $\mathcal{X}$ is solved by an element z_i of $V \cup \mathcal{C}_i$, i.e., either a vertex of the graph or any data point, see Hakimi (1964). We therefore proceed by first computing the distance matrix between all these points, which is then used for the entire algorithm. Such shortest-path distance computations in sparse graphs with thousands of points can be performed in (at most) a few seconds by various algorithms, see Chapter 25 in Cormen et al. (2009) and the concrete timing in Sect. 6.2. It is now easy to implement the function `optimClusterCenter`. For a given set of happy points of a cluster $\mathcal{C}_i$, pick the corresponding columns in the distance matrix, add them up and determine the minimal entry of the resulting vector. If there are several such entries, which due to choosing $p = 1$ can happen quite frequently, we pick one among them uniformly at random. The index of the obtained entry identifies the center point z_i.

Precomputing the distance matrix between all points of $V \cup \mathcal{C}_i$ in the graph case has the additional advantage that no distances have to be computed in the `optimPerm` step. It is, on the other hand, the main bottleneck of the procedure and may not be feasible in situations with very large graphs and data sets. In this case, we can resort to one of the various heuristics available, such as the single and multi-hub heuristics proposed (in principle) in Bandelt et al. (1994) and Koliander et al. (2018).

Details on `optimDelete`

This function deletes (i.e., moves to $\aleph$) any $z_i \in \mathcal{X}$ for which this operation decreases $\mathrm{cost}(\mathcal{C}_i)$.

We denote by k_{happy}, k_{miser} and $k_{\aleph}$ the numbers of data points in $\mathcal{C}_i$ that are happy, miserable and at $\aleph$, respectively. Write furthermore c_{happy} for the total cost of matching the happy points to z_i. If z_i stays in $\mathcal{X}$, the cluster incurs an overall total cost of

$$c_{\text{happy}} + k_{\text{miser}} \cdot 2C^p + k_{\aleph} \cdot C^p$$

as opposed to

$$k_{\text{happy}} \cdot C^p + k_{\text{miser}} \cdot C^p$$

if we delete z_i. Subtracting $k_{\text{miser}} \cdot C^p$ from both expressions, this leads to the deletion condition

$$k_{\text{happy}} C^p < c_{\text{happy}} + (k - k_{\text{happy}}) C^p.$$

Since $c_{\text{happy}} \geq 0$, a sufficient condition for deletion is $2k_{\text{happy}} < k$. We use this as a quick pretest, which allows us to avoid computing c_{happy} sometimes. The full deletion procedure is presented in Algorithm 3.

Algorithm 3: `optimDelete`: move center points from $\mathcal{X}$ to $\aleph$ if it decreases cost.

```
for $i \leftarrow 1$ to $n$ do
    if $z_i \in \mathfrak{N}$ then
        happypoints $\leftarrow$
          {points in $\mathcal{C}_i$ that are happily matched to $z_i$};
        $k_{\text{happy}} \leftarrow$ #happypoints;
        if $2 * k_{\text{happy}} < k$ then
            $z_i \leftarrow \aleph$;            // shortcut deletion
        else
            $c_{\text{happy}} \leftarrow \sum_{x \in \text{happypoints}} d'(x, z_i)^p$;
            if $k_{\text{happy}} * C^p < c_{\text{happy}} + (k - k_{\text{happy}}) * C^p$ then
              $z_i \leftarrow \aleph$;
        end
    end
end
return $\{z_1, \ldots, z_n\}$;
```

Details on `optimAdd`

This function adds (i.e., moves to $\mathcal{X}$) any $z_i \in \aleph$ for which it finds a way to do so that decreases $\mathrm{cost}(\mathcal{C}_i)$. Pseudocode is given in Algorithm 4.

As a compromise between computational simplicity and finding a good location in $\mathcal{X}$, we first sample a proposal location $\tilde{z}$ uniformly from all miserable data points (i.e., points from *any* cluster that are currently in a miserable match with their center). Before we consider moving z_i to $\tilde{z}$, we rebuild the cluster $\mathcal{C}_i$ in such a way that this move has a better chance of being accepted.

The corresponding procedure is performed by the `optimizeCluster`-function in the pseudocode: For each data pattern, ξ_j pick the miserable point that is closest to $\tilde{z}$ (if it has any) and exchange it with the corresponding point $x_{\pi_j(i),j}$ that is currently in $\mathcal{C}_i$. Since the point coming from the other cluster was miserable before, the cost of that cluster cannot increase by this exchange. The cost of the cluster $\mathcal{C}_i$ can increase only if it loses a point located at $\aleph$ in the exchange. In this case, the cost increases by C^p, which is compensated by the fact that the cost of the other cluster must decrease, either from $2C^p$ to C^p if its center is in $\mathcal{X}$, or from C^p to 0 if its center is at $\aleph$. Thus, the total cost remains the same, but C_i has an additional point in $\mathcal{X}$ now, which makes the successful addition of z_i to $\mathcal{X}$ more likely.

To further decrease the prospective cluster cost after addition, we update the proposal $\tilde{z}$ by recentering it in its new cluster using the appropriate `optimClusterCenter`-function introduced in `optimBary` (applied to the set of points of the new cluster that are in a happy match with $\tilde{z}$).

Finally, check whether the cost of the new cluster based on the updated $\tilde{z}$ is smaller than the same cost based on $z_i = \aleph$, which is C^p times the number $k_{\mathcal{X}}$ of non-$\aleph$ points in the new cluster. Set z_i to $\tilde{z}$ if this is the case.

Algorithm 4: `optimAdd`: move center points from $\aleph$ to $\mathcal{X}$ if it decreases cost.

```
alephindex ← $\{i \in [n];\ z_i = \aleph\}$;
if alephindex! = $\emptyset$ then
    supply ← {data points miserably matched to some $z_i$};
    for $i$ in alephindex do
        if supply = $\emptyset$ then break;
        $\tilde{z}$ ← sample(supply, 1);  // draw uniformly
          at random from supply
        supply ← supply \ $\{\tilde{z}\}$;
        newcluster, newperm ←
          optimizeCluster($\tilde{z}$, perm, $i$);
        newhappypoints ←
          {pts in newcluster with happy new match to $\tilde{z}$};
        if newhappypoints ! = $\emptyset$ then
          $\tilde{z}$ ← optimClusterCenter(newhappypoints);
        $k_{\mathcal{X}} \leftarrow \#\{x \in$ newcluster; $x \in \mathcal{X}\}$;
        $c_{\text{new}} \leftarrow \sum_{x \in \text{newcluster}} d'(x, \tilde{z})^p$;
        if $c_{\text{new}} < k_{\mathcal{X}} * C^p$ then
            $z_i \leftarrow \tilde{z}$;
            perm ← newperm;
            supply ← supply \ $\{x \in$ supply; $x$ is happy$\}$;
        end
    end
end
return $\{z_1, \dots, z_n\}$;
```

4.2 An improved `kMeansBary` algorithm

For obtaining an algorithm with a reduced computational cost, we cut down on steps that are costly, but are not expected to influence the resulting local optimum in a decisive way. Since for now we treat the location problem at the cluster level (performed by `optimBary`) as very general, allowing a wide range of metric spaces $(\mathcal{X}, d)$, we focus here on saving computations in the functions `optimPerm`, `optimDelete`, and `optimAdd`.

We have realized that by far the most additions and deletions of points take place in the first two iterations of the original algorithm (see also Fig. 3). Especially checking for addition of points is costly and after the first few iterations very rarely successful. Therefore, we limit such checking henceforward to the first $N_{\text{del/add}} = 5$ iteration steps. Some further heuristics could be applied in `optimAdd`, but the gain in computation time is not so large and they can significantly change the outcome, which is why we decided against implementing them.

In `optimPerm`, we cannot avoid doing matchings. However, the auction algorithm we use allows to solve a relaxation of the problem by stopping the ε-scaling method early. In general, the auction algorithm with ε-scaling based on a decreasing sequence $(\varepsilon_1, \dots, \varepsilon_l)$ returns successively improved solutions that are guaranteed to lie within $n\,\varepsilon_i$ of the optimal total cost after the i-th step, see Bertsekas (1988, Proposition 1). By representing rescaled distances as integers in $\{0, 1, \dots, 10^9\}$, an optimal matching is obtained in the l-th step if $\varepsilon_l < 1/n$. Our improved algorithm is based on the same ε-vector as the original algorithm, which has components $\varepsilon_i = \frac{1}{n+1}10^{l-i}$, $1 \le i \le l$, where l is chosen in such a way that $10^7 \le \varepsilon_1 < 10^8$. As a first improvement, we use the subsequence $(\varepsilon_{a_{\mathtt{it}}}, \varepsilon_{a_{\mathtt{it}}+1}, \dots, \varepsilon_{b_{\mathtt{it}}})$, where a and b are prespecified vectors of indices $\in \{1, 2, \dots, l\}$. A simple choice for a and b that tends to decrease the runtime noticeably is $a_{\mathtt{it}} = 1$ and $b_{\mathtt{it}} = \min\{\mathtt{it}, l\}$. Pseudocode for this is presented in Algorithm 5.

In practice, we settled for a somewhat more sophisticated improvement. We choose the vectors $a = (1, 1, 1, 3, 3, 3, \dots, 3, 4)$ and $b = (1, 2, 3, 4, 6, 8, \dots, 2\lfloor\frac{l-1}{2}\rfloor, l)$, and we use the sequence $(\varepsilon_{a_j}, \varepsilon_{a_j+1}, \dots, \varepsilon_{b_j})$, where $j = \mathtt{it}$ for $\mathtt{it} \in \{1, 2, 3\}$, and then j is increased by 1 each time the algorithm would otherwise converge or if the cost increases (which can only happen as long as the matchings are not optimal).

This strategy was chosen after analyzing the calculations of the algorithm with respect to the time each calculation takes. In the first two to three iterations, there are a lot of changes in the positions of the barycenter points. Especially in the first iteration, many points are deleted and added, which completely changes the assignments. Therefore, we have to begin the assignment calculation with ε_1 and to get more sensible results we get more precise with each of the first three iterations. After three iterations, there are usually no big changes to the barycenter anymore, so we can reuse the assignment from the iteration before as a sensible starting solution and can omit ε_1 and ε_2 in return. Leaving out the first entries of ε too soon increases the runtime. Every time the algorithm converges, but has no guaranteed optimal assignment (i.e., $b_j < l$), j is increased by 1, meaning that the next two entries of ε are used too, until the end of ε is reached. Then, we can safely leave out the first three entries of ε without increasing the runtime, because at this point the assignments from one iteration to the next only change very little.

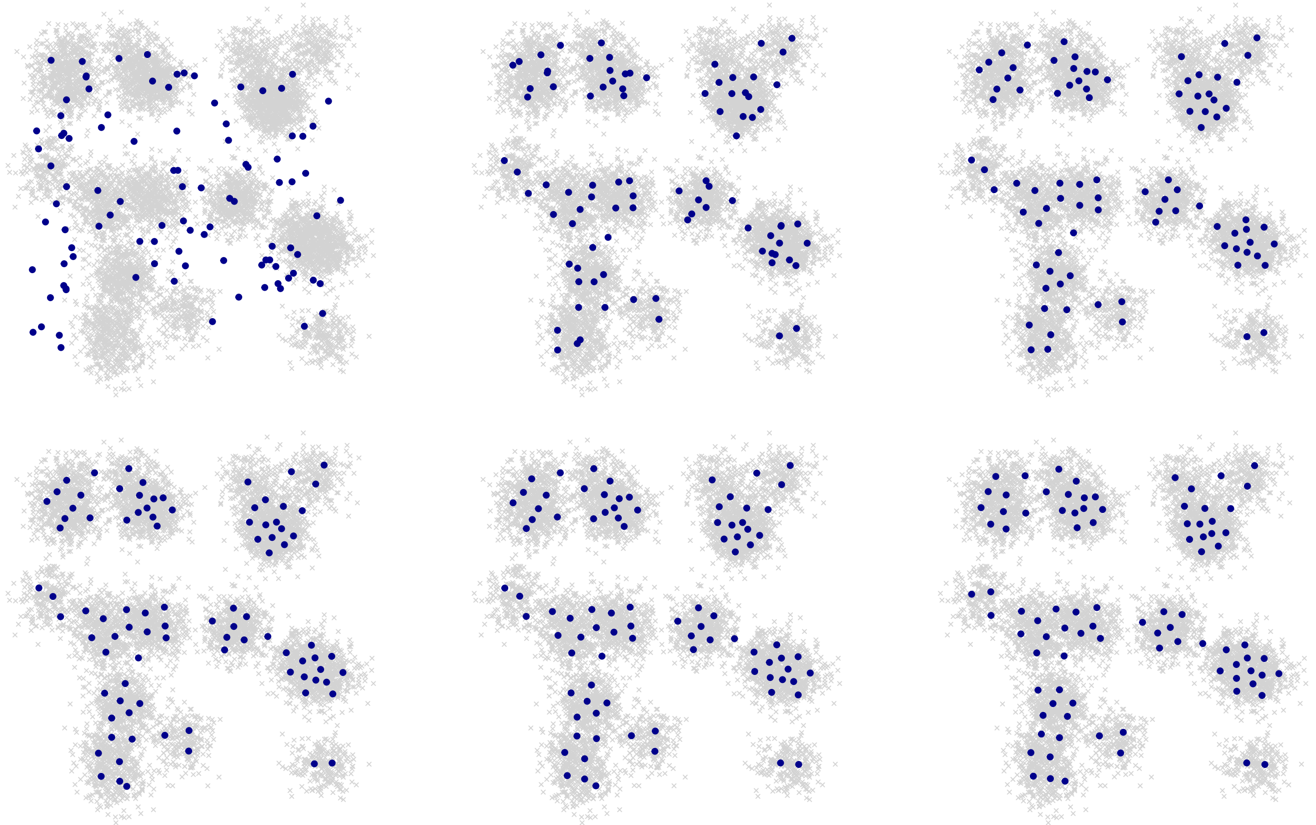


Fig. 3 Stepwise evolution of the barycenter for $k = 80$, $m_{\#} = 100$. In the first iteration, 32 points are deleted and 25 added. After that only movements take place

4.3 Practical aspects

As it turns out, the upper bound $\tilde{n}$ on the cardinality of the barycenter from Theorem 4 is often far too large in practice. For efficiency reasons, we typically run the algorithm with a number $n \geq \max\{n_j;\ 1 \leq j \leq k\}$ that is much smaller than $\tilde{n}$. We generate a starting point pattern center by picking $\frac{1}{k}\sum_{j=1}^{k}|\xi_j|$ points uniformly at random from the underlying observation window. In a first step, all point patterns are filled to n points by adding points at $\aleph$. Then, Algorithm 1 or 5 is run.

Figure 3 shows a typical run of Algorithm 1 in the case of an i.i.d. sample $\xi_1, \ldots, \xi_k$ of point patterns in $\mathbb{R}^2$ generated from a similar distribution as studied in Sect. 5. We use Euclidean distance and $p = 2$. The current barycenter is marked by blue points. Typically, the random starting point pattern is not a good approximation to the resulting barycenter. Therefore, many points are deleted in the first iteration. Many other ones are added at or moved to more cost efficient spots. Regardless of the starting pattern, the algorithm typically attains a reasonably looking configuration after a single iteration. After that hardly any points are added or deleted any more. The algorithm mostly moves a few individual barycenter points around each time.

5 Simulation study

In this section, we present a simulation study for evaluating the algorithms described in Sect. 4 for point patterns $\xi_1, \ldots, \xi_k$ in $\mathbb{R}^2$ using squared Euclidean cost.

Unfortunately, it is not feasible for larger data examples to compute the actual barycenter as a ground truth. To illustrate this, consider the special case where all point patterns have the same cardinality n and are contained in a subset of $\mathbb{R}^2$ of radius C. Assume further that we know that there is a barycenter that also has cardinality n (which need not be the case). In this situation, it is easy to see that instead of solving the minimization problem (13), we only need to minimize

$$\sum_{j=1}^{k}\sum_{i=1}^{n}\|x_{\pi_j(i),j} - z_i\|^2 \tag{14}$$

in $\pi_1, \ldots, \pi_k \in S_n$ and $z_1, \ldots, z_n \in \mathbb{R}^2$. This is the assignment version of the problem of finding a barycenter of the discrete probability measures $\frac{1}{n}\xi_1, \ldots, \frac{1}{n}\xi_k$ with respect to the Wasserstein metric W_2. An exact algorithm for this problem can be found in Anderes et al. (2016) and has been tremendously improved in Borgwardt and Patterson (2018).

Algorithm 5: `kMeansBary2`. A simple version of the improved `kMeansBary`-algorithm. Dependence on data pplist suppressed for simplicity.

```
Input  : center, pplist, $\delta > 0$, $N$ are as in the original
         kMeansBary-algorithm;
         $N_{\mathrm{del/add}}$ number of iterations during
           which we perform delete/add steps.
Output : Locally optimal pseudo-barycenter center.
$l \leftarrow \lceil \log_{10}(10^8/\frac{1}{n+1}) \rceil$;
epsvec $\leftarrow (\frac{1}{n+1} 10^{l-i})_{1 \le i \le l}$;
perm, cost $\leftarrow$ optimPerm(center, epsvec);
for it $\leftarrow 1$ to $N$ do
    costold $\leftarrow$ cost;
    center $\leftarrow$ optimBary(perm, center);
    if it $\le N_{\mathrm{del/add}}$ then
        center $\leftarrow$ optimDelete(perm, center);
        center $\leftarrow$ optimAdd(perm, center);
    end
    if it $< l$ then
        perm, cost $\leftarrow$ optimPerm(center, epsvec[1 : it]);
    else
        perm, cost $\leftarrow$ optimPerm(center, epsvec);
        if costold $-$ cost $< \delta$ then break;
          // difference always nonnegative
    end
end
return center;  // warn if the loop has run out
```

Nevertheless, the computation times increase still rapidly in the problem size and reach minutes to hours for problem sizes well smaller than our smallest examples below.

Since we are not able to compare the results of our algorithm to the actual barycenter for larger examples, we assess the range of the final objective function values. In addition, we evaluate the time performance of the default algorithm and compare both objective function values and timings to the improved algorithm.

As problem instances, we created sets of k point patterns in $\mathbb{R}^2$ having mean cardinality of $m_\#$ in each pattern. The cardinalities n_j, $j \in [k]$, of the individual point patterns were generated by one of the following methods:

(i) by setting $n_j = m_\#$ *(deterministic cardinality)*
(ii) by sampling n_j from a binomial distribution with mean $m_\#$ and variance ≈ 1
(low-variance cardinality)
(iii) by sampling n_j from a Poisson distribution with parameter $m_\#$
(high-variance cardinality)

The points were distributed according to a balanced mixture of $N \in \{5, 10, 15\}$ rotationally symmetric normal densities centered at fixed locations in $[0, 1]^2$ and having standard deviation $\sigma \in \{0.05, 0.1, 0.2\}$. Figure 4 gives examples under the three center scenarios for $k = 20$, deterministic cardinality $n_j = m_\# = 20$ and $\sigma = 0.05$.

We chose five $(k, m_\#)$ pairs $(20, 20)$, $(20, 50)$, $(50, 20)$, $(50, 50)$ and $(100, 100)$, which in combination with N varying in $\{5, 10, 15\}$, σ in $\{0.05, 0.1, 0.2\}$ and the three cardinality distributions yield a total of $5 \times 3^3 = 135$ scenarios. We created 100 instances for each scenario.

Our algorithms from Sect. 4 were run from ten starting solutions whose cardinalities matched the mean number of data points and whose points were sampled uniformly at random from $[0, 1]^2$. In a pilot experiment, this tended to give somewhat better local minima than starting from a random sample of all data points combined. The starting point patterns were independently chosen for each instance, but the same for both algorithms. In all cases, the penalty C was set to 0.1.

Tables 1, 2, 3 and 4 summarize the performance of our two algorithms. For the clarity of presentation, we leave

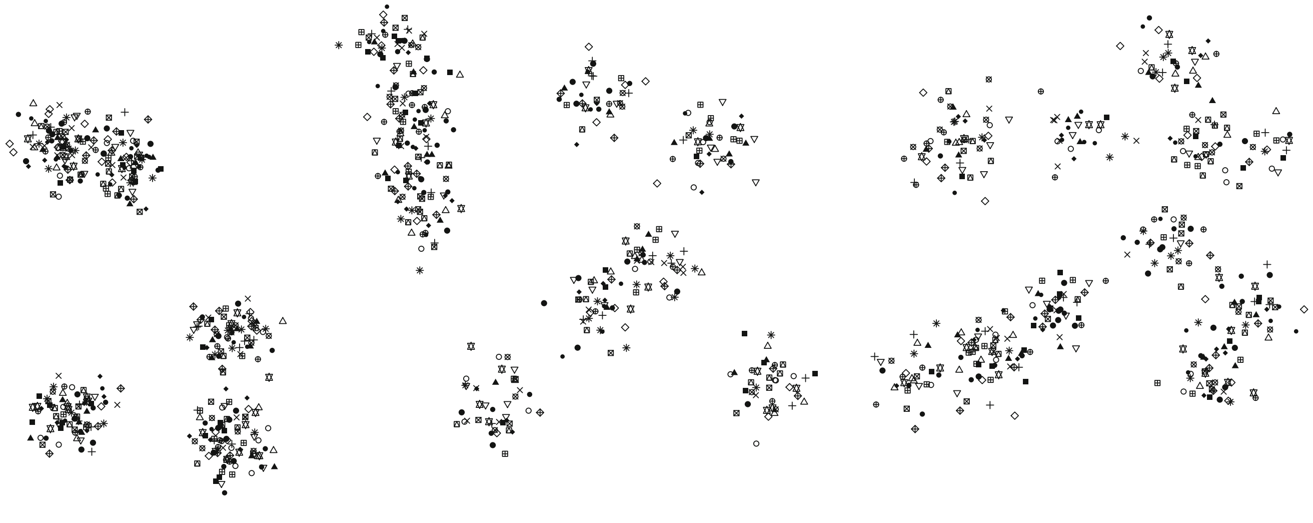


Fig. 4 20 point patterns with 20 points each from the three different center scenarios $N = 5, 10, 15$ for $\sigma = 0.05$

Table 1 *Original algorithm.* Maximum relative deviations from the minimum objective function value among ten starting solutions *given in percent*. Means taken over 100 instances, with 0.05- and 0.95-quantiles in parentheses. The first block of four rows corresponds to the deterministic cardinality, the second block to the high-variance cardinality

N	σ	20/20	20/50	50/20	50/50	100/100
5	0.05	4.59 (2.67, 7.41)	3.82 (1.93, 5.88)	4.00 (2.59, 5.93)	3.54 (2.18, 5.04)	3.27 (1.80, 4.78)
	0.2	2.06 (1.03, 3.21)	2.41 (1.42, 3.29)	1.21 (0.62, 2.03)	1.71 (1.06, 2.35)	1.32 (0.74, 1.89)
15	0.05	3.22 (1.72, 5.18)	3.31 (2.09, 5.00)	2.50 (1.25, 4.22)	2.74 (1.78, 3.89)	2.66 (1.56, 3.72)
	0.2	1.54 (0.56, 2.50)	2.19 (1.28, 3.11)	0.62 (0.01, 1.21)	1.50 (0.91, 2.21)	0.99 (0.52, 1.43)
5	0.05	3.48 (1.78, 5.46)	2.35 (1.40, 3.91)	2.82 (1.34, 4.50)	1.86 (1.04, 3.15)	1.16 (0.63, 1.81)
	0.2	1.81 (1.02, 3.01)	2.41 (1.46, 3.41)	0.97 (0.56, 1.49)	1.54 (0.97, 2.16)	1.02 (0.56, 1.66)
15	0.05	2.73 (1.48, 4.28)	2.52 (1.59, 3.89)	2.25 (1.22, 3.59)	2.04 (1.18, 3.01)	1.41 (0.79, 2.14)
	0.2	1.16 (0.42, 1.99)	2.13 (1.15, 3.13)	0.42 (0.00, 1.01)	1.42 (0.82, 2.03)	1.14 (0.61, 1.76)

Table 2 *Improved algorithm.* Maximum relative deviations from the minimum objective function value of the original algorithm (both based on the same ten starting solutions) *given in percent*. Means over 100 instances, with 0.05- and 0.95-quantiles in parentheses. The first block of four rows corresponds to the deterministic cardinality, the second block to the high-variance cardinality

N	σ	20/20	20/50	50/20	50/50	100/100
5	0.05	3.84 (1.93, 6.02)	3.77 (2.06, 5.76)	3.81 (2.02, 5.88)	3.80 (2.07, 5.63)	3.38 (1.88, 5.13)
	0.2	1.66 (0.88, 2.69)	2.49 (1.67, 3.40)	1.01 (0.45, 1.70)	1.77 (1.13, 2.42)	1.35 (0.84, 1.92)
15	0.05	3.23 (1.89, 4.81)	3.41 (2.24, 4.89)	2.61 (1.39, 4.04)	2.79 (1.67, 4.07)	2.64 (1.62, 3.82)
	0.2	1.04 (0.37, 1.81)	2.21 (1.33, 3.35)	0.40 (0.01, 0.90)	1.59 (0.99, 2.32)	0.99 (0.55, 1.41)
5	0.05	3.32 (1.80, 5.27)	2.06 (1.09, 3.12)	2.55 (1.36, 4.01)	1.67 (0.85, 2.68)	1.08 (0.49, 1.74)
	0.2	1.53 (0.63, 2.55)	2.28 (1.32, 3.40)	0.79 (0.34, 1.33)	1.53 (0.85, 2.28)	1.02 (0.59, 1.60)
15	0.05	2.49 (1.22, 4.09)	2.50 (1.63, 3.57)	2.19 (1.11, 3.57)	2.08 (1.20, 3.17)	1.33 (0.68, 2.30)
	0.2	0.82 (0.26, 1.46)	2.26 (1.44, 3.48)	0.29 (0.00, 0.77)	1.39 (0.84, 1.96)	1.12 (0.64, 1.64)

Table 3 *Original algorithm.* Total times in seconds *for ten runs* with random starting patterns. Means over 100 instances, with 0.05- and 0.95-quantiles in parentheses. The first block of four rows corresponds to the deterministic cardinality, the second block to the high-variance cardinality

N	σ	20/20	20/50	50/20	50/50	100/100
5	0.05	0.48 (0.47, 0.50)	0.89 (0.85, 0.93)	1.02 (0.99, 1.05)	2.37 (2.24, 2.52)	20.79 (19.42, 22.32)
	0.2	0.51 (0.49, 0.53)	1.14 (1.07, 1.23)	1.05 (0.99, 1.12)	3.55 (3.26, 3.96)	37.95 (34.70, 42.22)
15	0.05	0.50 (0.49, 0.51)	0.95 (0.91, 1.00)	1.06 (1.03, 1.09)	2.60 (2.44, 2.80)	24.25 (22.51, 25.93)
	0.2	0.49 (0.47, 0.51)	1.14 (1.08, 1.22)	0.91 (0.85, 0.99)	3.51 (3.12, 3.85)	37.77 (33.86, 42.21)
5	0.05	0.58 (0.53, 0.64)	1.27 (1.11, 1.53)	1.37 (1.24, 1.55)	4.12 (3.38, 4.97)	39.66 (33.33, 46.96)
	0.2	0.59 (0.53, 0.67)	1.58 (1.37, 1.88)	1.30 (1.11, 1.47)	5.62 (4.64, 6.86)	65.86 (54.51, 77.49)
15	0.05	0.59 (0.55, 0.65)	1.32 (1.16, 1.56)	1.41 (1.26, 1.64)	4.31 (3.59, 5.22)	45.37 (38.88, 51.93)
	0.2	0.55 (0.49, 0.61)	1.55 (1.37, 1.81)	1.06 (0.94, 1.23)	5.73 (4.72, 7.23)	66.25 (56.73, 79.16)

out the “middle values” $N = 10$ and $\sigma = 0.1$, as well as the low-variance cardinality distribution. The corresponding performance results lie up to minor random fluctuations between the values shown. The original purpose of including the low-variance cardinality case was to detect whether a slight departure from equal cardinalities would cause substantial differences in the performance. As it turned out, this was not the case.

We first consider the original algorithm presented in Sect. 4. Table 1 gives the maximum relative deviation from the minimum $d_{\min}$ of the resulting objective function values among the ten starting solutions, i.e., $\frac{d_{\max}-d_{\min}}{d_{\min}}$. We can see that the maximal objective function value among the ten runs rarely exceeds the minimum value by more than 5%. This percentage is rather higher for the deterministic and low-variance cardinalities and when clusters in the (unmarked) superposition of the point patterns are well separated (small

Table 4 *Improved algorithm.* Total times in seconds *for ten runs* with random starting patterns. Means over 100 instances, with 0.05- and 0.95-quantiles in parentheses. The first block of four rows corresponds to the deterministic cardinality, the second block to the high-variance cardinality

N	σ	20/20	20/50	50/20	50/50	100/100
5	0.05	0.48 (0.47, 0.48)	0.83 (0.81, 0.85)	0.97 (0.96, 0.99)	2.05 (1.99, 2.12)	13.68 (13.02, 14.57)
	0.2	0.50 (0.49, 0.50)	0.93 (0.89, 0.98)	1.00 (0.98, 1.02)	2.47 (2.33, 2.65)	18.25 (16.87, 20.57)
15	0.05	0.49 (0.48, 0.49)	0.85 (0.82, 0.88)	0.98 (0.97, 1.00)	2.08 (1.99, 2.20)	13.24 (12.47, 14.26)
	0.2	0.50 (0.49, 0.50)	0.92 (0.89, 0.95)	0.96 (0.92, 0.99)	2.42 (2.27, 2.57)	17.62 (16.08, 19.33)
5	0.05	0.56 (0.52, 0.61)	1.14 (1.00, 1.33)	1.24 (1.14, 1.36)	3.21 (2.82, 3.73)	22.47 (19.86, 25.53)
	0.2	0.58 (0.54, 0.63)	1.23 (1.09, 1.43)	1.21 (1.12, 1.30)	3.64 (3.25, 4.18)	29.41 (25.54, 34.47)
15	0.05	0.56 (0.52, 0.60)	1.12 (0.99, 1.27)	1.24 (1.13, 1.37)	3.17 (2.76, 3.71)	22.64 (19.72, 25.59)
	0.2	0.56 (0.53, 0.61)	1.19 (1.05, 1.36)	1.14 (1.04, 1.27)	3.60 (3.06, 4.35)	29.11 (24.94, 34.51)

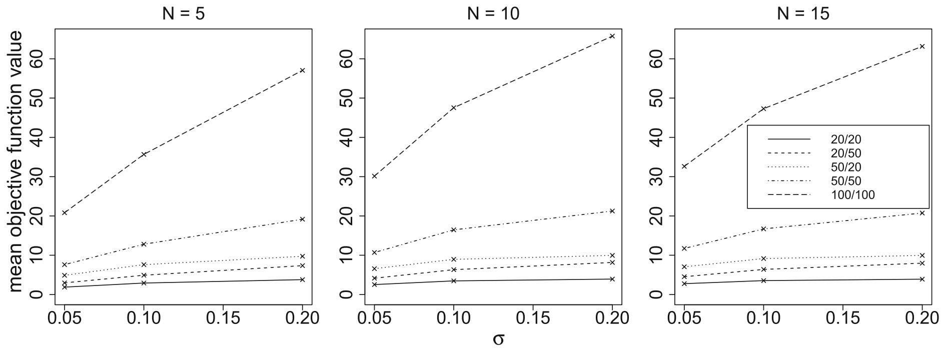


Fig. 5 Mean objective function values over all instances as function of σ for different N

N and σ). This may well be explicable by the fact that typically many pairs can be matched over short distances in these situations such that wrong clustering decisions come typically at a higher relative cost. Figure 5 supports this by showing that the total objective function values within each problem size are lower for well separated clusters.

A further smaller experiment following up on the scenarios that exhibited the poorest performance for ten starting patterns showed that the margin of 5% increases to 8% when basing the maximum relative deviation from the minimum on 100 starting patterns.

For the improved algorithm from Sect. 4.2, we compute the maximum relative deviation of its objective function values from the minimum $d_{\min}$ of the corresponding values of the original algorithm, i.e., $\frac{d^*_{\max}-d_{\min}}{d_{\min}}$, where $d^*_{\max}$ is the maximum of the objective function values of the improved algorithm. As seen in Table 2, the performance is no worse than for the original algorithm in spite of the reduced amount of computations performed.

We finally turn to the computation times. We present the total runtimes in seconds for the ten runs with different starting patterns. This corresponds to the realistic situation of selecting as (pseudo-)barycenter the solution with the smallest local minimum in ten runs. It also provides some more stability for the means and quantiles given in Tables 3 and 4.

Table 3 gives the runtimes for the original algorithm. We see that individual runs of as large scenarios as 100 patterns with 100 points on average only take a few seconds.

From Table 4, we see that the runtimes for the improved algorithm are even considerably lower, and for some of the larger problems they have less than half of the original runtimes (at virtually no loss with regard to the objective function value as we have seen before). It is to be expected that this ratio becomes even smaller if the problem size is further increased.

Let us finally compare our algorithm to an algorithm that treats point patterns as empirical measures and tackles the Wasserstein-2 barycenter problem for these measures. As noted in the introduction, it is not realistic to treat even our smallest examples with exact algorithms for this problem. A selection of approximate algorithms can be found in Peyré and Cuturi (2019). See also the alternating algorithm in Borgwardt (2019), which includes a factor-2 performance guarantee. For our comparison, we choose Algorithm 2 in Cuturi and Doucet (2014), which alternates between solving transport problems and using gradient descent to calculate a discrete barycenter with a prescribed maximal number of support points. It allows to restrict the set of weights for the support points to a closed convex set Θ and thus provides an approximate solution of the problem (14) if we set $\Theta = \{(1/n, \ldots, 1/n)\}$.

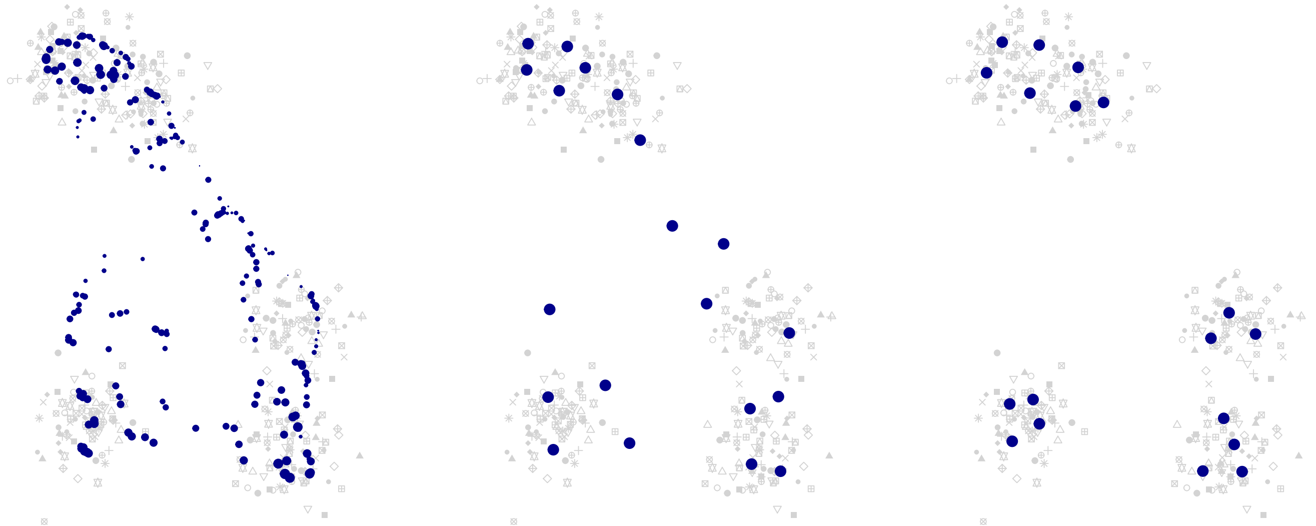


Fig. 6 Barycenters for one of our simulated data sets (20 patterns with 20 points each). From left to right: Cuturi–Doucet algorithm without constraints, Cuturi–Doucet algorithm with (maximally) 20 support points and equal masses, typical result from `kMeansBary` algorithm based on a single start. The areas of the disks are proportional to the masses

We thank Florian Heinemann for allowing us to use his R implementation with underlying C++ code of this algorithm. Figure 6 shows an example that compares the Cuturi–Doucet algorithm without constraint using the theoretically maximal number of support points (according to Anderes et al. (2016)), the Cuturi–Doucet algorithm with full constraint and our algorithm.

To evaluate how the algorithms perform on our objective function (13), we have run both the fully constrained Cuturi–Doucet algorithm and our `kMeansBary` algorithm (with a single starting value) on the smallest scenarios used in the simulation study. These are 900 instances of 20 patterns with exactly 20 points (deterministic cardinality) and 900 instances of 20 patterns whose cardinalities are Poisson with mean 20 (high-variance cardinality).

We report the ratio of the total TT-objective function (13) between the solution of `kMeansBary` and the Cuturi–Doucet algorithm, where again $C = 0.1$. For the case of deterministic cardinality, the ratio was 0.729 on average, with a minimum of 0.554 and a maximum of 0.871. For the high-variance cardinality, the results are very similar with an average of 0.732 and a minimum of 0.541 and a maximum of 0.866. So on average the objective function values attained by the point patterns returned by the Cuturi–Doucet algorithm are about 37% larger than the ones attained by `kMeansBary`. This increase is reflected in the example in Fig. 6.

At the same time, the average runtime of the Cuturi–Doucet algorithm is more than twice the runtime of `kMeansBary`. This may well be due to the fact that the former is not particularly optimized for the constrained setting we use.

Overall, our comparison yields that the Cuturi–Doucet algorithm is not well suited for our problem, which is simply due to the fact that this algorithm was designed for a somewhat different problem. We expect similar results when comparing with other algorithms that compute (approximate) Wasserstein-2 barycenters.

6 Applications

The following analyses are all performed in R, see R Core Team (2019), with the help of the package spatstat, see Baddeley et al. (2015).

6.1 Street theft in Bogotá

We investigate a data set of person-related street thefts in Bogotá, Colombia, during the years 2012–2017. This data set is part of a huge data set based on a large number of types of crimes collected by the Dirección de Investigación Criminal e Interpol (DIJIN), a department of the Colombian National Police. We acknowledge DIJIN and the General Santander National Police Academy (ECSAN) for allowing us to use this data. In particular, the cases of street theft in Bogotá consist of muggings, which involve the use of force or threat, as well as pickpocketing. They do not include theft of vehicles, breaking into cars, etc. Here, we focus on the locality of Kennedy, a roughly 7.5 km × 7.5 km administrative ward in the west of the city, because this area is considered by the police as being more dangerous with a higher average number of crime events compared to the rest of Bogotá. The

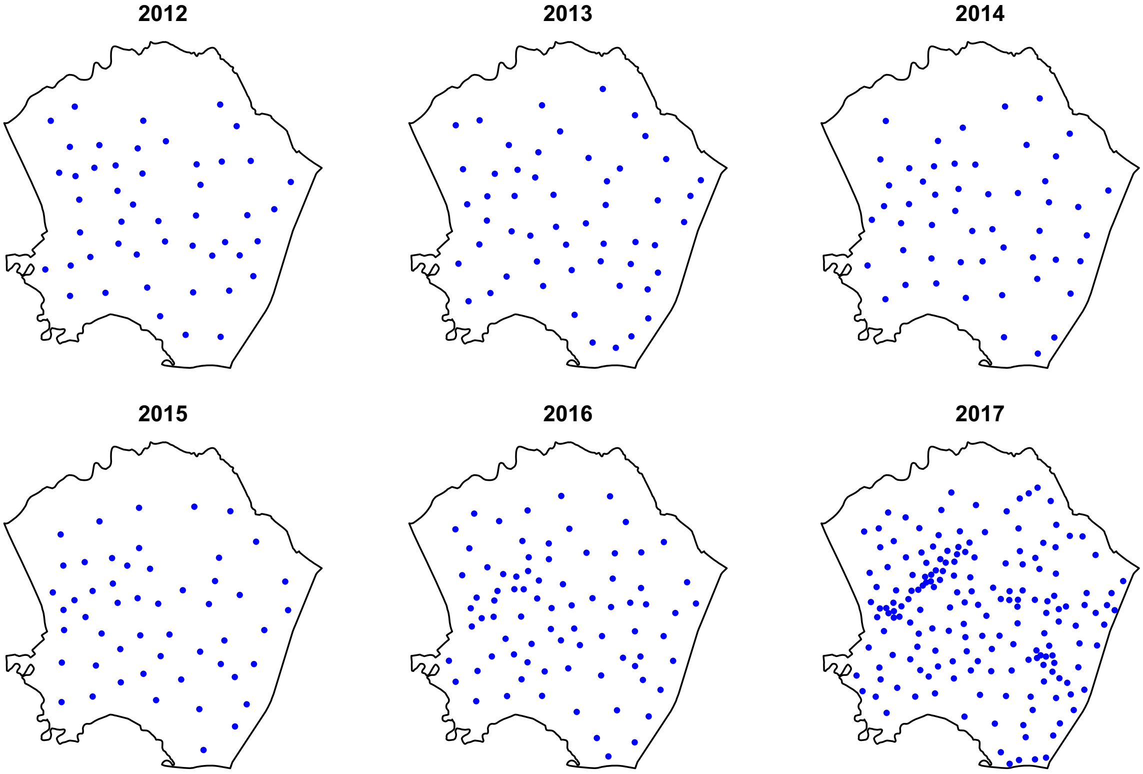


Fig. 7 Barycenters of weekly street thefts in the localidad of Kennedy in Bogotá. The cardinalities are 48, 53, 52, 52, 80 and 175, respectively

total number of street thefts in Kennedy for the considered period is 25 840.

Since a plot of weekly numbers of crimes reveals no clear seasonal pattern and since weekly patterns (and hence their barycenters) are of a good size to be interpreted graphically, we compute yearly barycenters for these weekly patterns. Thus, we may think of a barycenter pattern as representing a "typical week" of street thefts in the corresponding year. As penalty parameter, we chose 1000 m. Since street information was not directly available to us, we chose Euclidean distance as a metric and set $p = 2$ to be able to relate to our simulation results in the previous section.

Each barycenter was computed based on 100 starting patterns with cardinalities regularly scattered over the integer numbers between the 0.45 to the 0.7 quantiles of the weekly number of data points for the corresponding year. We chose this somewhat asymmetrically around the median, because the mean number of thefts (the theoretical number of points in the barycenter if the penalty becomes large) was typically quite a bit to the right of the median, and also because our algorithm is somewhat better at deleting than at adding points.

Figure 7 depicts the obtained barycenters, which except for the last pattern have cardinalities just slightly below the average weekly numbers of muggings of 51.7, 57.6, 52.8, 54.4, 82.5 and 196.9, respectively. The barycenters for the years 2012–2015 seem to be largely similar. Then, in 2016 we start seeing patterns of denser structures forming along a line to the west and a center in the south-east of Kennedy. These can be actually identified as a main street and a major intersection in the densely populated parts of Kennedy.

6.2 Assault cases in Valencia

As a second application, we analyze cases of assault in Valencia, Spain, reported to the police in the years 2010–2017. Since the addresses of the assaults and the street network are available, we treat this data as point patterns on a graph using shortest-path distance and $p = 1$. We acknowledge the local police in Valencia city together with the 112 emergency phone that kindly provided us the data after cleaning and removing any personal information.

We split up the graph and analyze the four central districts of Ciutat Vella, Eixample, Extramurs and El Pla del Real separately. For this, we assigned each assault case to its district, but added also streets from other districts at the boundary, in order to enable more natural shortest-path computations. The north-south and east-west extensions of the districts vary roughly between 1.6 and 3.3 km.

In the time domain, we split up the assault data by year and season into seven winter patterns (data from December, January and February) and eight summer patterns (data from

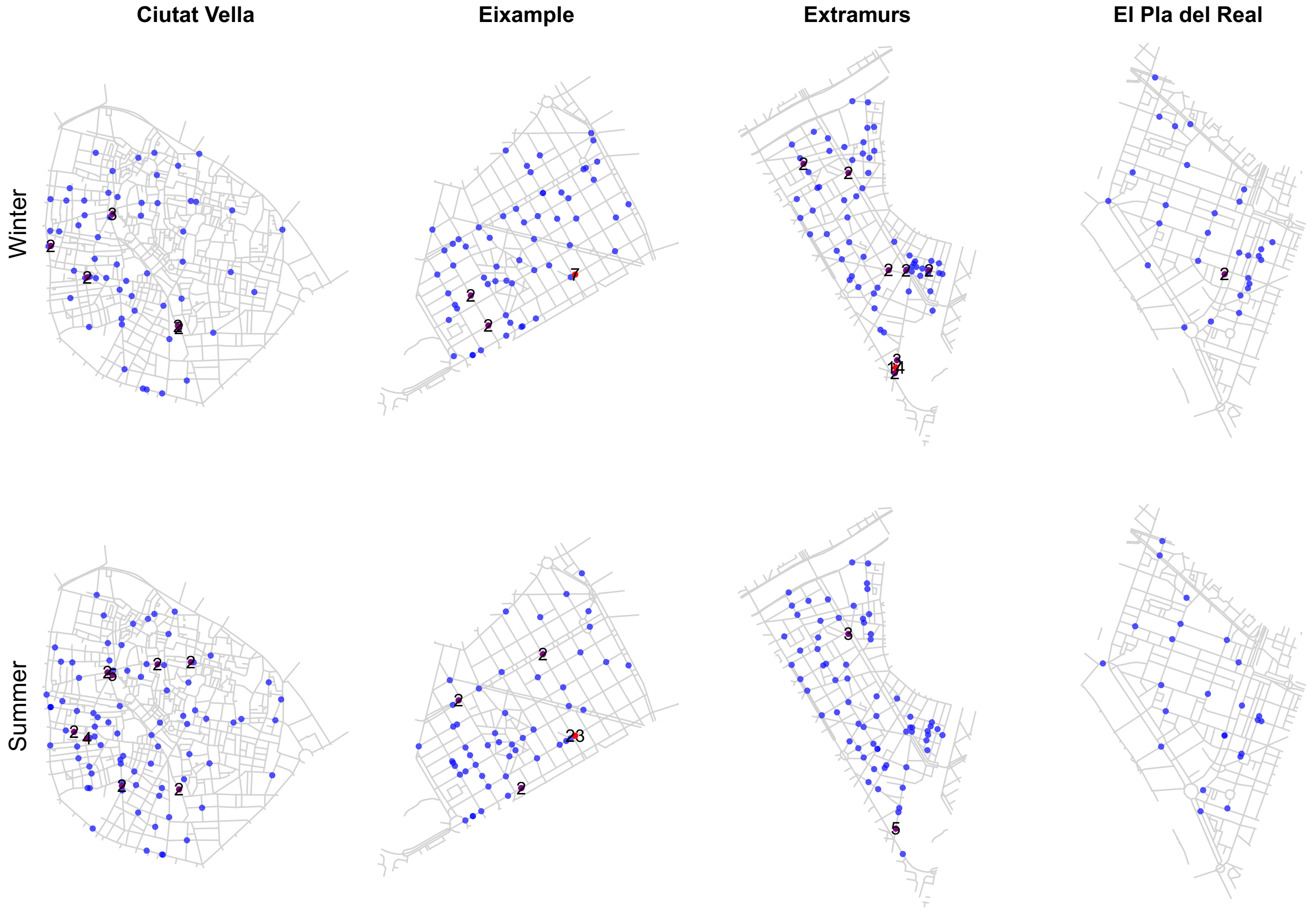


Fig. 8 Barycenters of cases of assault for different districts of Valencia in winter and in summer. The numbers indicate multiplicities if there are several points at a single location. The cardinalities of the barycenters are 68, 69, 88, 30 (winter) and 103, 79, 74, 24 (summer)

June, July and August), discarding for the present analysis data from the intermediate seasons, as well as from January and February 2010 and December 2017. We then computed barycenters per main season and district, obtaining "typical" assault patterns for summer and winter for each of the four districts considered, see Fig. 8. The penalty was chosen as 800 m with respect to the shortest-path distance.

As mentioned in Sect. 4.1 when describing the subroutine `optimBary` that finds cluster centers on networks, we can calculate all distances that are relevant to the algorithm beforehand. For this, we use the corresponding functionality built into the `linnet` objects in spatstat. On a standard laptop with a 1.6 GHz Intel i5 processor, the computation took only about four seconds for the largest data set, which is Ciutat Vella in summer with a total of 2494 vertices (1676 street crossings plus 818 data points).

For the starting patterns in each district in summer and in winter, we chose n random points, where n ranged from 0.8 times to 1.15 times the median cardinality of the data point patterns. Since our present implementation of the `kmeansbary` algorithm on graphs runs without `optimDelete` and `optimAdd` steps, we based each barycenter on a large sample of 500 starting patterns for each n. This resulted in an overall total of 101 500 calls to our algorithm for the eight scenarios, which on average took 0.57 seconds each, using the precomputed distance matrices. One calculation in the largest setting (Ciutat Vella in summer) takes about 0.82 seconds and in the smallest (El Pla del Real in summer) about 0.08 seconds. The increase in the objective function was only up to 1% when decreasing the total number of calls to our algorithm by a factor of 20, resulting in a total computation time of well under one hour.

Due to the choice of $p = 1$, it happens quite frequently that there are several optimal centers for some of the clusters obtained after convergence of the `kmeansbary` algorithm. In this case, we take the average of their coordinates and project the result back onto the graph in order to obtain a somewhat more balanced result. The resulting point does not necessarily realize the same cluster cost as the original center points, but on a real street network it is not to be expected that the cost becomes considerably worse. In fact, for the data considered, the results hardly differed at all.

Considering the barycenters in Fig. 8, it seems that there are no very clear effects of the season on the assaults. Nevertheless, we may discern a number of differences between summer and winter in the four districts, which even in this relatively small data set would be considerably harder to spot in a plot of the raw data.

In the first district (Ciutat Vella), there are substantially more assaults in summer, but their spatial distribution in the barycenter is more or less similar. In Eixample, we see a concentration of assault cases in summer in the Barrio Ruzafa in the southern half of the district, whereas cases are more or less equally spread in winter. A notable feature is the occurrence of 23 barycenter points at a single crossing in summer and 7 points at the same crossing in winter with further points close by. This is due to the cul-de-sac visible in Fig. 8, which in reality forms sort of a backyard that makes the area easy for assaults, especially in summer time when there are more people (especially tourists) moving around those parts of the city. The spot is well-known to the police and in recent years the number of assaults has decreased due to police interventions. The barycenters clearly reflect this (former) assault hot spot.

In the district of Extramurs, both barycenters are more or less spread over the whole district, with two clusters of assaults occurring in the east and south. Both clusters are much more pronounced in winter. In the district of El Pla del Real, there is some concentration in the winter month in the east and south-east. Apart from that the only noticeable difference is that there are substantially more assaults in winter than in summer, which may well be related to the fact that this is a popular student district.

7 Discussion and outlook

In this paper, we have introduced the p-th-order TT- and RTT-metrics, which allow us to measure distances between point patterns in an intuitive way, generalizing several earlier metrics. We have investigated q-th-order barycenters with respect to the TT-metric and presented two variants of a heuristic algorithm. These variants return local minimizers of the Fréchet functional that mimic properties of the actual barycenter well and attain consistent objective function values. They are computable in a few seconds for medium-sized problems, such as 100 patterns of 100 points.

For the proof of Theorem 4, it was necessary to set $p = q$. While such a choice may seem natural, we point out that due to the separate interpretations of p as the order for matching points in the metric on $\mathfrak{N}_{\mathrm{fin}}$ (higher p tends to balance out the matching distances) and q as the order of the empirical moment in $\mathfrak{N}_{\mathrm{fin}}$, it may well be desirable to combine $p \neq q$.

In the present paper, we have only dealt with the descriptive aspects of barycenters. It is thus clear that our applications in Sect. 6 can only be seen as explorative studies. In order to determine whether differences between group barycenters are statistically significant, we need to take the distribution of the point patterns around their barycenters into account and perform appropriate hypothesis tests.

Fortunately, the Fréchet functional (7) provides us with a natural quantification of scatter around the barycenter. For $q = 2$, it is quite common to refer to

$$\mathrm{Var}(\xi_1, \ldots, \xi_k) = \min_{\zeta \in \mathfrak{N}_{\mathrm{fin}}} \frac{1}{k} \sum_{j=1}^{k} \tau(\xi_j, \zeta)^2$$

as (empirical) Fréchet variance, due to Equation (8). Detailed asymptotic theory for performing an analysis of variance (ANOVA) in metric spaces based on comparing Fréchet variances has been recently developed in Dubey and Müller (2019a). The application and adaptation of this theory for the point pattern space and an investigation of the performance of our heuristic algorithm in this context will be the subject of a future paper.

Based on the computation of barycenters further more advanced procedures in statistics and machine learning become possible. This includes barycenter-based dimension reduction techniques, such as Wasserstein dictionary learning, see Schmitz et al. (2018), and functional principal component analysis of point patterns evolving in time, see Dubey and Müller (2019b).

Acknowledgements Open Access funding provided by Projekt DEAL.

Appendix: Proofs left out in the main text

Lemma A.1 *Let $C > 0$, $\widetilde{C} \in (0, 2C]$ and let $(\mathcal{X}, d)$ be a metric space with $\mathrm{diam}(\mathcal{X}) = \sup_{x,y \in \mathcal{X}} d(x, y) \leq 2C$. For $k \in \mathbb{N}$ set $\mathcal{X}' = \mathcal{X} \cup \{\aleph_1, \ldots, \aleph_k\}$, where $\aleph_1, \ldots \aleph_k \notin \mathcal{X}$ are pairwise different, and define*

$$d'(x, y) = \begin{cases} d(x, y) & \text{if } x, y \in \mathcal{X}; \\ C & \text{if } \{x, y\} \cap \mathcal{X} \neq \emptyset \\ & \text{and } \{x, y\} \cap \{\aleph_1, \ldots, \aleph_k\} \neq \emptyset; \\ \widetilde{C} & \text{if } \{x, y\} \subset \{\aleph_1, \ldots, \aleph_k\} \text{ and } x \neq y; \\ 0 & \text{if } x = y = \aleph_i \text{ for some } i \in [k]. \end{cases}$$

Then, $(\mathcal{X}', d')$ is a metric space.

Proof Identity and symmetry properties of the map $d'\colon \mathcal{X}' \times \mathcal{X}' \to \mathbb{R}_+$ follow immediately. Since d is a metric on $\mathcal{X}$ and $\tilde{d}(x, y) = \widetilde{C}$ if $x \neq y$ (and 0 otherwise) defines a metric $\tilde{d}$ on $\mathcal{Y} = \{\aleph_1, \ldots, \aleph_k\}$, we only have to check the triangle inequality for a few special cases. If $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ (or vice versa), then $d(x, y) = C$. Since one of $d(x, z)$ and $d(z, y)$ has to be $= C$ regardless of $z \in \mathcal{X}'$, we obtain $d(x, y) \leq d(x, z) + d(z, y)$. If $x, y \in \mathcal{X}$ and $z \in \mathcal{Y}$, then

$$d(x, y) \leq \operatorname{diam}(\mathcal{X}) \leq 2C = d(x, z) + d(z, y).$$

Likewise, if $x, y \in \mathcal{Y}$ and $z \in \mathcal{X}$, then

$$d(x, y) \leq \widetilde{C} \leq 2C = d(x, z) + d(z, y).$$

□

Proof of Theorem 1 Since $\bar{\tau}(\xi, \eta) = \frac{1}{n^{1/p}}\tau(\xi, \eta)$, it is enough to show the statement for τ.

Let $\pi \in S_n$ be a permutation that minimizes $\sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p$. Writing $I = \big\{i \in [m];\ d(x_i, y_{\pi(i)}) < 2^{1/p}C\big\}$, we obtain

$$d'(x_i, y_{\pi(i)}) = \begin{cases} d(x_i, y_{\pi(i)}) & \text{if } i \in I; \\ 2^{1/p}C & \text{if } i \in [m]\backslash I; \\ C & \text{if } i \in [n]\backslash[m]. \end{cases} \tag{A.15}$$

Therefore, enumerating I in arbitrary order as $\{i_1, \ldots, i_l\}$ for some $l \in [m]$ and setting $j_r := \pi(i_r)$ for $r \in [l]$, we have

$$\begin{aligned} &\sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \\ &\quad = \sum_{r=1}^{l} d(x_{i_r}, y_{j_r})^p + (m - l)(2^{1/p}C)^p + (n - m)C^p \\ &\quad = (m + n - 2l)C^p + \sum_{r=1}^{l} d(x_{i_r}, y_{j_r})^p. \end{aligned} \tag{A.16}$$

Thus, $\tau(\xi, \eta)^p \leq \min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p$.

Conversely, let $(i_1, \ldots, i_l; j_1, \ldots, j_l) \in S(m, n)$ minimize $(m + n - 2l)C^p + \sum_{r=1}^{l} d(x_{i_r}, y_{j_r})^p$. This implies $d(x_{i_r}, y_{j_r}) \leq 2^{1/p}C$ for all $r \in [l]$, because otherwise we could obtain a smaller value by removing i_r, j_r from the vector. Writing $I = \{i_1, \ldots, i_l\}$, $J = \{j_1, \ldots, j_l\}$ it implies also that $d(x_i, y_j) \geq 2^{1/p}C$ for all $i \in [m]\backslash I$ and $j \in [n]\backslash J$, because otherwise we could obtain a smaller value by adding i, j to the vector. Let then $\pi \in S_n$ be any permutation satisfying $\pi(i_r) = \pi(j_r)$ for all $r \in [l]$. With this π, we obtain exactly the d'-distances in (A.15) for all $i \in [n]$ and hence (A.16) holds again. Thus, $\min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p \leq \tau(\xi, \eta)^p$. □

Proof of Theorem 2 We start with the map $\tau\colon \mathfrak{N}_{\text{fin}} \times \mathfrak{N}_{\text{fin}} \to \mathbb{R}_+$. If $\xi = \eta$, then $m = n$ and there is a permutation $\pi \in S_n$ such that $x_i = y_{\pi(i)}$ for $1 \leq i \leq n$. Hence $\tau(\xi, \eta) = 0$, choosing $l = n$ and $(i_1, \ldots, i_l; j_1, \ldots, j_l) = (1, \ldots, n, \pi(1), \ldots, \pi(n))$. If on the other hand $\tau(\xi, \eta) = 0$, we must have $l = m = n$ to be able to achieve $m + n - 2l = 0$ and there must be $(i_1, \ldots, i_n; j_1, \ldots, j_n) \in S(n, n)$ such that $d(x_{i_r}, y_{j_r}) = 0$ for $1 \leq r \leq n$. Since the d is a metric, this yields $\xi = \sum_{r=1}^{n} \delta_{i_r} = \sum_{r=1}^{n} \delta_{j_r} = \eta$. The symmetry of τ is immediately clear from the symmetric form of (2).

For the proof of the triangle inequality, we use the metric space $(\mathcal{X}', d')$ introduced before Theorem 1. Let $\xi, \eta, \zeta \in \mathfrak{N}_{\text{fin}}$. After filling up patterns to the maximum of the three cardinalities by adding points at the auxiliary location $\aleph$, we may assume that $\xi = \sum_{i=1}^{n} \delta_{x_i}$, $\eta = \sum_{j=1}^{n} \delta_{y_j}$ and $\zeta = \sum_{k=1}^{n} \delta_{z_k}$ have the same cardinality. Noting that given two point patterns of the same cardinality we may add any number of extra points located at $\aleph$ to both of them without changing their τ-distance, Theorem 1 yields that there are $\pi_1, \pi_2 \in S_n$ such that

$$\tau(\xi, \zeta) = \left(\sum_{i=1}^{n} d'(x_i, z_{\pi_1(i)})^p\right)^{1/p} \quad \text{and}$$
$$\tau(\zeta, \eta) = \left(\sum_{i=1}^{n} d'(z_i, y_{\pi_2(i)})^p\right)^{1/p}.$$

Then, $\pi = \pi_2 \circ \pi_1 \in S_n$ matches the points of ξ and η in such a way that

$$d'(x_i, y_{\pi(i)}) \leq d'(x_i, z_{\pi_1(i)}) + d'(z_{\pi_1(i)}, y_{\pi_2(\pi_1(i))})$$

and Theorem 1 and the triangle inequality for the ℓ_p-norm yields that

$$\begin{aligned} \tau(\xi, \eta) &\leq \left(\sum_{i=1}^{n} d'(x_i, y_{\pi(i)})^p\right)^{1/p} \\ &\leq \left(\sum_{i=1}^{n} d'(x_i, z_{\pi_1(i)})^p\right)^{1/p} \\ &\quad + \left(\sum_{i=1}^{n} d'(z_{\pi_1(i)}, y_{\pi_2(\pi_1(i))})^p\right)^{1/p} \\ &= \tau(\xi, \zeta) + \tau(\zeta, \eta). \end{aligned}$$

We turn to the map $\bar{\tau}\colon \mathfrak{N}_{\text{fin}} \times \mathfrak{N}_{\text{fin}} \to \mathbb{R}_+$. Since $\bar{\tau}(\xi, \eta) = \frac{1}{\max\{|\xi|, |\eta|\}^{1/p}}\tau(\xi, \eta)$, we may inherit the identity and symmetry properties for $\bar{\tau}$ directly from τ. To show the triangle inequality, let $\xi = \sum_{i=1}^{m_1} \delta_{x_i}$, $\eta = \sum_{j=1}^{m_2} \delta_{y_j}$ and $\zeta = \sum_{k=1}^{n} \delta_{z_k}$ be in $\mathfrak{N}_{\text{fin}}$ and set $m_* = \max\{m_1, m_2\}$. If $n \leq m_*$, we obtain the desired result from the triangle inequality of τ as

$$\begin{aligned}\bar{\tau}(\xi,\eta) &= \frac{1}{m_*^{1/p}}\tau(\xi,\eta)\\ &\le \frac{1}{m_*^{1/p}}\big(\tau(\xi,\zeta)+\tau(\zeta,\eta)\big)\\ &\le \frac{1}{\max\{m_1,n\}^{1/p}}\tau(\xi,\zeta)+\frac{1}{\max\{n,m_2\}^{1/p}}\tau(\zeta,\eta)\\ &\le \bar{\tau}(\xi,\zeta)+\bar{\tau}(\zeta,\eta).\end{aligned}$$

If $n > m_*$, we use a slightly different construction for the extended metric space. Let $\mathcal{X}' = \mathcal{X} \cup \{\aleph, \aleph'\}$ for two different $\aleph, \aleph' \notin \mathcal{X}$. Setting

$$d'(x,y) = \begin{cases} \min\{d(x,y), 2^{1/p}C\} & \text{if } x,y \in \mathcal{X},\\ C & \text{if } \{x,y\}\cap\mathcal{X} \neq \emptyset \text{ and}\\ & \{x,y\}\cap\{\aleph,\aleph'\} \neq \emptyset,\\ 2^{1/p}C & \text{if } \{x,y\} = \{\aleph,\aleph'\},\\ 0 & \text{if } x = y = \aleph\\ & \text{or } x = y = \aleph', \end{cases}$$

we obtain by Lemma A.1 that $(\mathcal{X}', d')$ is again a metric space. Setting $x_i = \aleph$ for $m_1+1 \le i \le n$ and $y_j = \aleph'$ for $m_2+1 \le j \le n$, we may define $\tilde{\xi} = \sum_{i=1}^{n}\delta_{x_i}$ and $\tilde{\eta} = \sum_{j=1}^{n}\delta_{y_j}$. Note that an optimal permutation $\pi_* \in S_{m_*}$ for $\bar{\tau}(\xi,\eta)$ can be extended to an optimal permutation $\tilde{\pi}_* \in S_n$ for $\bar{\tau}(\tilde{\xi},\tilde{\eta})$ by setting $\tilde{\pi}_*(i) = i$ for $m_*+1 \le i \le n$. Furthermore, for any $s, c \ge 0$ with $s \le m_* c$, we have $\frac{1}{m_*}s \le \frac{1}{n}\big(s+(n-m_*)c\big)$. Combining these two facts, we obtain

$$\begin{aligned}\bar{\tau}(\xi,\eta)^p &= \frac{1}{m_*}\sum_{i=1}^{m_*} d'(x_i, y_{\pi_*(i)})^p\\ &\le \frac{1}{n}\left(\sum_{i=1}^{m_*} d'(x_i, y_{\pi_*(i)})^p + (n-m_*)\cdot 2C^p\right)\\ &= \bar{\tau}(\tilde{\xi},\tilde{\eta})^p,\end{aligned}$$

and therefore

$$\bar{\tau}(\xi,\eta) \le \bar{\tau}(\tilde{\xi},\tilde{\eta}) \le \bar{\tau}(\tilde{\xi},\zeta)+\bar{\tau}(\zeta,\tilde{\eta}) = \bar{\tau}(\xi,\zeta)+\bar{\tau}(\zeta,\eta),$$

where the second inequality holds since the cardinalities of all point patterns are equal and the equality holds by two more applications of Theorem 1. □

Proof of Theorem 3 The equivalence for (a) was already used in Diez et al. (2012). We give a quick argument for the sake of completeness. We may assume without loss of generality that, in an admissible path $P = (\xi_0, \ldots, \xi_N)$ for the minimization problem (4),

- only moves from $x \in \xi$ to $y \in \eta$ occur;
- only points $y \in \eta$ are added;
- only points $x \in \xi$ are deleted;

because if any of these conditions were violated, the total cost of the path could only become larger (for the first item we use the triangle inequality for d). The minimization (4) is then equivalent to choosing $l \in \{0, 1, \ldots, \min\{m,n\}\}$ points to be moved from ξ-points with indices $i_1, \ldots, i_l \in [m]$ to η-points with indices $j_1, \ldots, j_l \in [n]$, respectively, at cost $d(x_{i_r}, y_{j_r})$ for each move. The remaining $m-l$ points of ξ are deleted at cost C per deletion, and the remaining $n-l$ points of η are added at cost C per addition. This yields exactly the minimization problem (2).

The equivalence (b) is an immediate consequence of Theorem 1. □

References

Agueh, M., Carlier, G.: Barycenters in the Wasserstein space. SIAM J. Math. Anal. **43**, 904–924 (2011)

Anderes, E., Borgwardt, S., Miller, J.: Discrete Wasserstein barycenters: optimal transport for discrete data. Math. Methods Oper. Res. **84**(2), 389–409 (2016)

Baddeley, A., Rubak, E., Turner, R.: Spatial Point Patterns: Methodology and Applications with R. Chapman and Hall/CRC, Boca Raton (2015)

Bandelt, H.-J., Crama, Y., Spieksma, F.C.R.: Approximation algorithms for multi-dimensional assignment problems with decomposable costs. Discrete Appl. Math. **49**, 25–50 (1994)

Bertsekas, D.P.: The auction algorithm: a distributed relaxation method for the assignment problem. Ann. Oper. Res. **14**, 105–123 (1988)

Błaszczyszyn, B., Haenggi, M., Keeler, P., Mukherjee, S.: Stochastic Geometry Analysis of Cellular Networks. Cambridge University Press, Cambridge (2018)

Borgwardt, S.: An LP-based, strongly polynomial 2-approximation algorithm for sparse Wasserstein barycenters. Preprint (2019). arXiv:1704.05491v5

Borgwardt, S., Patterson, S.: Improved linear programs for discrete barycenters. INFORMS J Optim (2018). arXiv:1803.11313

Chiaraviglio, L., Cuomo, F., Maisto, M., Gigli, A., Lorincz, J., Zhou, Y., Zhao, Z., Qi, C., Zhang, H.: What is the best spatial distribution to model base station density? A deep dive into two European mobile networks. IEEE Access **4**, 1434–1443 (2016)

Chizat, L.: Unbalanced Optimal Transport: Models, Numerical Methods, Applications. Ph.D. thesis, PSL Research University (2017)

Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.-X.: Scaling algorithms for unbalanced optimal transport problems. Math. Comput. **87**(314), 2563–2609 (2018)

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 3rd edn. MIT Press, Cambridge (2009)

Cuturi, M., Doucet, A.: Fast computation of Wasserstein barycenters. In: Xing, E.P., Jebara, T. (eds.) Proceedings of the 31st International Conference on Machine Learning, pp. 685–693 (2014)

del Barrio, E., Cuesta-Albertos, J.A., Matrán, C., Mayo-Íscar, A.: Robust clustering tools based on optimal transportation. Stat. Comput. **29**, 139–160 (2019)

Diez, D.M., Schoenberg, F.P., Woody, C.D.: Algorithms for computing spike time distance and point process prototypes with application to feline neuronal responses to acoustic stimuli. J. Neurosci. Methods **203**(1), 186–192 (2012)

Diggle, P.J.: Statistical Analysis of Spatial and Spatio-Temporal Point Patterns. Chapman and Hall/CRC, Boca Raton (2013)

Dubey, P., Müller, H.-G.: Fréchet analysis of variance for random objects. Preprint **106**(4), 803–821 (2019a)

Dubey, P., Müller, H.-G.: Functional models for time-varying random objects. Preprint (2019b). arXiv:1907.10829

Fréchet, M.: Les éléments aléatoires de nature quelconque dans un espace distancié. Ann. Inst. H. Poincaré **10**, 215–310 (1948)

Hakimi, S.L.: Optimum locations of switching centers and the absolute centers and medians of a graph. Oper. Res. **12**(3), 456–458 (1964)

Koliander, G., Schuhmacher, D., Hlawatsch, F.: Rate-distortion theory of finite point processes. IEEE Trans. Inf. Theory **64**(8), 5832–5861 (2018)

Konstantinoudis, G., Schuhmacher, D., Ammann, R., Diesch, T., Kuehni, C., Spycher, B.D.: Bayesian spatial modelling of childhood cancer incidence in Switzerland using exact point data: a nationwide study during 1985–2015. Preprint (2019). https://www.medrxiv.org/content/early/2019/07/15/19001545

Kuhn, H.W.: The Hungarian method for the assignment problem. Naval Res. Logist. Quart. **2**, 83–97 (1955)

Liero, M., Mielke, A., Savaré, G.: Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. Invent. Math. **211**(3), 969–1117 (2018)

Lin, Z., Müller, H.-G.: Total variation regularized Fréchet regression for metric-space valued data. Preprint (2019). arXiv:1904.09647

Lombardo, L., Opitz, T., Huser, R.: Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. Stoch. Environ. Res Risk Assess. **32**(7), 2179–2198 (2018)

Luenberger, D.G., Ye, Y.: Linear and Nonlinear Programming, third edn. Springer, New York (2008)

Mateu, J., Schoenberg, F.P., Diez, D.M., Gonzáles, J.A., Lu, W.: On measures of dissimilarity between point patterns: classification based on prototypes and multidimensional scaling. Biom. J. **57**(2), 340–358 (2015)

Moradi, M., Mateu, J.: First and second-order characteristics of spatio-temporal point processes on linear networks. J. Comput. Graph. Stat. (2019, to appear)

Moradi, M., Rodriguez-Cortes, F., Mateu, J.: On kernel-based intensity estimation of spatial point patterns on linear networks. J. Comput. Graph. Stat. **27**(2), 302–311 (2018)

Müller, R., Schuhmacher, D.: ttbary: barycenter methods for spatial point patterns. R package version 0.1-1. (2019) https://cran.r-project.org/package=ttbary

Petersen, A., Müller, H.-G.: Fréchet regression for random objects with Euclidean predictors. Ann. Stat. **47**(2), 691–719 (2019)

Peyré, G., Cuturi, M.: Computational optimal transport Foundations and Trends®. Mach. Learn. **11**(5–6), 355–607 (2019)

R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2019)

Rakshit, S., Davies, T., Moradi, M., McSwiggan, G., Nair, G., Mateu, J., Baddeley, A.: Fast kernel smoothing of point patterns on a large network using 2d convolution. Int. Stat. Rev. (2019)

Samartsidis, P., Eickhoff, C.R., Eickhoff, S.B., Wager, T.D., Barrett, L.F., Atzil, S., Johnson, T.D., Nichols, T.E.: Bayesian log-Gaussian Cox process regression: applications to meta-analysis of neuroimaging working memory studies. J. R. Stat. Soc. Ser. C **68**(1), 217–234 (2019)

Schmitz, M.A., Heitz, M., Bonneel, N., Ngole, F., Coeurjolly, D., Cuturi, M., Peyré, G., Starck, J.-L.: Wasserstein dictionary learning: optimal transport-based unsupervised nonlinear dictionary learning. SIAM J. Imaging Sci. **11**(1), 643–678 (2018)

Schoenberg, F.P., Tranbarger, K.E.: Description of earthquake aftershock sequences using prototype point patterns. Environmetrics **19**(3), 271–286 (2008)

Schuhmacher, D.: Stein's method for approximating complex distributions, with a view towards point processes. In: Schmidt, V. (ed.) Stochastic Geometry, Spatial Statistics and Random Fields, Vol. II: Models and Algorithms. Lecture Notes in Mathematics, vol. 2120, pp. 1–30. Springer (2014)

Schuhmacher, D., Vo, B.-T., Vo, B.-N.: A consistent metric for performance evaluation of multi-object filters. IEEE Trans. Signal Process. **56**(8, part 1), 3447–3457 (2008)

Schuhmacher, D., Xia, A.: A new metric between distributions of point processes. Adv. Appl. Probab. **40**(3), 651–672 (2008)

Victor, J.D., Purpura, K.P.: Metric-space analysis of spike trains: theory, algorithms and application. Netw. Comput Neural Syst. **8**, 127–164 (1997)

Weiszfeld, E.: Sur le point pour lequel la somme des distances de *n* points donnés est minimum. Tohoku Math. J. **43**, 355–386 (1937)

APPENDIX B

ANOVA for Data in Metric Spaces, with Applications to Spatial Point Patterns

ANOVA for Data in Metric Spaces, with Applications to Spatial Point Patterns

Raoul Müller[*][†][‡] Dominic Schuhmacher[‡] Jorge Mateu[§][¶]

February 18, 2022

Abstract

We give a review of recent ANOVA-like procedures for testing group differences based on data in a metric space and present a new such procedure. Our statistic is based on the classic Levene's test for detecting differences in dispersion. It uses only pairwise distances of data points and and can be computed quickly and precisely in situations where the computation of barycenters ("generalized means") in the data space is slow, only by approximation or even infeasible. We show the asymptotic normality of our test statistic and present simulation studies for spatial point pattern data, in which we compare the various procedures in a 1-way ANOVA setting. As an application, we perform a 2-way ANOVA on a data set of bubbles in a mineral flotation process.

1 Introduction

Real-world statistical data is often not Euclidean, involving components that are most suitably analyzed in a more complicated space. Examples include spaces of point patterns and more general subsets, trees and more general graphs, functions and images.

In recent years a number of methods have been proposed for analyzing group differences of such data by generalizing classical analysis of variance (ANOVA) ideas to more complex data spaces. Examples include Cuevas et al. (2004) for functional data, Huckemann et al. (2009) for data on Riemannian manifolds and Ramón et al. (2016) for point pattern data. A common feature of the underlying spaces is that there is typically a more or less natural concept of distance between data points available. In addition to the more obvious choices of distances on function spaces and Riemannian manifolds, suitable metrics for tree spaces, graph spaces and point pattern spaces can be found in Billera et al. (2001), Ginestet et al. (2017) and Müller et al. (2020), respectively.

In the present paper we focus on generalized ANOVA-procedures for metric spaces without using any more special structure of the space. There is a number of preceding articles that work in similar generality.

Anderson (2001) proposes to perform ANOVA based on pairwise dissimilarities of observations rather than Euclidean distances between observations and their group means, and introduces the name PERMANOVA for this procedure. While not directly referring to any more abstract spaces than $\mathbb{R}^d$, that article clearly discusses the abstract template of doing non-Euclidean ANOVA without using a centroid object. We discuss this further in Subsection 3.1.

[*] Work supported by DFG RTG 2088.

[†] raoul.mueller@uni-goettingen.de

[‡] Institute for Mathematical Stochastics, University of Göttingen, 37077 Göttingen, Germany.

[§] Work partially funded by Ministry of Science and Innovation with grant PID2019-107392RB-I00, and by grant UJI-B2021-37 from University Jaume I.

[¶] Department of Mathematics, University Jaume I, 12071 Castellón, Spain.

Anderson (2006) proposes multidimensional scaling followed by a Levene's test (using the centroid object in the principal coordinate space) for detecting differences of within-group dispersions (scatter, variability); this is referred to as PERMDISP, see Anderson (2017). Anderson et al. (2017) and Hamidi et al. (2019) correct the PERMANOVA statistic for heteroscedasticity in the unbalanced setting based on the variants of classical ANOVA by Brown–Forsythe and Welch, respectively.

In an independent line of research, Dubey and Müller (2019) design an ANOVA procedure on metric spaces using Fréchet means as centroid objects. They propose to use as statistic the sum of an ANOVA-term and a Levene-term. We discuss this further in Subsection 3.2.

In the present paper we formulate Anderson's PERMANOVA on general metric spaces. We simply refer to the resulting method as Anderson ANOVA, because the use of M (due to the use of $\mathbb{R}^d$ in Anderson's work) seems inappropriate in our context and the use of PER (referring to the fact that a permutation test is performed) does not distinguish it from the other methods used. Rather than pursuing the PERMDISP method mentioned above, we introduce a new test for detecting differences of within-group dispersion based on Levene's procedure and refer to it as L-test. Our test statistic works directly with the pairwise distances between observations without using any kind of group centroid, neither in the original metric space nor in any principal coordinate space. We show that it has an asymptotic χ^2_1-distribution, but we recommend using it with a permutation test just as the other statistics.

We also study the two summands used by Dubey and Müller (2019) as separate test statistics for detecting differences in location and dispersion, respectively. We refer to Table 1 for an overview of the methods discussed.

	location	*dispersion*
pairwise distances	Anderson, Subsec. 3.1	New L-test, Section 4
Fréchet means	Dubey–Müller, Subsec. 3.2	Dubey–Müller, Subsec. 3.2

Table 1: Overview of the non-Euclidean ANOVA methods studied in this paper. Procedures targeting *location* are derived from the classic ANOVA statistic, whereas those targeting *dispersion* are derived from the classic Levene's statistic (ANOVA statistic for "deviations"). The rows distinguish whether computationally a procedure is based on (simple arithmetics of) *pairwise distances* or on a centroid object (here a *Fréchet mean*) in the metric space.

Although the methods described are applicable in general metric spaces, our central goal in undertaking this research was to be able to perform ANOVA for point pattern data, see also the discussion section of Müller et al. (2020). Among all the metric spaces mentioned above, we therefore focus in the later part of the present paper on the space of finite point patterns equipped with the TT-metric from Müller et al. (2020). As in many other spaces, exact Fréchet means can be computed within reasonable time only for (very) small data sets and one typically has to resort to a heuristic algorithm that finds only local minima of the Fréchet functional. We present simulation studies to compare the powers of the four tests across various situations and to understand the quality of approximation by the limiting χ^2_1-distribution from a practical point of view. We also present an application of a 2-way ANOVA on a data set of bubbles in a mineral flotation process.

The plan of the paper is as follows: Section 2 contains a brief reminder of central aspects of classical ANOVA including Levene's test. In Section 3 we give a rather detailed presentation of Anderson ANOVA in metric spaces and the two summands proposed by Dubey and Müller. In Section 4 we introduce our new L-statistic, discuss its relation to the other methods and the original Levene's test, and show its asymptotic distribution. Section 5 is a short overview of the metric space of point patterns. In Sections 6 and 7 we present the simulation studies and the real-world data example. The paper ends with some further conclusions in Section 8.

2 Classic ANOVA

For self-containedness and easy reference we briefly remind the reader of some facts and formulae in the context of the classical ANOVA going back to Fisher (1925). Details can be found in Scheffé (1967).

One-Way ANOVA

Given independent observations $x_{ij} \in \mathbb{R}$, $1 \le j \le n_i$, $1 \le i \le k$, from k potentially different distributions $P_1, \dots, P_k$, we do the following sum-of-squares decomposition

$$\mathrm{TSS} = \mathrm{MSS} + \mathrm{RSS},$$

where

$$\begin{aligned} \mathrm{TSS} &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{\cdot\cdot})^2 && \textit{(total sum of squares)} \\ \mathrm{MSS} &= \sum_{i=1}^{k} n_i (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 && \textit{(model sum of squares)} \\ \mathrm{RSS} &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot})^2 && \textit{(residual sum of squares).} \end{aligned}$$

Here $\bar{x}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$ denotes the i-th group mean and $\bar{x}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}$ denotes the overall mean. We write $n = \sum_{i=1}^{k} n_i$ for the total number of observations.

Assume for now that the group distributions P_i are Gaussian with the same variance. Under the null hypothesis that also the means are the same (hence all data comes from the same normal distribution), it is well-known that the ANOVA statistics

$$F = \frac{n-k}{k-1} \frac{\mathrm{MSS}}{\mathrm{RSS}}, \tag{1}$$

describing the ratio between the variability explained by the model and the total variability in the data, is F-distributed with $k-1$ and $n-k$ degrees of freedom. Since $T \sim F(d_1, d_2)$ implies $d_1 T \xrightarrow{\mathcal{D}} \chi^2_{d_1}$ as $d_2 \to \infty$, we obtain

$$(k-1)F \xrightarrow{\mathcal{D}} \chi^2_{k-1} \quad \text{as } n \to \infty. \tag{2}$$

The *asymptotic* result remains true even if $P_1 = P_2 = \ldots = P_k$ is non-Gaussian, but has second moments and there are $\lambda_1, \dots, \lambda_k > 0$ such that the ratios of group sizes satisfy $\frac{n_i}{n} \to \lambda_i$, see e.g. Wooldridge (2010), Section 3.6.2.

Remark 1. *Strictly speaking ANOVA techniques are designed for inference within a linear model of different group means plus errors. Based on an error distribution P with mean zero, one considers the model equations*

$$x_{ij} = \mu_i + \varepsilon_{ij}, \quad 1 \le j \le n_i, 1 \le i \le k,$$

where $\mu_i \in \mathbb{R}$ are the different group means and ε_{ij} are i.i.d. P-distributed error terms. In terms of the group distributions above this means that $P_i = P * \delta_{\mu_i}$, i.e. P_i is obtained by shifting P by μ_i. Note that the asymptotic χ^2_{k-1}-test does not need this assumption since in any case the null hypothesis just correspond to having k times the same distribution. At the same time we cannot expect this test to achieve high power against all alternatives that have substantially different group distributions (see also the paragraph on Levene's test below). We will take up this point when discussing ANOVA on metric spaces, where typically "shifting the distribution" is meaningless (but may have an intuitive counterpart).

Two-Way ANOVA

As soon as more than one grouping factor is involved, important design decisions come into play, such as if factors are (partially) nested or if we allow for interaction terms between several factors on the same level. ANOVA has a long-standing history with many different designs. As an example which is pursued further in later sections we remind the reader of the balanced two-way ANOVA (two main factors, with interaction terms, same number $\tilde{n}$ of observations for each factor combination).

Given independent observations $x_{i_1 i_2 j} \in \mathbb{R}$, $1 \leq j \leq \tilde{n}$, $1 \leq i_1 \leq k_1$, $1 \leq i_2 \leq k_2$ from groups obtained by crossing a Factor a with k_1 levels and a Factor b with k_2 levels (with $n_{i_1 i_2} := \tilde{n}$ observations for each combination), we can perform a finer sum-of-squares decomposition

$$\mathrm{TSS} = \mathrm{SSa} + \mathrm{SSb} + \mathrm{SSi} + \mathrm{RSS},$$

splitting up the model sum of squares into sums of squares for the individual factors and an interaction sum of squares. In formulae:

$$\begin{aligned}
\mathrm{TSS} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \sum_{j=1}^{\tilde{n}} (x_{i_1 i_2 j} - \bar{x}_{...})^2 \\
\mathrm{RSS} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \sum_{j=1}^{\tilde{n}} (x_{i_1 i_2 j} - \bar{x}_{i_1 i_2 \cdot})^2 \\
\mathrm{SSa} &= \sum_{i_1=1}^{k_1} k_2 \tilde{n} (\bar{x}_{i_1 \cdot\cdot} - \bar{x}_{...})^2 \\
\mathrm{SSb} &= \sum_{i_2=1}^{k_2} k_1 \tilde{n} (\bar{x}_{\cdot i_2 \cdot} - \bar{x}_{...})^2 \\
\mathrm{SSi} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \tilde{n} (\bar{x}_{i_1 i_2 \cdot} - \bar{x}_{i_1 \cdot\cdot} - \bar{x}_{\cdot i_2 \cdot} + \bar{x}_{...})^2,
\end{aligned}$$

where the various means are taken over the dot components while keeping the given indices fixed. Set $n = \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} n_{i_1 i_2} = k_1 k_2 \tilde{n}$.

In addition to performing an omnibus test for group differences as for one-way ANOVA, we may then test for effects of Factor a and b separately, as well as for an interaction effect. The corresponding statistics are

$$Fa = \frac{n - k_1 k_2}{k_1 - 1} \frac{\mathrm{SSa}}{\mathrm{RSS}}, \quad Fb = \frac{n - k_1 k_2}{k_2 - 1} \frac{\mathrm{SSb}}{\mathrm{RSS}}, \quad Fi = \frac{n - k_1 k_2}{(k_1 - 1)(k_2 - 1)} \frac{\mathrm{SSi}}{\mathrm{RSS}}.$$

If the observations come from Gaussian distributions with equal variances, each of the three statistics is F-distributed again under the corresponding null hypothesis that different levels of the factor or interaction to be tested do not lead to different shifts in mean. The degrees of freedom can be read from the denominator and the numerator, respectively, of the first ratio in each statistic.

Levene's Test

The test first proposed in Levene (1960) was originally developed as a preliminary test to check for equal variances *before* applying the basic ANOVA F-test in the Gaussian setting. This was important, as it was well-known at the time that for the goal of inference about differences in the means of the various groups (see Remark 1), the size of the F-test can depart substantially from its nominal size if group variances are not equal.

Levene (1960) proposed to use as test statistic the usual ANOVA statistic, but to replace the observations x_{ij} by the absolute differences from their group means $z_{ij} = |x_{ij} - \bar{x}_{i\cdot}|$, i.e.

$$\widetilde{F} = \frac{n-k}{k-1} \cdot \frac{\sum_{i=1}^{k} n_i (\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot})^2}. \tag{3}$$

If the observations are independently sampled from the same Gaussian distributions, it is plausible that $\widetilde{F}$ is still approximately F-distributed, because the dependence between the z_{ij} is small even at moderate group sizes. This was confirmed by simulation in Levene (1960). Brown and Forsythe (1974a) present a larger simulation experiment suggesting that replacing the $\bar{x}_{i\cdot}$ in the definition of z_{ij} by a trimmed mean or median leads to a more robust test for non-Gaussian data.

Current best practice suggests to perform a Welch-modified ANOVA directly if the assumption of equal variance is unclear as it results only in a small loss of power in the case where the variances are indeed equal. We refer to Gastwirth et al. (2009) for a comprehensive presentation on Levene's test including this question and many further developments.

Levene's test and its variants remain highly important today as differences in variances (or some other measure of dispersion) are often in the center of attention in their own rights. In the rest of the paper we present tests on differences in "location" of groups and differences in "dispersion" of groups, both based on inter-point distances in a metric space. Our goal is to combine one of either kind in order to detect group differences in some universality.

3 Non-Euclidean ANOVA

In this and the next sections we assume that our data lies in a general metric space $(\mathcal{X}, d)$. We present existing methods of testing for group differences based on ANOVA-like ideas. For the presentation we focus on generalizations of 1-way ANOVA, but provide further information on which methods can easily be extended to more complex designs. We always assume having $n = \sum_{i=1}^{k} n_i$ independent observations $x_{ij} \in \mathcal{X}$, $1 \leq j \leq n_i$, $1 \leq i \leq k$ from k potentially different distributions $P_1, \ldots, P_k$ on $\mathcal{X}$ (with Borel σ-algebra).

3.1 Anderson ANOVA

Anderson (2001) argues, in the context of data sets in ecology, that traditional multivariate analogues of ANOVA are too stringent in their assumptions. These are typically based on similar statistics as (1), but with absolute values replaced by Euclidean norms, see e.g. Mardia et al. (1979) Section 12.3. We may avoid the use of means of observations by writing TSS − RSS instead of MSS and replacing the sums of squared deviations from the mean with the help of the formula

$$\sum_{j=1}^{m} \|y_j - \overline{y}\|^2 = \frac{1}{2m} \sum_{j_1,j_2=1}^{m} \|y_{j_1} - y_{j_2}\|^2 = \frac{1}{m} \sum_{j_1,j_2=1}^{m,<} \|y_{j_1} - y_{j_2}\|^2,$$

where we indicate by "<" in the summation bound that the sum is to be taken over strictly ordered summands only, here $j_1 < j_2$. Anderson proposes to replace the pairwise Euclidean distances by more general dissimilarities between observations and performs a permutation test.

In our context we simply use the pairwise distances in the metric space. Thus

$$
\begin{aligned}
\mathrm{TSS} &= \frac{1}{n}\Big(\sum_{i_1,i_2=1}^{k,<}\sum_{j_1=1}^{n_{i_1}}\sum_{j_2=1}^{n_{i_2}} d^2(x_{i_1j_1},x_{i_2j_2}) + \sum_{i=1}^{k}\sum_{j_1,j_2=1}^{n_i,<} d^2(x_{ij_1},x_{ij_2})\Big)\\
\mathrm{RSS} &= \sum_{i=1}^{k}\frac{1}{n_i}\sum_{j_1,j_2=1}^{n_i,<} d^2(x_{ij_1},x_{ij_2})\\
\mathrm{MSS} &= \mathrm{TSS} - \mathrm{RSS}
\end{aligned}
$$

and the final Anderson ANOVA statistic becomes

$$
F_{\mathrm{A}} = \frac{n-k}{k-1}\frac{\mathrm{MSS}}{\mathrm{RSS}}.
$$

It has been noted in various places that this statistic may suffer from type I error inflation (in terms of a null hypothesis of equal *means* in Euclidean space) and substantial loss of power in the unbalanced setting if the groups are heteroscedastic; see e.g.Alekseyenko (2016). Anderson et al. (2017) and Hamidi et al. (2019) propose improvements based on the classical ANOVA variants by Brown and Forsythe (1974b) and Welch (1951), respectively. In the former, the F-statistic is replaced by

$$
F_{\mathrm{BF}} = \frac{\mathrm{MSS}}{\sum_{i=1}^{k}(1-\frac{n_i}{n})\frac{1}{n_i(n_i-1)}\sum_{j_1,j_2=1}^{n_i,<} d^2(x_{ij_1},x_{ij_2})}.
$$

For the simulation studies in Section 6 we concentrate on the balanced setting, for which Anderson F_A performs typically well even in presence of heteroscedacity. We therefore do not discuss these improvements further, which in the balanced setting do not change the statistic.

3.2 Fréchet ANOVA

Dubey and Müller (2019) introduce ANOVA-like terms that use distances in the metric d to Fréchet means rather than absolute differences to averages. For observation $y_1, \dots y_m \in \mathcal{X}$ the Fréchet mean is defined as

$$
\bar{y} = \operatorname*{argmin}_{z\in\mathcal{X}} \sum_{i=1}^{m} d^2(y_i, z). \tag{4}
$$

One of the assumptions in Dubey and Müller (2019) is that all Fréchet means considered exist and are unique. For our usual set of observations we denote by $\bar{x}_{i\cdot}$ the Fréchet mean of $x_{i1}, \dots, x_{in_i}$, $i = 1, \dots, k$ and by $\bar{x}_{\cdot\cdot}$ the Fréchet mean of all observations. Following the notation in Dubey and Müller (2019), we write the Fréchet variance for the i-th group and the total Fréchet variance as

$$
\hat{V}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} d^2(x_{ij}, \bar{x}_{i\cdot}) \quad \text{and} \quad \hat{V}_p = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} d^2(x_{ij}, \bar{x}_{\cdot\cdot}),
$$

respectively. While $\hat{V}_i$ is the mean of $d^2(x_{ij}, \bar{x}_{i\cdot})$, $j = 1, \dots, n_i$, we also require the corresponding variance

$$
\hat{\sigma}_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i} d^4(x_{ij}, \bar{x}_{i\cdot}) - \hat{V}_i^2.
$$

Setting $\lambda_i = \frac{n_i}{n}$ one finally obtains

$$U_n = \sum_{i_1, i_2 = 1}^{k, <} \frac{\lambda_{i_1} \lambda_{i_2}}{\hat{\sigma}_{i_1}^2 \hat{\sigma}_{i_2}^2} (\hat{V}_{i_1} - \hat{V}_{i_2})^2$$

$$F_n = \hat{V}_p - \sum_{i=1}^{k} \lambda_i \hat{V}_i$$

$$T = \frac{nU_n}{\sum_{i=1}^{k} \frac{\lambda_i}{\hat{\sigma}_i^2}} + \frac{nF_n^2}{\sum_{i=1}^{k} \lambda_i^2 \hat{\sigma}_i^2} =: T_L + T_F.$$

In the Euclidean setting of Section 2 the term F_n is equal to $\frac{1}{n}(\text{TSS} - \text{RSS})$ and the denominator of T_F is then an estimator for the variance of $\frac{1}{n}\text{RSS}$, so that T_F has close ties to the ANOVA F-statistic. The unweighted summands $(\hat{V}_{i_1} - \hat{V}_{i_2})^2$ of U_n are similar in spirit to the terms $(\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot})^2$ from the definition of Levene's statistic, and in fact it appears that in the Euclidean case T_L corresponds exactly to a simpler variant of Welch's ANOVA applied to $d^2(x_{ij}, \bar{x}_{i\cdot})$, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$; see the computation in Formulae (8)–(16) in Hamidi et al. (2019). Thus T_L has close ties to Levene's statistic.

Dubey and Müller show under a list of conditions pertaining to existence and uniqueness of theoretical and empirical Fréchet means and the complexity of the metric space (in terms of entropy integrals) that

$$\frac{nU_n}{\sum_{i=1}^{k} \frac{\lambda_i}{\hat{\sigma}_i^2}} \xrightarrow{\mathcal{D}} \chi_{k-1}^2 \quad \text{and} \quad \frac{nF_n^2}{\sum_{i=1}^{k} \lambda_i^2 \hat{\sigma}_i^2} \xrightarrow{\mathcal{D}} 0 \quad \text{as } n \to \infty.$$

The authors advocate the simple addition of the two terms in order to obtain a single test statistic T, maybe with weights if there is prior information available whether to rather look out for inequality of Fréchet means or of Fréchet variances. However, due to the unbalanced convergence of the two terms and the fact that the reason for the concrete normalization (especially) of T_F remains a bit inscrutable to us, we prefer to analyze the two summands separately in Section 6.

4 A New Non-Euclidean Method of Levene Type

What appears to be missing is a test for detecting differences of within-group dispersion that is based directly on pairwise distances between observations in the metric space. The idea of the PERMDISP-test mentioned in the introduction, i.e. performing multidimensional scaling and applying Levene's test in the principal coordinate space, is to some extent applicable here. However, it is rather an indirect method and it is methodologically not on the same level as the Anderson F_A. Indeed multidimensional scaling can be applied in combination with *any* Euclidean procedure, so the PERMDISP-method should be rather paired up with the analog method of multidimensional scaling plus applying Euclidean (M)ANOVA. What is more, it contains an unwelcome tuning parameter, the number of principal coordinates, which is not easy to choose, but may be crucial. Instead we propose the following test of Levene type for data in a metric space.

4.1 Form and Properties

We assume the same setup as in the previous section, i.e. there are $n = \sum_{i=1}^{k} n_i$ independent observations $x_{ij} \in \mathcal{X}$, $1 \leq j \leq n_i$, $1 \leq i \leq k$ from k potentially different distributions $P_1, \ldots, P_k$ on $\mathcal{X}$. Set $N_i = \binom{n_i}{2}$ and $N = \sum_{i=1}^{k} N_i$. As a surrogate for the individual deviation terms z_{ij} from Levene's statistic (3), which in a general metric space would require the use of a Fréchet or similar mean, we use $d_{i,\{j_1,j_2\}} := \frac{1}{2}d(x_{ij_1}, x_{ij_2})$. To simplify the notation, we enumerate the two-element subsets of $\{1,\ldots,n_i\}$ by $j = 1,\ldots,N_i$ and use d_{ij} rather than $d_{i,\{j_1,j_2\}}$ for the j-th half-distance in the i-th group.

In a first step we assume that $n_1 = \ldots = n_k$ (balanced case) and emulate the statistics (3) by setting

$$L := \frac{N-k}{k-1} \frac{\sum_{i=1}^{k} n_i(\bar{d}_{i\cdot} - \bar{d}_{\cdot\cdot})^2}{\sum_{i=1}^{k}\sum_{j=1}^{N_i}(d_{ij} - \bar{d}_{i\cdot})^2} \tag{5}$$

where

$$\bar{d}_{i\cdot} = \frac{1}{N_i}\sum_{j=1}^{N_i} d_{ij} \quad \text{and} \quad \bar{d}_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} d_{ij}$$

denote the i-th group mean and the overall mean over pairwise distances, respectively.

Typographically the main fractions of Equations (5) and (3) are very similar, but the way they use the data x_{ij} is quite different in that we replace $z_{ij} = |x_{ij} - \bar{x}_{i\cdot}|$, $1 \leq j \leq n_i$ by $d_{i,\{j_1,j_2\}} = \frac{1}{2}d(x_{ij_1}, x_{ij_2})$, $1 \leq j_1 < j_2 \leq n_i$. Note that we keep n_i in the numerator rather than replacing it by N_i, which might have seemed more natural at first glance. The reason is the substantial dependence of the random variables $d_{i,\{j_1,j_2\}}$ (as opposed to the less substantial dependence between the z_{ij}) for each i, which implies that n_i, not N_i, is the correct scaling factor; see Subsection 4.2. Note further that, for the same reason, the main denominator is not the most natural choice here, but it is convenient since it keeps the statistic similar to the original Levene statistic, is fast to compute and empirically performs no worse than the more natural choice discussed in Subsection 4.2.

There are various ways how one might generalize (5) to general group sizes. We propose using

$$L := \frac{N-k}{k-1} \frac{\frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_i n_j(\bar{d}_{i\cdot} - \bar{d}_{j\cdot})^2}{\sum_{i=1}^{k}\sum_{j=1}^{N_i}(d_{ij} - \bar{d}_{i\cdot})^2}. \tag{6}$$

Direct computation shows that Equations (6) and (5) agree in the balanced case, but not in general; see Remark 8. The statistic (6) performs well in several respects: it allows for an asymptotic distribution (χ^2_{k-1} up to a deterministic factor, see Corollary 3), is still fast to compute and shows a reasonable performance for unequal group sizes, though it may well be that a more judicious scaling that takes more proper care of different group sizes would be superior.

We briefly come back to this last point in Section 6, but do not go much deeper in the present paper because based on additional considerations, both theoretical and from simulation studies, we do not see any clear improvements when choosing different normalizations.

In spite of the limit distribution which we compute in the next section, we recommend performing a permutation test as for the other statistics considered. For this we permute the observations, not only their distances, i.e. new permutations use distances that are potentially different from the pairwise within-group distances of the original data. As a consequence not only the RSS changes with permutations, but also the TSS.

It is easy enough to generalize the construction of the above test statistic to more complex experimental designs. As an example we take up the balanced two-way ANOVA from Section 2 and form the corresponding Levene-type statistics for $(\mathcal{X}, d)$. For the specific statistics see Section 7.1.

4.2 Limit Distribution

In this subsection we derive asymptotic distributions for the statistic L from (6) and for the related statistic

$$\widetilde{L} := \frac{N^* - k}{k-1} \frac{\frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_i n_j (\bar{d}_{i\cdot} - \bar{d}_{j\cdot})^2}{4\,T_n}, \tag{7}$$

where $N^* = \sum_{i=1}^{k} n_i(n_i-1)^2$ and

$$T_n = \sum_{i=1}^{k} \sum_{\substack{j_1,j_2,j_3=1 \\ j_1 \notin \{j_2,j_3\}}}^{n_i} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)\big(d_{i,\{j_1,j_3\}} - \bar{d}_{i\cdot}\big). \tag{8}$$

The previous formula makes it necessary to use the more complicated notation $d_{i,\{j_1,j_2\}} = \frac{1}{2}d(x_{ij_1}, x_{ij_2})$ from the beginning of Subsection 4.1. Note that $\frac{1}{N^*-k}T_n$ is a natural group based estimator of $\operatorname{Cov}\big(\frac{1}{2}d(X_1,X_2), \frac{1}{2}d(X_1,X_3)\big)$, where X_1, X_2, X_3 are three independent random variables sampled from the distribution of the group. The normalization by $N^* - k$ rather than N^* is simply modeled after the bias correcting term for independent data points.

In spite of the ANOVA-like construction, we cannot use the asymptotic theory for ANOVA directly, because the distances $d_{i,\{j_1,j_2\}}$, our "data", stem from dependent random variables for each i. This dependence is taken into account by using the factor $\frac{n_i n_j}{n}$ rather than N_i or N_j in the numerator and by normalizing with $\frac{1}{N^*-k}4\,T_n$ in (7), which then still allows to obtain the asymptotic χ^2_{k-1}-distribution for $(k-1)\widetilde{L}$. In contrast $(k-1)L$ converges "only" towards a multiple of χ^2_{k-1} that depends on parameters of the group distribution.

Theorem 2. *Assume that the Borel σ-algebra for $(\mathcal{X}, d)$ is countably generated. In the usual 1-way setup of Subsection 4.1 assume that $P_1 = \ldots = P_k = P$ for a distribution P that is not a Dirac distribution and satisfies $\int_{\mathcal{X}}\int_{\mathcal{X}} d^2(x,y)\,P(dx)\,P(dy) < \infty$. Suppose that there are $\lambda_i > 0$ such that $n_i/n \to \lambda_i$ for every i as $n \to \infty$. Then we have*

$$(k-1)\,\widetilde{L} \xrightarrow{\mathcal{D}} \chi^2_{k-1} \quad \text{as } n \to \infty.$$

Corollary 3. *Under the conditions of Theorem 2, we obtain*

$$(k-1)\,L \xrightarrow{\mathcal{D}} \frac{4\gamma^2}{\sigma^2}\chi^2_{k-1} \quad \text{as } n \to \infty,$$

where with independent $X, Y, Z \sim P$ we have

$$\gamma^2 = \operatorname{Cov}(d(X,Y), d(X,Z));$$
$$\sigma^2 = \operatorname{Var}(d(X,Y)).$$

Proof of Theorem 2. Under the null hypothesis our data is generated by independent $\mathcal{X}$-valued random elements $X_{ij} \sim P$, $1 \le j \le n_i$, $1 \le i \le k$, and the distances $d_{i,\{j_1,j_2\}}$ are realizations of the random variables $\frac{1}{2}d(X_{ij_1}, X_{ij_2})$, $1 \le j_1 < j_2 \le n_i$, $1 \le i \le k$. Under the conditions on P we have asymptotic normality of the U-statistics

$$U_i = U_i^{(n)} = \binom{n_i}{2}^{-1} \sum_{j_1,j_2=1}^{n_i,<} \tfrac{1}{2}d(X_{ij_1}, X_{ij_2}), \quad i = 1, \ldots, k \tag{9}$$

by a straightforward generalization of Hoeffding's theorem to random elements in $\mathcal{X}$, see Theorem 5 in the appendix. More precisely, we have with $X, Y, Z \sim P$ independent that

$$\sqrt{n_i}\big(U_i - \tfrac{1}{2}\mathbb{E}d(X,Y)\big) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\gamma^2) \quad \text{as } n_i \to \infty, \tag{10}$$

where $\gamma^2 = \mathrm{Cov}(d(X,Y), d(X,Z)) = \mathrm{Var}(\mathbb{E}(d(X,Y) \mid X)) = 4\gamma_h^2$ in the notation of the appendix with $h = \frac{1}{2}d$. In view of the 1-way ANOVA construction, on which L is based, we define the "design matrix" $D = D_n \in \mathbb{R}^{n\times k}$ by

$$D' := \begin{pmatrix} 1 & \dots & 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \dots & 1 & \cdots & 0 & \cdots & 0 \\ & \vdots & & & & & \ddots & & \vdots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{k\times n}, \tag{11}$$

where the i-th row has exactly n_i ones, and the "contrast matrix"

$$C := \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & & 0 & -1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(k-1)\times k}. \tag{12}$$

Setting $\Delta = \lim_{n\to\infty} \frac{1}{n} D_n' D_n = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, we obtain with $U = U^{(n)} = (U_1, \dots, U_k)'$ by independence of the components and $n_i \to \infty$ as $n \to \infty$ (since $\lambda_i > 0$) that

$$Z_n := \gamma^{-1} \sqrt{n}\, \Delta^{1/2} \big(U - \mathbb{E}U\big) \xrightarrow{\mathcal{D}} \mathcal{N}_k(0, I_k) \quad \text{as } n \to \infty. \tag{13}$$

Setting $\nu = (n_1, \dots, n_k)'$, we may further compute $C'(C(D'D)^{-1}C')^{-1}C = D'D - \frac{1}{n}\nu\nu'$ (see Lemma 7 in the Appendix for the calculation) and therefore

$$\widetilde{L} = \frac{N^* - k}{k-1} \frac{U'C'(C(D_n'D_n)^{-1}C')^{-1}CU}{4\,T_n}. \tag{14}$$

Since $\mathbb{E}U = \frac{1}{2}\mathbb{E}d(X,Y) \cdot \mathbb{1} \in \mathbb{R}^k$ and $C \cdot \mathbb{1} = 0$, we obtain

$$(k-1)\widetilde{L} = \gamma^2 \frac{Z_n' \big(\frac{1}{n} W_n\big) Z_n}{\frac{4}{N^*-k} T_n},$$

where $W_n := \Delta^{-1/2} C'(C(D_n'D_n)^{-1}C')^{-1} C \Delta^{-1/2}$. Note that

$$W := \lim_{n\to\infty} \frac{1}{n} W_n = \Delta^{-1/2} C'(C\Delta^{-1}C')^{-1} C \Delta^{-1/2}$$

is a symmetric and idempotent matrix of rank $k-1$, and therefore $Z'WZ \sim \chi^2_{k-1}$ for $Z \sim \mathcal{N}_k(0, I_k)$ by Lemma 9 from the Appendix. Using (13) it is straightforward to show with the help of the continuous mapping theorem that

$$Z_n' \big(\tfrac{1}{n} W_n\big) Z_n \xrightarrow{\mathcal{D}} \chi^2_{k-1}.$$

So it suffices to show that $\frac{1}{N^*-k} T_n \xrightarrow{p} \gamma^2_{d/2}$. For this we note that the normalized inner sum of (8) satisfies

$$\begin{aligned} &\frac{1}{n_i(n_i-1)^2} \sum_{\substack{j_1,j_2,j_3=1 \\ j_1 \notin \{j_2,j_3\}}}^{n_i} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)\big(d_{i,\{j_1,j_3\}} - \bar{d}_{i\cdot}\big) \\ &\quad = \underbrace{\frac{n_i(n_i-1)(n_i-2)}{n_i(n_i-1)^2}}_{\longrightarrow 1} \underbrace{\frac{1}{n_i(n_i-1)(n_i-2)} \sum_{j_1,j_2,j_3=1}^{n_i,\neq} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)\big(d_{i,\{j_1,j_3\}} - \bar{d}_{i\cdot}\big)}_{\longrightarrow \mathrm{Cov}(\frac{1}{2}d(X_1,X_2), \frac{1}{2}d(X_1,X_3)) = \gamma^2_{d/2}} \qquad (15) \\ &\qquad + \underbrace{\frac{1}{n_i - 1}}_{\longrightarrow 0} \underbrace{\frac{1}{n_i(n_i-1)} \sum_{j_1,j_2}^{n_i,\neq} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)^2}_{\longrightarrow \mathrm{Var}(\frac{1}{2}d(X_1,X_2)) = \sigma^2_{d/2}}, \end{aligned}$$

where convergence of the averages is almost surely and follows after expansion of the products by the strong law of large numbers for U-statistics using the prerequisite $\mathbb{E}(d(X_1, X_2)^2) < \infty$; see Hoeffding (1961).

Thus for the total term

$$\frac{1}{N^* - k} T_n = \frac{1}{N^* - k} \sum_{i=1}^{k} n_i(n_i - 1)^2 \cdot \frac{1}{n_i(n_i - 1)^2} \sum_{\substack{j_1, j_2, j_3 = 1 \\ j_1 \notin \{j_2, j_3\}}}^{n_i} \big(d_{i,\{j_1,j_2\}} - \bar{d}_{i\cdot}\big)\big(d_{i,\{j_1,j_3\}} - \bar{d}_{i\cdot}\big) \longrightarrow \gamma^2_{d/2}.$$

□

Proof of Corollary 3. This follows from Theorem 2 because

$$L = 4 \frac{\frac{1}{N^* - k} T_n}{\frac{1}{N-k} \sum_{i=1}^{k} \sum_{j=1}^{N_i} (d_{ij} - \bar{d}_{i\cdot})^2} \widetilde{L},$$

where the numerator is a consistent estimator of $\gamma^2/4$ and the denominator is a consistent estimator of $\sigma^2/4$, see (15). □

5 Metric Space of Finite Point Patterns

In Sections 6 and 7 we apply the four statistics from Table 1 for the space of finite point patterns equipped with the metric introduced in Müller et al. (2020). For self-containedness we give a short summary of the relevant concepts and results, referring to the paper as MSM20.

For $n \in \mathbb{Z}_+$ write $[n] = \{1, 2, \ldots, n\}$ (including $[0] = \emptyset$). Denote by $\mathfrak{N}_{\text{fin}}$ the space of finite multisets on a complete separable metric space $(\mathcal{X}, \varrho)$. We refer to the elements $\xi = \{x_1, x_2, \ldots, x_n\} \in \mathfrak{N}_{\text{fin}}$ as point patterns, where $n \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ and $x_i \in \mathcal{X}$, $i \in [n]$. Note that $x_i = x_j$ for $i \neq j$ is allowed and that the point patterns can be identified with the counting measure $\sum_{i=1}^{n} \delta_{x_i}$, which is often helpful for theoretical considerations. We write $|\xi|$ to denote the total number of points in the pattern ξ.

Definition 4 (Definition 1 of MSM20)**.** *Let $C > 0$ and $p \geq 1$ be two parameters, referred to as* penalty *and* order*, respectively.*
For $\xi = \{x_1, \ldots, x_m\}, \eta = \{y_1, \ldots, y_n\} \in \mathfrak{N}_{\text{fin}}$ define the transport-transform (TT) metric *by*

$$d_{\text{TT}}(\xi, \eta) = d_{\text{TT}}^{(C,p)}(\xi, \eta) = \left(\min \left((m + n - 2l) C^p + \sum_{r=1}^{l} \varrho(x_{i_r}, y_{j_r})^p \right) \right)^{1/p}, \qquad (16)$$

where the minimum is taken over equal numbers of pairwise different indices $i_1, \ldots, i_l$ in $[m]$ and $j_1, \ldots, j_l$ in $[n]$, i.e. over the set

$$\begin{aligned} S(m,n) = \big\{ &(i_1, \ldots, i_l; j_1, \ldots, j_l)\,;\ l \in \{0, 1, \ldots, \min\{m, n\}\}, \\ &i_1, \ldots, i_l \in [m] \text{ pairwise different}, j_1, \ldots, j_l \in [n] \text{ pairwise different} \big\}. \end{aligned}$$

The distance $d_{\text{TT}}(\xi, \eta)$ can be computed by filling up the smaller point pattern with dummy points located at distance C until it has the same cardinality n as the larger point pattern and then solving a standard assignment problem with cost $\min\{d(x,y), 2^{1/p}C\}$ between points x, y (MSM20, Theorem 1). The classical worst-case complexity of this is $O(n^3)$ (MSM20, Remark 1), which can be somewhat improved to order $n^{2.5}$ up to polylogarithmic factors (Lee and Sidford, 2014). Practical computation times for well over $n = 1000$ points are less than one second (R package ttbary, Müller and Schuhmacher, 2021, using the auction algorithm from Bertsekas, 1988).

The TT-metric can be interpreted as an unbalanced Wasserstein metric (Remark 3). Computing Fréchet means in Wasserstein spaces is a topic of active research; see e.g. Borgwardt and Patterson (2020), Borgwardt and Patterson (2021), Heinemann et al. (2021) and references therein for recent developments the space of discrete measures. In our context an additional increase in difficulty comes from the constraint that the result must be a discrete measure with integer cardinality. In MSM20 we therefore apply an alternating heuristics to obtain local minima of the Fréchet functional in (4). The resulting "pseudo-barycenters" are obtained much faster and appear to be of good quality (consistent objective function values and results conform with intuition), but are by no means perfect and still require considerable computation time for hundreds of patterns with hundred of points (Table 1–4 in MSM20).

A related metric that we take up in Section 7 is the relative TT-metric defined as

$$d_{\mathrm{RTT}}(\xi,\eta) = d_{\mathrm{RTT}}^{(C,p)}(\xi,\eta) = \frac{1}{\max\{|\xi|,|\eta|\}^{1/p}} d_{\mathrm{TT}}^{(C,p)}(\xi,\eta). \tag{17}$$

This metric is in a sense more robust to individual outliers if there are many points. In particular note that $d_{\mathrm{RTT}}(\xi_N, \xi_N \cup \zeta) \to 0$ as $N \to \infty$ if $|\xi_N| \to \infty$ and ζ is a fixed point pattern.

In view of the conditions for Theorem 2, completeness and separability are inherited from $(\mathcal{X}, \varrho)$ to $(\mathfrak{N}_{\mathrm{fin}}, d_{\mathrm{TT}})$ and $(\mathfrak{N}_{\mathrm{fin}}, d_{\mathrm{RTT}})$. This is straightforward to see after checking that $d_{\mathrm{RTT}}(\xi_N, \xi) \to 0$ iff $d_{\mathrm{TT}}(\xi_N, \xi) \to 0$ iff $|\xi_N| \to |\xi|$ and each point x of ξ is approximated by exactly one point of ξ_N (if x is a multipoint of cardinality k this means that there is a total of exactly k points in ξ_N, possibly forming multipoints of their own, that converge towards x). The condition $\int_{\mathcal{X}} \int_{\mathcal{X}} d(x,y)\, P(dx)\, P(dy) < \infty$ is always satisfied for d_{RTT} because it is bounded by C. Since $d_{\mathrm{TT}}(\xi,\eta) \leq C \max\{|\xi|,|\eta|\}^{1/p}$ it is satisfied for d_{TT} if $\Xi \sim P$ satisfies $\mathbb{E}|\Xi|^{2/p} < \infty$, which is for example the case for all point process distributions considered in Section 6.

For the simulation study in the next section it is helpful to understand some basic probability measures on $\mathfrak{N}_{\mathrm{fin}}$. Suppose that $\mathcal{R} \subset \mathbb{R}^d$ is compact (in the next section we only use a unit square in $\mathbb{R}^2$). A random element in the metric space $(\mathfrak{N}_{\mathrm{fin}}, d_{\mathrm{RTT}})$, equipped with its Borel σ-algebra is called a *point process*, i.e. a point process is a measurable map from a probability space to $\mathfrak{N}_{\mathrm{fin}}$. The Borel σ-algebra coincides with the smallest σ-algebra that makes $\xi \mapsto \xi(A)$ measurable for every measurable $A \subset \mathcal{R}$, which is the usual σ-algebra considered on $\mathfrak{N}_{\mathrm{fin}}$; see Proposition 9.1.IV in Daley and Vere-Jones (2008).

We say a point process Ξ satisfies *complete spatial randomness (CSR)* if it is a Poisson process with intensity measure $\nu = \lambda \,\mathrm{Leb}^d$, where $\lambda \geq 0$ and Leb^d is Lebesgue measure (on $\mathcal{R}$). This means that $\Xi(A) \sim \mathrm{Po}(\nu(A))$ for all measurable $A \subset \mathcal{R}$ and that $\Xi(A_1), \ldots, \Xi(A_l)$ are independent for all $l \in \mathbb{N}$ and all measurable $A_1, \ldots, A_l \subset \mathcal{R}$ that are pairwise disjoint. See e.g. Section 2.4 in Daley and Vere-Jones (2003) for more details on the Poisson process.

6 Simulation Study

We tested the different statistics from Table 1 for various point process distributions and present the results in what follows. First we investigate the practical use of our asymptotics in Subsection 6.1. In spatial statistics there are usually two fundamentally different ways how distributions can deviate from CSR. One is spatial inhomogeneity of points, i.e. points may be more or less likely to occur in different regions of the space. The ability of tests to detect deviations from CSR against various spatially inhomogeneous alternatives is studied in Subsection 6.2. The other way is interaction of points, i.e. presence of points in one region may excite or inhibit the presence of other points nearby. In Subsection 6.3 we study how well the statistics discern between various interaction strengths in homogeneous Strauss processes.

For the evaluations in Subsections 6.2 and 6.3 we perform permutation tests. These are based on generating M independent uniform permutations of the indices of the data points resulting in alternative split-ups of the data into k groups of sizes n_i, $1 \leq i \leq k$. We then determine the rank r of the statistic-value for the original split-up within the statistic-values of the alternative split up (from $r = 1$ for the highest value to $r = M + 1$ for the lowest value). It is easily checked (and well-known) that $p = \frac{r}{M+1}$ is an honest p-value (i.e. $\mathbb{P}(p \leq \alpha) \leq \alpha$ for every $\alpha \in (0, 1)$). We reject the null if $p \leq 0.05$.

In Subsections 6.2 and 6.3 we have $k = 2$, $n_i = \tilde{n} = 20$ and use $M = 999$ permutations if no barycenter computation is needed and $M = 99$ permutations if barycenter computation is needed. In view of the $\binom{40}{20} \approx 1.4 \cdot 10^{11}$ possible split-ups, this means that there is a high degree of randomization in each individual test. The small $M = 99$ was necessary due to the large computational burden of computing pseudo-barycenters in point pattern space (see Section 5). For statistics that do not require barycenter computation, choosing $M = 999$ typically results in much faster computation time than the choice of $M = 99$ for statistics that do require barycenter computation. For reproducibility of individual test results, a higher M or (where possible) comparing within all possible split-ups into groups would be desirable in both cases.

Preferring exact permutation tests over tests based on the limit χ^2-distribution is in agreement with the recommendations from previous papers and corresponds to our own experience. However, the χ^2-approximation of our L-statistic is quite fast as we can see in Subsection 6.1, where we compare the finite sample distributions of the new L- and the Dubey–Müller statistics.

In all tests we use as the underlying space $\mathcal{R} = [0, 1]^2 \subset \mathbb{R}^2$ with the Euclidean metric. The significance level is always $\alpha = 0.05$. Furthermore we choose as order $p = 2$ and as penalty $C = 0.25$, which means that $\sqrt{2} \cdot 0.25 \approx 0.35$ is the maximal contribution that a single matched point pair makes to the TT-distance, i.e. the actual Euclidean distances are cut off at this value. In applications the choice of C is often based on the physical reality of the data and possibly the goal of the analysis. For the present simulation study we tried not to restrict a substantial proportion of matching distances while keeping the contribution of additional points reasonably low. Table 2 gives an overview of how many pairs are matched above and below the cutoff distance for various values of C based on pairwise comparisons of 1000 point patterns simulated according to CSR with intensity $\lambda = 35$. For $C = 0.25$ we have for every matching above the cutoff distance $1/0.038 \approx 26$ matchings below the cutoff distance.

	mean d_{TT}	below cutoff	above cutoff	unpaired
$C = 0.1$	0.309	11123388	4469722	3309250
relative		1	0.424	0.311
$C = 0.15$	0.393	13456877	2136233	3309250
relative		1	0.167	0.257
$C = 0.2$	0.457	14514420	1078690	3309250
relative		1	0.078	0.24
$C = 0.25$	0.512	15054688	538422	3309250
relative		1	0.038	0.233
$C = 0.3$	0.561	15347529	245581	3309250
relative		1	0.017	0.23
$C = 0.35$	0.609	15498175	94935	3309250
relative		1	0.006	0.229

Table 2: Pairwise comparison within 1000 patterns simulated from CSR on $[0, 1]^2$ with intensity $\lambda = 35$ for various penalties C. The columns give the mean d_{TT}-distance, and the (absolute and relative) number of matchings below cutoff and above cutoff, as well as the number of unpaired points for these $\binom{1000}{2}$ pairwise comparisons. Note that the relative numbers are with respect to the number of matchings below cutoff.

6.1 Asymptotics

In the present subsection we numerically assess the speed of convergence of our new $\widetilde{L}$ statistic under the null hypothesis of equal group distributions towards the χ^2_{k-1} distribution as presented in Subsection 4.2. For comparison we also consider the Fréchet T_L and T statistics, which were shown in Dubey and Müller (2019) to have a limiting χ^2_{k-1} distribution as well.

Our experiments are based on $k = 2$ groups, both simulated from the same distribution, which is either CSR(35) or the Strauss hard core distribution with $\lambda = 35$. These are the extreme distributions having either no interaction or very strong interaction in Subsection 6.3. As group size we consider $\tilde{n} = 5, 20, 50, 200$. The computation of the Fréchet T and T_L depend on the calculation of a barycenter. For this we used the heuristic algorithm presented in Müller et al. (2020). The calculation of an exact barycenter is computationally infeasible for this kind of data. To compensate that we do not get the optimal solution, we did 5 restarts in every barycenter calculation and used the best of the 5 solutions as the barycenter.

Figure 1 shows QQ-plots for the empirical distributions of our new Levene statistic $\widetilde{L}$, the Fréchet statistic T_L and the Fréchet statistic T on the y-axis and the theoretical χ^2_1 distribution on the x-axis. The data are the CSR(35) point patterns. In the first column the groups consist of $\tilde{n} = 5$ patterns, in the second column of $\tilde{n} = 20$ patterns and so on. For the two Levene statistics $\widetilde{L}$ and T_L we can see the computed quantiles approach the theoretical quantiles as the group size $\tilde{n}$ gets larger. Even for a medium group size $\tilde{n} = 50$ the computed quantiles are very close to the theoretical quantiles of a χ^2_1 distribution.

Similarly Figure 2 shows QQ-plots for hardcore Strauss distributed point patterns. Again the four columns correspond to the four group sizes $\tilde{n} = 5, 20, 50, 200$ and the three rows correspond to the three statistics. For this data the computed quantiles are already very close to the theoretical quantiles of a χ^2_1 distribution for $\tilde{n} = 20$ for the two Levene statistics.

In both cases the third row, the combined Fréchet statistic T, yields quantiles that are far from the theoretical quantiles. This is solely due to the summand T_F that is not considered in the second row.

In spite of the asymptotic results we use permutation based tests in what follows. This is in the tradition of previous methods, see e.g. Anderson (2017), Dubey and Müller (2019). Comparison of different statistics for different data sets are presented in the following two subsections.

6.2 Inhomogeneity

Here we compare $k = 2$ groups of $\tilde{n} = 20$ point patterns. Patterns in Group 2 are simulated from CSR with $\lambda = 35$. In Group 1, they are simulated from various inhomogeneous scenarios, i.e. from Poisson process distributions where the intensity function (the density of the measure ν with respect to Lebesgue measure) deviates more or less from a constant but still integrates up to 35 over the whole window $\mathcal{R} = [0, 1]^2$.

In Scenarios 1–3 the intensity is obtained by adding a number of rotation-invariant Gaussian distributions with different means but the same covariance matrix $\sigma^2 I$ and scaling to total mass 35. For simplicity we do not restrict the intensity to $\mathcal{R}$, but as can be seen from Figure 3 only very few points outside $\mathcal{R}$ occur. Scenarios 4–6 use as intensity an exponential function that is constant in the y-coordinate and induces a certain tendency for points to lie in the left part of the window rather than in the right part.

Table 3 provides more information about the chosen parameters. Figure 3 shows five example point patterns for each scenario. In addition we add a Scenario 0, which corresponds to simulating the first group also from CSR with $\lambda = 35$.

Table 4 gives the results in terms of numbers of rejections (out of 100) of the null hypothesis of equal distribution in both groups.

We observe that the direct ANOVA procedures perform much better than the Levene (or

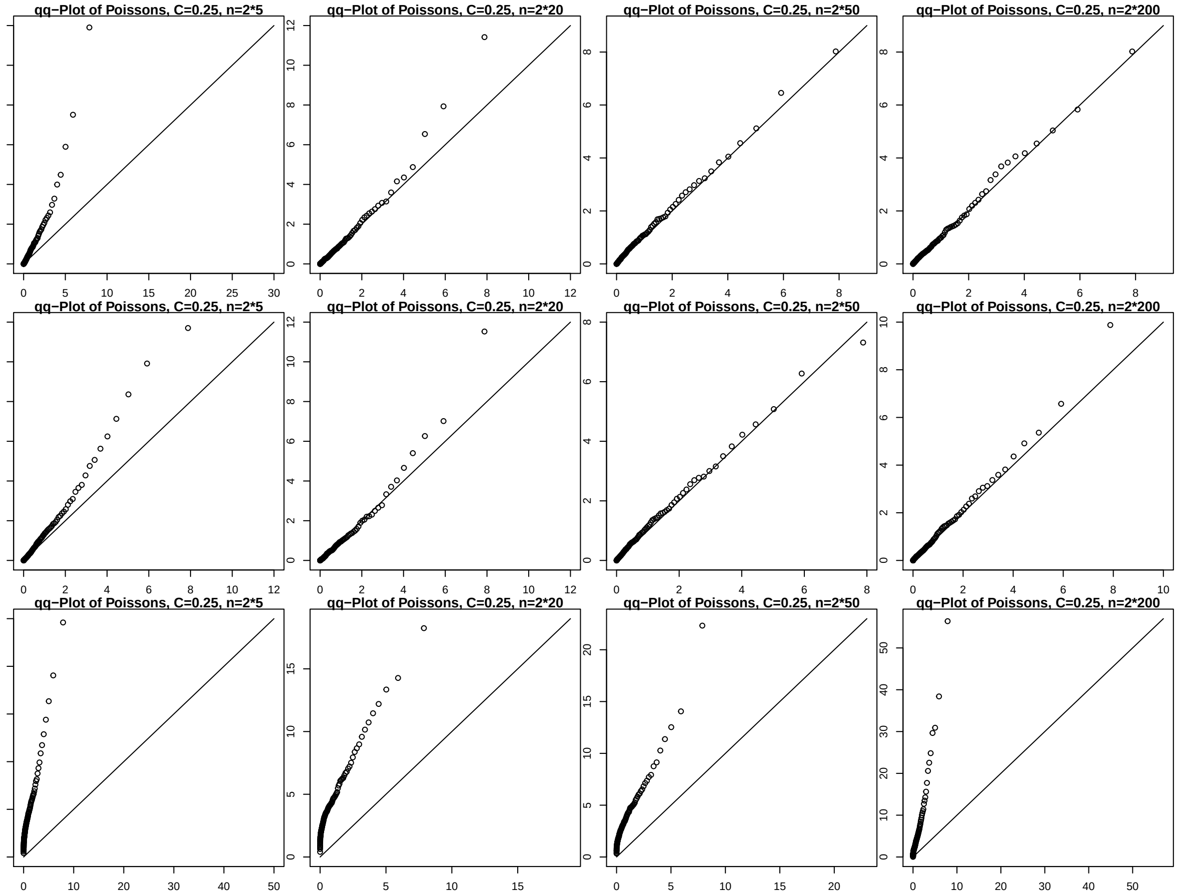


Figure 1: QQ-plots of the percentiles based on 500 statistics values (on the y-axis) versus χ^2_1-percentiles. Based on $k = 2$ groups of $\tilde{n} = 5, 20, 50, 200$ patterns from CSR(35). The first row is our new $\widetilde{L}$ statistic (7), the second and third rows are the Fréchet T_L statistic from Section 3.2 and the Fréchet T statistic, respectively.

Scenario	$\lambda(x, y)$ proportional to
1	$\sum_{i=1}^{3} \varphi_{\mu_i,0.075}(x, y)$
2	$\sum_{i=1}^{3} \varphi_{\mu_i,0.1}(x, y)$
3	$\sum_{i=1}^{4} \varphi_{\mu_i,0.1}(x, y)$
4	$\exp(-2x)$
5	$\exp(-1x)$
6	$\exp(-0.02x)$

Table 3: Overview of the Poisson process intensities for the six scenarios. The proportionality constant is chosen such that the expected number of points in each scenario is 35. By φ_{μ,σ^2} we denote the density of the bivariate normal distribution with mean $\mu \in \mathbb{R}^2$ and covariance matrix $\sigma^2 I$. The different μ_i used are seen in Figure 3.

indirect ANOVA) procedures. This is not so surprising, because the inhomogeneity experiment considers two groups of distributions that are different in terms of their location in the point pattern space. To see this intuitively, think about the distributions in Scenarios 1–6 (and 0 as a boundary case) in terms of producing locally perturbed versions of a typical point pattern, which is more or less any one of the example point patterns in Figure 3 (more appropriately one would rather think of an idealized version of these patterns, such as the Fréchet mean).

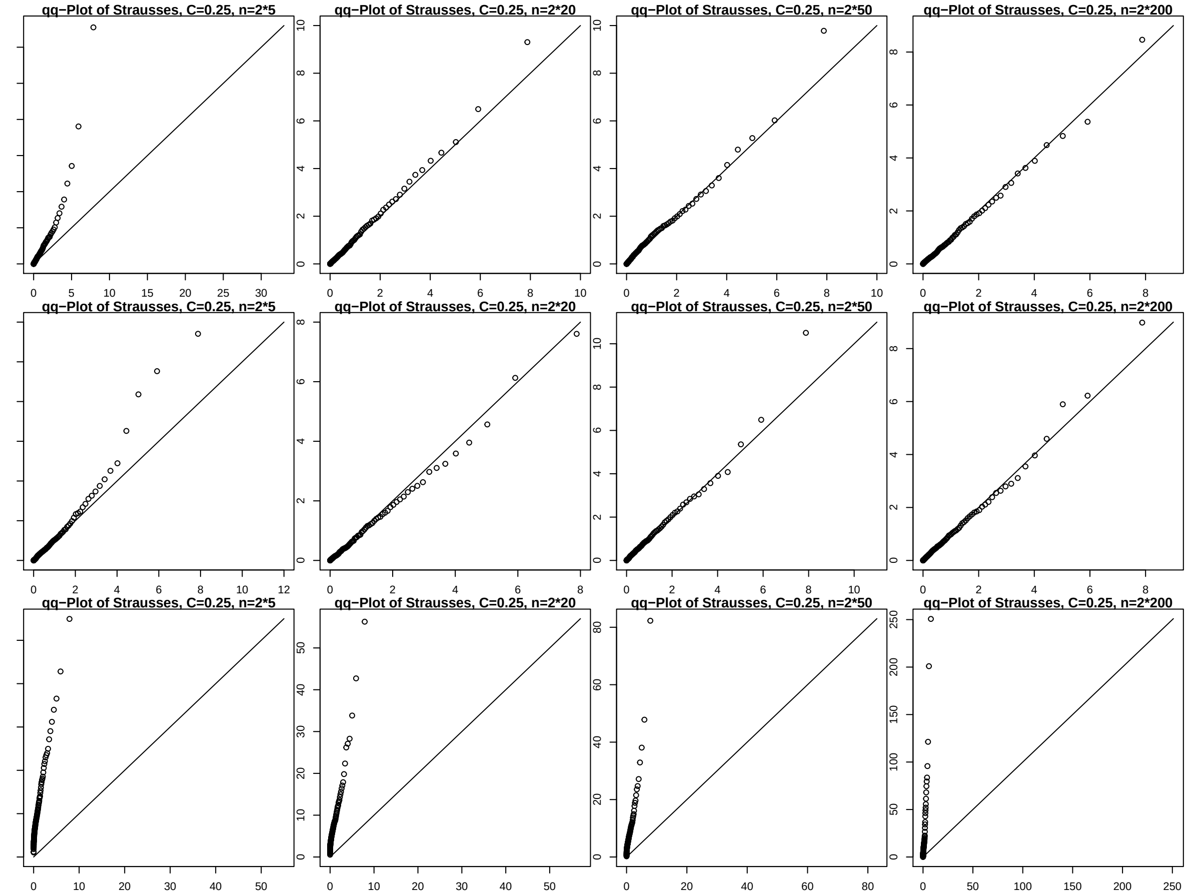


Figure 2: QQ-plots of the percentiles based on 500 statistics values (on the y-axis) versus χ^2_1-percentiles. Based on $k = 2$ groups of $\tilde{n} = 5, 20, 50, 200$ patterns from a Strauss hard core distribution with $\lambda = 35$. The first row is our new $\widetilde{L}$ statistic (7), the second and third rows are the Fréchet T_L statistic from Section 3.2 and the Fréchet T statistic, respectively.

Among the direct ANOVA methods, Anderson F_A performs substantially better than Fréchet T_F and has still a reasonable chance to detect the faint differences between Scenarios 6 and 0 when presented with the 20 patterns from each group. Our new L-test performs somewhat better than the Fréchet L-test, but both tests are only able to detect the inhomogeneity (with reasonable probability) when it is very obvious (Scenarios 1–3).

6.3 Interaction between Points

Again we compare $k = 2$ groups of $\tilde{n} = 20$ point patterns. This time the group distributions differ in the degree of point interaction. For this we consider the distribution of the homogeneous Strauss process on the unit square $\mathcal{R} = [0, 1]^2$, which is obtained by specifying the density $f \colon \mathfrak{N}_{\text{fin}} \to \mathbb{R}_+$,

$$f(\xi) := c \cdot \beta^{|\xi|} \cdot \gamma^{s_R(\xi)},$$

with respect to CSR with intensity 1 on $\mathcal{R}$, where

$$s_R(\xi) = \sum_{\{x,y\} \subset \xi} \mathbb{1}\{\|x - y\| \leq R\}$$

is the number of pairs of points at distance $\leq R$ from one another. Here $R > 0$ is the range of the interaction, $\beta > 0$ is the so-called activity (which controls the intensity of the process via an increasing function, that is however only accessible numerically) and $\gamma \in [0, 1]$ is the

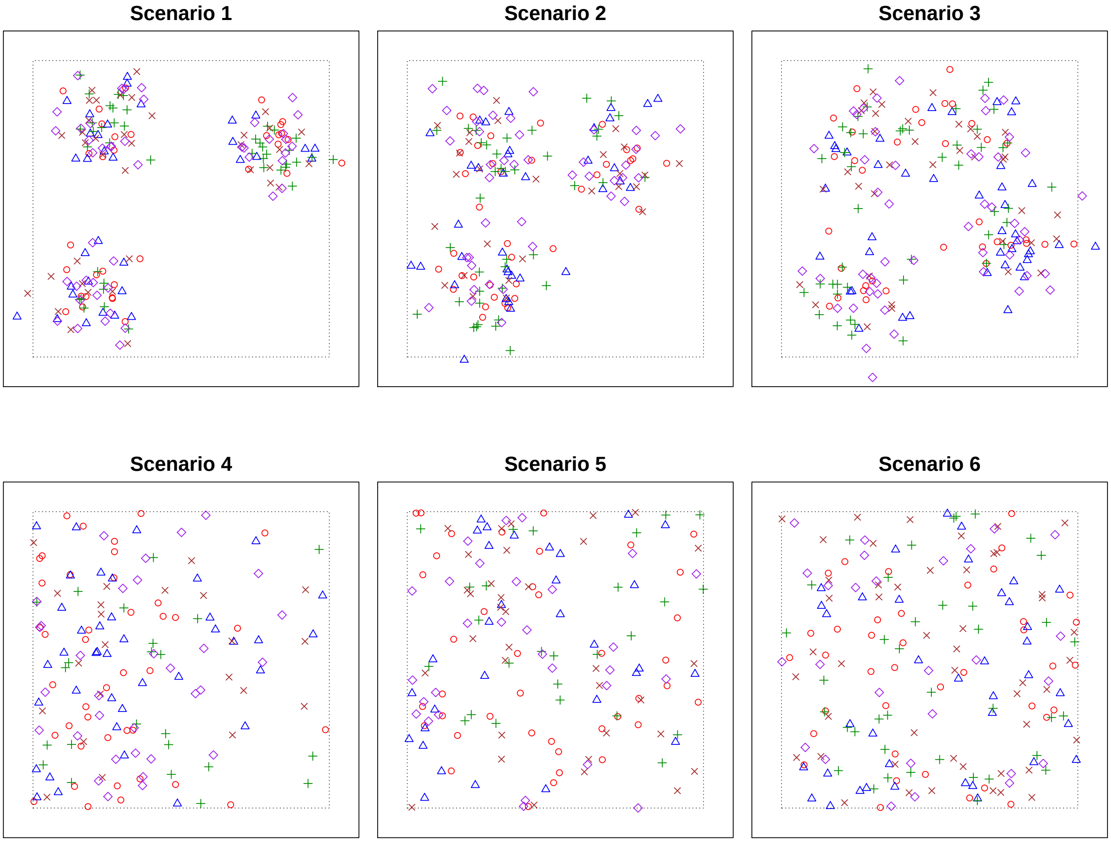


Figure 3: Five example patterns for each of the six scenarios together with the window $[0,1]^2$ (dotted line) in which the homogeneous Poissons processes of the second group are sampled.

strength of the interaction. The constant c normalizes the density to an overall integral of 1 and is also not available in closed form. We write Strauss$(\beta, \gamma; R)$ for this point process distribution. Intuitively a Strauss$(\beta, \gamma; R)$-process is obtained from a CSR(β) process by penalizing each outcome according to a factor γ per R-close point pair. Correspondingly we have Strauss$(\beta, 1; R) =$ CSR(β) (regardless of R). At the other end of the spectrum Strauss$(\beta, 0; R)$ is the distribution of a hard core process with no points allowed within distance R of other points.

For the simulation we set $R = 0.1$ and consider scenarios based on the six different values $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$. The activity β is adapted so that each time $\lambda = 35$. Figure 4 shows one realization for each of the six scenarios.

We perform two different experiments here. In the first one the patterns in Group 1 are sampled from CSR(35) corresponding to $\gamma = 1$, in the second one they are sampled from the mentioned hard core process with $\lambda = 35$ corresponding to $\gamma = 0$. The patterns in Group 2

Scenario	1	2	3	4	5	6	0
Anderson F_A	100	100	100	100	99	39	2
new L	93	76	77	14	7	9	3
Fréchet T_F	100	100	100	99	11	0	4
Fréchet T_L	59	24	47	13	9	12	4

Table 4: Numbers of rejections of the null hypothesis “equal distribution in both groups” based on 100 data sets per column. In each data set the first group is sampled from the scenario indicated in the column and the second group is sampled from Scenario 0.

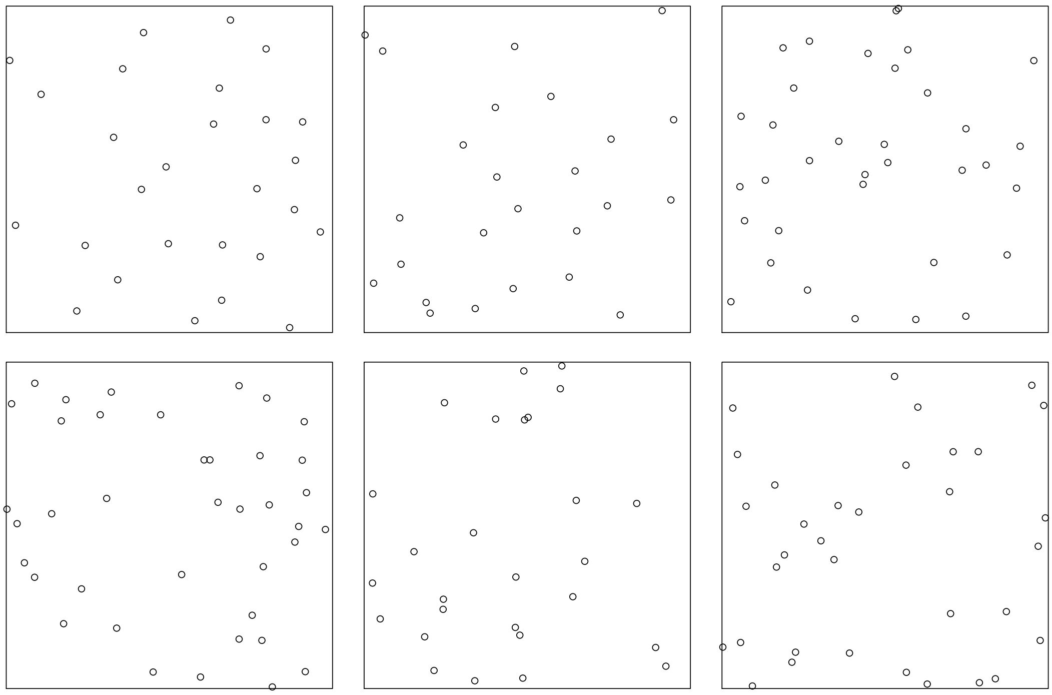


Figure 4: Simulations from Strauss($\beta, \gamma; 0.1$)-distributions, where rowwise from left to right $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$ and β is adjusted such that $\lambda = 35$. For $\gamma = 0$ we have a realization of a hard core process, for $\gamma = 1$ a realization from CSR.

are sampled in both experiment from each of the six γ-values in turn. The results are listed in Table 5.

$\gamma = 1$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	98	41	13	8	4	5
new L	100	100	95	67	20	3
Fréchet T_F	100	98	78	45	20	4
Fréchet T_L	100	99	88	45	20	4

$\gamma = 0$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	3	55	90	98	97	99
new L	11	60	96	100	100	100
Fréchet T_F	6	57	91	97	100	100
Fréchet T_L	9	33	82	95	99	100

Table 5: Numbers of rejections of the null hypothesis “equal distribution in both groups” based on 100 data sets per column. In each data the point patterns in both groups are sampled from a Strauss distribution with $\lambda = 35$ and $R = 0.1$. The first group is sampled using $\gamma = 1$ or $\gamma = 0$ as indicated on the top left of the table and the second group uses γ as indicated in the column.

In contrast to the situation in the previous subsection (different inhomogeneity), we now observe that the *indirect* ANOVA procedures, i.e. the Levene-type tests perform considerably better than the direct ANOVA procedures. Again this is intuitively understandable because a small γ in the Strauss process leads to less dispersion, both in terms of a smaller variance for the total number of points and also with respect to typical distances of points from one another: for small γ the points are quite regularly placed, whereas for larger γ there are erratic patches that are free of points leading typically to some points that have to be matched over longer

distances, which in the squared Euclidean metric has quite some influence. A small γ will also lead to smaller average distances than a larger γ (either between point patterns or relative to a barycenter), which may explain why the difference in the performance of the indirect and direct ANOVA tests is somewhat less pronounced than in the inhomogeneity experiment.
Note again that the powers of the tests based on pairwise distances are slightly better than those of the tests based on barycenters.

7 Applications

In this section we apply our Levene's test to a real data example. We investigate the location of bubbles in a mineral flotation experiment. The structure of the data calls for a two factor design. We establish a distance based two-way Levene's test and compare its performance to existing methods. The classical two-way ANOVA design can be found in Section 2.

7.1 Balanced Two-Way Levene's Test

As mentioned in Subsection 4.1 it is easy to generalize statistic (5) to a two-way design, that will further be useful for the bubble data analyzed in the next Subsection.

Suppose we have independent observations $x_{i_1 i_2 j} \in \mathcal{X}$, $1 \leq j \leq \tilde{n}$, $1 \leq i_1 \leq k_1$, $1 \leq i_2 \leq k_2$ from groups obtained by crossing a Factor a with k_1 levels and a Factor b with k_2 levels with $\tilde{n}$ observations for each combination. In a similar way as above we denote by $d_{i_1 i_2 j}$ the j-th half-distance in the group (i_1, i_2), where $j = 1, \ldots, \tilde{N} := \binom{\tilde{n}}{2}$. Set then

$$\begin{aligned}
\mathrm{RSS} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \sum_{j=1}^{\tilde{N}} (d_{i_1 i_2 j} - \bar{d}_{i_1 i_2 \cdot})^2 \\
\mathrm{MSS} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \tilde{n}(\bar{d}_{i_1 i_2 \cdot} - \bar{d}_{\cdots})^2 \\
\mathrm{SSa} &= \sum_{i_1=1}^{k_1} k_2 \tilde{n}(\bar{d}_{i_1 \cdot\cdot} - \bar{d}_{\cdots})^2 \\
\mathrm{SSb} &= \sum_{i_2=1}^{k_2} k_1 \tilde{n}(\bar{d}_{\cdot i_2 \cdot} - \bar{d}_{\cdots})^2 \\
\mathrm{SSi} &= \sum_{i_1=1}^{k_1} \sum_{i_2=1}^{k_2} \tilde{n}(\bar{d}_{i_1 i_2 \cdot} - \bar{d}_{i_1 \cdot\cdot} - \bar{d}_{\cdot i_2 \cdot} + \bar{d}_{\cdots})^2,
\end{aligned}$$

where the various means are taken over the dot components in the usual way. Note that we never use any distances between observations of different factor combinations.

In addition to the omnibus test for group differences as in one-way ANOVA, we may then perform Levene-type tests for effects of Factor a and b separately, as well as for an interaction effect. The corresponding statistics are

$$L = \frac{N - k_1 k_2}{(k_1 k_2 - 1)} \frac{\mathrm{MSS}}{\mathrm{RSS}}, \quad La = \frac{N - k_1 k_2}{k_1 - 1} \frac{\mathrm{SSa}}{\mathrm{RSS}}, \quad Lb = \frac{N - k_1 k_2}{k_2 - 1} \frac{\mathrm{SSb}}{\mathrm{RSS}}, \quad Li = \frac{N - k_1 k_2}{(k_1 - 1)(k_2 - 1)} \frac{\mathrm{SSi}}{\mathrm{RSS}}.$$

7.2 Bubble Data

We consider the data from González et al. (2021) which provides locations of bubbles in a mineral flotation experiment, where the interest is analysing if the spatial distribution might be affected by frother concentrations and volumetric airflow rates. Indeed, the data set consists of

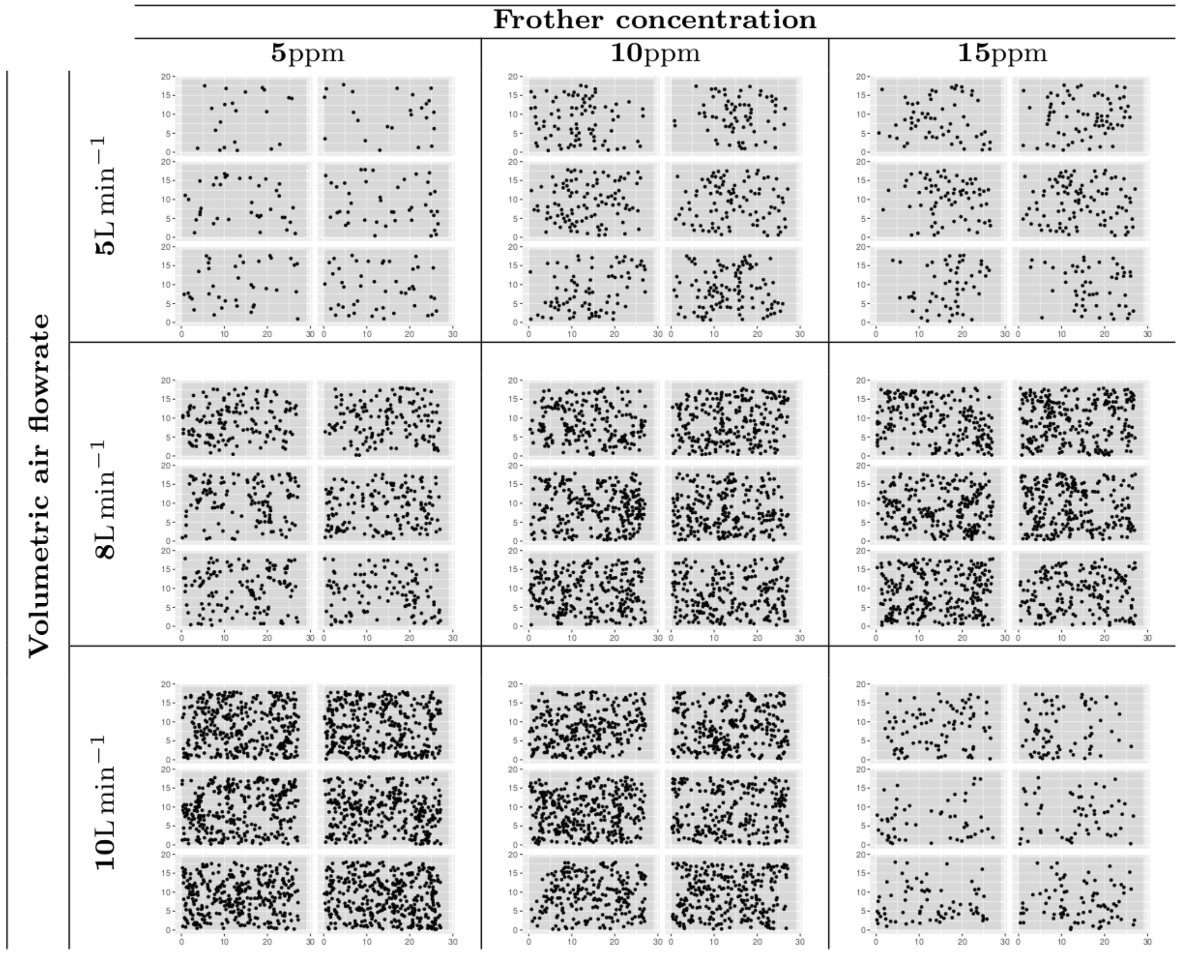


Figure 5: Arrangement of floating bubbles data. Rows represent the three frother concentration levels and columns the three volumetric air flowrate levels (treatments). Each cell contains six spatial point patterns (responses).

54 images containing a total of 8385 floating bubbles. The images of bubbles can be regarded as spatial point patterns where the centroids of the bubbles correspond to the points. In addition, we have three frother concentration levels (5 ppm, 10 ppm, 15 ppm) as well as three volumetric airflow rate levels (5 l/min, 8 l/min, 10 l/min), and we have six replicates of point patterns at each combination of levels of such factors. The treatment combinations of the experiment, as well as the observed bubble point patterns, are represented in Figure 5.

We used the two-way design of Levene's statistic from Section 7.1 to test for influence of the different factors, interaction and differences between the groups. For comparison we also used the two factor statistics from Anderson (2001), we performed a two factor ANOVA on the number of points per pattern, and finally complemented our analysis with a two factor ANOVA with K-functions, so as to link our analysis with that of González et al. (2021). We did a permutation test with 999 permutations.

In Section 6, the cutoff was always fixed to $C = 0.25$. This was a reasonable value for point patterns with expected 35 points in the unit square. In the bubble data, the number of points per observed pattern ranges from 21 to 353. With such a great variability in the number of points we suggest adjusting the cutoff to prevent that distances between two patterns are dominated by their different numbers of points. For the results presented in this section we computed the mean number of points of the tested patterns $\bar{n}$ and used the cutoff $\bar{C} = 0.25 \cdot 35/\bar{n}$ for the computations of d_{TT}. For more details to the cutoff see (16).

p-values	FC	VA	Interaction	Overall
Anderson F_A	0.003	0.001	0.001	0.001
new L	0.001	0.001	0.001	0.001
Number of points	0.001	0.001	0.001	0.001
K-functions	0.005	0.001	0.001	0.001**

** this is the p-value for the sum of both factors, not the overall ANOVA statistic.

Table 6: Results for the different tests for the bubble data. Quantiles are obtained by a permutation test with 999 permutations. The cutoff is $C = 0.0564$, the maximal radius for the K-functions is $r = 0.15$.

p-values	FC	VA	Interaction	Overall
Anderson F_A	0.043	0.001	0.019	0.001
new L	0.001	0.001	0.001	0.001
Number of points	0.001	0.002	0.001	0.001
K-functions	0.002	0.022	0.002	0.006**

** this is the p-value for the sum of both factors, not the overall ANOVA statistic.

Table 7: Results for the different tests for the bubble data, leaving out the frother concentration of 15ppm. Quantiles are obtained by a permutation test with 999 permutations. The cutoff is $C = 0.0636$, the maximal radius for the K-functions is $r = 0.15$.

The p-values of the permutation tests are shown in Tables 6 and 7. In particular, Table 6 shows results for the whole data set, while Table 7 depicts results for only part of the data, leaving out the third column, i.e. any patterns from frother concentration of 15 ppm. In both cases, our new Levene, Anderson F_A, the ANOVA on the number of points per pattern, and the ANOVA for K-functions detect significant influence of each of the two factors and the interaction. We already recommended to always perform both, the tests for differences in variability and the test for differences of means. In the second test scenario, both Levene's test and Anderson F_A detect significance for the frother concentration and the interaction of both parameters for our usual significance level of 5%. But the relative difference between the p-values of the two tests is very large. For the smaller significance level of 1%, our Levene's test still detects significance where Anderson F_A does not. So the test for differences of means might not be enough in a practical application. This is particularly important in cases where, as it is the case for the bubble data, the number of points plays a crucial role in the behavior and structure of the point patterns.

We see that for this data apparently the numbers of points per pattern contain enough information to detect significant influence of the factors. This is not very surprising since the number of points per pattern is similar in the 6 patterns of a single cell, but the differences between cells are large.
This observation is reinforced by a classical multidimensional scaling (mds). Based on the TT-distances between the point patterns, we translated every point pattern into a single point in $\mathbb{R}^2$. The mds was applied first for the whole bubble data set, see Figure 6, and then for a subset of the data consisting of the first and second columns, leaving out the data with a frother concentration of 15 ppm, see Figure 7. This is the same data that we used for our analyses in Tables 6 and 7. The three levels of the air flow are encoded by the colors 'red', 'green' and 'blue', same color means same air flow rate, and the three levels of the frother concentration are encoded by the symbols 'circle', 'triangle' and 'cross'. When we compare these plots to the images of the point patterns in Figure 5 we can see that the multidimensional scaling sorts the point patterns from left to right in ascending order by their number of points per pattern. In Figure 7 we can see that the points that correspond to the data with a frother concentration of

5 ppm, i.e. the circles, and the data with a frother concentration of 10 ppm, i.e. the triangles, are scattered differently. The (coordinate-wise) means of the triangles and circles are similar, but we can see that the circles are more scattered along both axes. We conjecture that it is this difference in scatter that our Levene's test is able to detect in Table 7, whereas the Anderson F_A only barely detects a slight difference in means.

8 Conclusions and Discussion

In this paper we gave an overview of some ANOVA procedures that can be used for data in general metric spaces. We introduced a new method that is similar to Levene's test and compared it to existing methods with regard to point pattern data in Section 6. In the studies, see Tables 4 and 5 for the results, we compared the distance-based ANOVA from Anderson (2001), the distance-based Levene's test, (5), introduced in this paper and the tests based on the ANOVA statistic T_F and the Levene statistic T_L from Dubey and Müller (2019).
The latter proposed in their paper the combined statistic $T = T_L + T_F$. In our simulations we wanted to put a focus on the two fundamentally different ways of "location" and "dispersion" in which group distributions can differ, even in an abstract metric space. We therefore did our tests with the statistics T_L and T_F seperately, which allows us to have a direct comparison between the distance-based statistics and the statistics of Dubey and Müller (2019). We also did the tests with the proposed combined statistic T, and for completeness we give the results in Tables 8 and 9. In the tests for differences in interaction, comparing Tables 9 and 5, we can see that the performance of the statistic T is "between" the performance of T_L and T_F. For the tests of inhomogeneity, see Tables 8 and 4, the combined statistic T performs almost as good as the better statistic of T_L and T_F, which is in this case T_F.
For the presented scenarios and the chosen parameters all the statistics worked well for their designated purpose, in particular the ANOVA statistics for detecting inhomogeneity and the Levene's statistics for detecting differences in point interaction. But for different scenarios this might not be the case. For a cutoff of $C = 0.1$ instead of the proposed $C = 0.25$, we observed that the statistic T and the Anderson F_A do not work as well anymore in detecting differences in interaction, see Tables 10 and 11. The performance of T is considerably worse than the performance of T_L and T_F is working even more poorly as well. The distance based ANOVA statistic of Anderson also performs very poorly compared to the cutoff of $C = 0.25$, while our statistic L is working even better with the smaller cutoff.

For future research it would be interesting to take a closer look at the (co)variance estimator γ of our $\widetilde{L}$ statistic. This estimator is not unbiased and it remains open if a statistic with an unbiased estimator works even better, in particular for the asymptotics.
There are also more complex designs for the 2-factor ANOVA, with different sums of squares or designs that allow for different group sizes. Our L statistic could be generalized to more complex 2-factor designs, or even k-factor designs.
Additionally it would be interesting to test our statistics with more kinds of data. On the one hand there are different kinds of point pattern data, e.g. marked point patterns, but also data from other metric spaces, e.g. image data or graph data.

In Section 6 we already mentioned that the computation of the Fréchet T statistic is more time consuming than the distance based tests, because of the barycenter calculation. If we take for example the data from Section 6.3, the distance based Anderson F_A and our new L take about 8 seconds and 2 seconds, respectively, for 100 permutation tests with 999 permutations each. The calculation of 100 permutation tests of Fréchet T with 99 permutations each takes about 45 minutes. The workhorse computation of all tests is done in C++. However, parts of the overhead for Anderson F_A and Fréchet T are programmed in R and a complete implementation in C++ might improve the runtime a little bit. These numbers are merely meant to give a

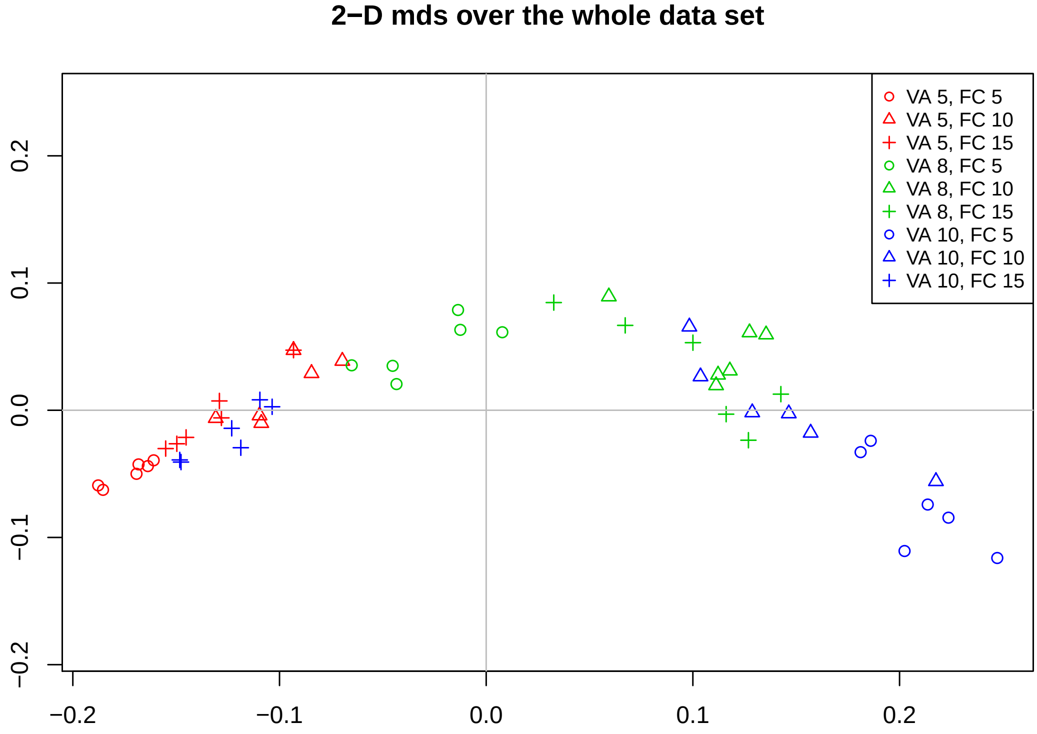


Figure 6: The bubble data after a multidimensional scaling into two dimensions based on the distance matrix w.r.t the TT-metric.

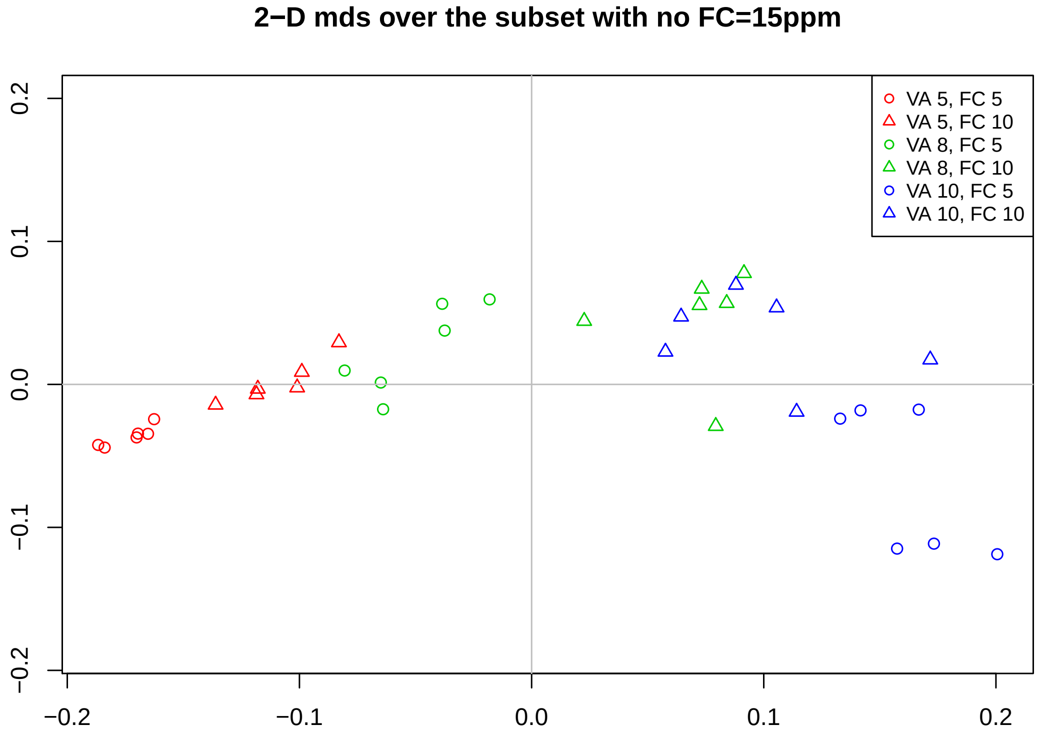


Figure 7: The bubble data without the third column, i.e. without data where FC=15ppm, after a multidimensional scaling into two dimensions based on the distance matrix w.r.t the TT-metric.

general impression of the order of magnitude of the runtimes. The distance based tests take only a few seconds, while for the Fréchet tests the computation of the barycenters takes many minutes, even with the fast heuristics instead of the exact solution and with fewer permutations.

Overall we find that the new L in combination with the Anderson F_A has a similar performance and allows for considerably faster computation than the other methods in settings where the computation of barycenters is costly.

Scenario	1	2	3	4	5	6	0
Fréchet T	100	100	100	98	11	2	6

Table 8: Performance of the sum statistic Fréchet T. Numbers of rejections of the null hypothesis "equal distribution in both groups" based on 100 data sets per column for the 7 scenarios of inhomogeneity.

$\gamma = 1$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Fréchet T	100	98	77	48	16	2
$\gamma = 0$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Fréchet T	5	48	91	98	100	100

Table 9: Performance of the sum statistic Fréchet T. Numbers of rejections of the null hypothesis "equal distribution in both groups" based on 100 data sets per column for the different scenarios of interaction between points.

$\gamma = 1$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	37	16	8	5	12	3
new L	100	100	96	69	15	6
Fréchet T_F	49	31	10	4	5	9
Fréchet T_L	100	99	87	47	13	7
Fréchet T	99	84	38	12	7	9

Table 10: C=0.1, The first scenario: Groups of Poisson-distributed point patterns vs groups of Strauss-distributed point patterns with 6 different gammas: $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$, $R = 0.1$, $\lambda = 35$, $\alpha = 0.05$, 20 patterns per group. Numbers indicate how many times the hypothesis "equal distributions in both groups" is rejected out of 100 times. The tests should see no difference between groups of Poisson-patterns and Strauss-patterns with $\gamma = 1$.

$\gamma = 0$ vs.	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$	$\gamma = 1$
Anderson F_A	5	11	35	41	40	50
new L	5	99	100	100	100	100
Fréchet T_F	8	14	26	41	40	53
Fréchet T_L	7	87	100	100	100	100
Fréchet T	7	31	73	95	100	100

Table 11: C=0.1, The second scenario: Groups of Strauss-distributed point patterns with a fixed $\gamma = 0$ vs groups of Strauss-distributed point patterns with 6 different gammas: $\gamma = 0, 0.2, 0.4, 0.6, 0.8, 1$, $R = 0.1$, $\lambda = 35$, $\alpha = 0.05$, 20 patterns per group. Numbers indicate how many times the hypothesis "equal distributions in both groups" is rejected out of 100 times. The tests should see no difference for $\gamma = 0$.

References

Alekseyenko, A. V. (2016). Multivariate Welch t-test on distances. *Bioinformatics*, 32(23):3552–3558.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral ecology*, 26(1):32–46.

Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62(1):245–253.

Anderson, M. J. (2017). Permutational multivariate analysis of variance (PERMANOVA). *Wiley statsref: statistics reference online*, pages 1–15.

Anderson, M. J., Walsh, D. C., Robert Clarke, K., Gorley, R. N., and Guerra-Castro, E. (2017). Some solutions to the multivariate Behrens–Fisher problem for dissimilarity-based analyses. *Australian & New Zealand Journal of Statistics*, 59(1):57–79.

Bertsekas, D. P. (1988). The auction algorithm: A distributed relaxation method for the assignment problem. *Annals of operations research*, 14(1):105–123.

Billera, L. J., Holmes, S. P., and Vogtmann, K. (2001). Geometry of the space of phylogenetic trees. *Advances in Applied Mathematics*, 27(4):733–767.

Borgwardt, S. and Patterson, S. (2020). Improved linear programs for discrete barycenters. *Informs Journal on Optimization*, 2(1):14–33.

Borgwardt, S. and Patterson, S. (2021). On the computational complexity of finding a sparse Wasserstein barycenter. *Journal of Combinatorial Optimization*, 41(3):736–761.

Brown, M. B. and Forsythe, A. B. (1974a). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364–367.

Brown, M. B. and Forsythe, A. B. (1974b). The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16(1):129–132.

Cuevas, A., Febrero, M., and Fraiman, R. (2004). An ANOVA test for functional data. *Computational statistics & data analysis*, 47(1):111–122.

Daley, D. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I.* Springer, 2nd edition. Elementary theory and methods.

Daley, D. and Vere-Jones, D. (2008). *An introduction to the theory of point processes. Vol. II.* Springer, 2nd edition. General theory and structure.

Denker, M. and Keller, G. (1983). On U-statistics and v. Mises' statistics for weakly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 64(4):505–522.

Dubey, P. and Müller, H.-G. (2019). Fréchet analysis of variance for random objects. *Biometrika*, 106(4):803–821.

Fisher, R. (1925). *Statistical Methods for Research Workers.* Oliver & Boyd.

Gastwirth, J. L., Gel, Y. R., and Miao, W. (2009). The impact of Levene's test of equality of variances on statistical theory and practice. *Statistical Science*, 24(3):343–360.

Ginestet, C. E., Li, J., Balachandran, P., Rosenberg, S., and Kolaczyk, E. D. (2017). Hypothesis testing for network data in functional neuroimaging. *The Annals of Applied Statistics*, pages 725–750.

González, J. A., Lagos-Álvarez, B. M., and Mateu, J. (2021). Two-way layout factorial experiments of spatial point pattern responses in mineral flotation. *TEST*, pages 1–30.

Hamidi, B., Wallace, K., Vasu, C., and Alekseyenko, A. V. (2019). W_d^*-test: robust distance-based multivariate analysis of variance. *Microbiome*, 7(1):1–9.

Heinemann, F., Munk, A., and Zemel, Y. (2021). Randomised Wasserstein barycenter computation: Resampling with statistical guarantees. *Preprint.* Available at https://arxiv.org/abs/2012.06397.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325.

Hoeffding, W. (1961). The strong law of large numbers for U-statistics. Technical Report, Mimeograph Series 302, Department of Statistics, University of North Carolina.

Huckemann, S., Hotz, T., and Munk, A. (2009). Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):593–603.

Lee, Y. T. and Sidford, A. (2014). Path finding methods for linear programming: Solving linear programs in Õ(vrank) iterations and faster algorithms for maximum flow. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 424–433. IEEE.

Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics*, pages 278–292.

Mardia, K., Kent, J., and Bibby, J. (1979). *Multivariate Analysis.* Academic Press, London.

Müller, R. and Schuhmacher, D. (2021). *ttbary: Barycenter Methods for Spatial Point Patterns.* R package version 0.2-0. https://CRAN.R-project.org/package=ttbary.

Müller, R., Schuhmacher, D., and Mateu, J. (2020). Metrics and barycenters for point pattern data. *Statistics and Computing*, 30:953–972.

Ramón, P., de la Cruz, M., Chacón-Labella, J., and Escudero, A. (2016). A new non-parametric method for analyzing replicated point patterns in ecology. *Ecography*, 39(11):1109–1117.

Scheffé, H. (1967). *The analysis of variance.* John Wiley & Sons, 5th printing, 1st edition.

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38(3/4):330–336.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data.* MIT press.

A Auxiliary Results Used for the Proof of Theorem 2

For completeness and self-containedness we state here (consequences of) results from the literature as well as some additional calculations needed for the proof of Theorem 2.

Firstly we formulate a straightforward generalization of Hoeffding's theorem for the asymptotic normality of U-statistics (univariate version of Theorem 7.1 in Hoeffding, 1948) for random elements in the general metric space $\mathcal{X}$ with countably generated Borel σ-algebra. See also Theorem 1(b) of Denker and Keller (1983), where this result is further generalized to (weakly) dependent sequences of random elements.

Theorem 5. *Let $(X_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence of $\mathcal{X}$-valued random elements. Let $h\colon \mathcal{X}^m \to \mathbb{R}$ be symmetric and non-degenerate in the sense that there are $x_2, \dots, x_m \in \mathcal{X}$ such that*

$$\mathbb{E}h(X_1, x_2, \dots, x_m) \neq 0.$$

Suppose further that $\mathbb{E}\big(h(X_1, \dots, X_m)^2\big) < \infty$. We write

$$U_n = \binom{n}{m}^{-1} \sum_{\substack{i_1,\dots,i_m=1\\ i_1<\dots<i_m}}^{n} h(X_{i_1}, \dots, X_{i_m}).$$

for the U-statistic with kernel h. Then

$$\sqrt{n}(U_n - \mathbb{E}(U_n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, m^2\gamma_h^2),$$

where for an independent copy $(\tilde{X}_2, \dots, \tilde{X}_m)$ of $(X_2, \dots, X_m)$

$$\gamma_h^2 = \operatorname{Cov}\big(h(X_1, X_2, \dots, X_m), h(X_1, \tilde{X}_2, \dots, \tilde{X}_m)\big) = \operatorname{Var}\big(\mathbb{E}(h(X_1, \dots, X_m) \,|\, X_1)\big).$$

Remark 6. *In the setting of Theorem 5 above, Theorem 5.2 of Hoeffding (1948) yields*

$$m^2\gamma_h^2 \leq n\operatorname{Var}(U_n) \leq m\operatorname{Var}(h(X_1, \dots, X_m))$$

for all $n \geq m$. The right hand bound is sharp for $n = m$ and $n\operatorname{Var}(U_n) \searrow m^2\gamma_h^2$ as $n \to \infty$.

The above inequality means in particular that for finite n the expression $\frac{m^2}{n}\gamma_h^2$ can only underestimate $\operatorname{Var}(U_n)$. The exact formula for $m = 2$ is

$$n\operatorname{Var}(U_n) = \frac{n-2}{n-1} \cdot 4\gamma_h^2 + \frac{1}{n-1} \cdot 2\operatorname{Var}(h(X_1, X_2)).$$

The next result is similar to classical ANOVA. For completeness we give its proof.

Lemma 7. *Let $C \in \mathbb{R}^{(k-1)\times k}$ as in (12), $D \in \mathbb{R}^{n\times k}$ as in (11), $U = (u_1, \dots, u_k)'$ and $\nu = (n_1, \dots, n_k)'$. We have*

$$C'(C(D'D)^{-1}C')^{-1}C = D'D - \frac{1}{n}\nu\nu'$$

and

$$U'(D'D - \frac{1}{n}\nu\nu')U = \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_i n_j (u_i - u_j)^2$$

Proof. Define

$$\nu_{(i)} := (n_1, \dots, n_i)' \ , \ \Lambda_{(i)} := \operatorname{diag}(\nu_{(i)}) \in \mathbb{R}^{i\times i} \ \text{ and } \ \mathbb{1}_{(i)} := (1, \dots, 1)' \in \mathbb{R}^i.$$

Then $\mathbb{1}_{(i)}\mathbb{1}'_{(i)}$ is the $i \times i$ matrix of 1's. We build up the equality step by step. Since $D'D = \Lambda_{(k)}$ and therefore

$$(D'D)^{-1} = (\Lambda_{(k)})^{-1} = \operatorname{diag}(1/n_1, \ldots, 1/n_k),$$

We obtain

$$C(D'D)^{-1}C' = (\Lambda_{(k-1)})^{-1} + \frac{1}{n_k} \cdot \mathbb{1}_{(k-1)}\mathbb{1}'_{(k-1)}$$

and

$$(C(D'D)^{-1}C')^{-1} = \Lambda_{(k-1)} - \frac{1}{n} \cdot \nu_{(k-1)}\nu'_{(k-1)}$$

and finally

$$C'(C(D'D)^{-1}C')^{-1}C = \Lambda_{(k)} - \frac{1}{n} \cdot \nu\nu'$$

When we multiply the vector U from left and right, the ij-th entry in the matrix is the coefficient of u_iu_j. This leads to

$$\begin{aligned}
&U'(D'D - \frac{1}{n}\nu\nu')U \\
&\quad= \sum_{i=1}^{k} n_i u_i^2 - \frac{1}{n}\sum_{i=1}^{k} n_i^2 u_i^2 - \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} 2n_in_ju_iu_j \\
&\quad= \frac{1}{n}\sum_{i=1}^{k} n_i u_i^2 \sum_{j=1}^{k} n_j - \frac{1}{2n}\sum_{i=1}^{k} n_i^2(u_i^2 + u_i^2) - \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} 2n_in_ju_iu_j \\
&\quad= \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=1}^{k} n_in_j(u_i^2 + u_j^2) - \frac{1}{2n}\sum_{i=1}^{k} n_i^2(u_i^2 + u_i^2) - \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} 2n_in_ju_iu_j \\
&\quad= \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_in_j(u_i^2 + u_j^2) - \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} 2n_in_ju_iu_j \\
&\quad= \frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_in_j(u_i - u_j)^2.
\end{aligned}$$

□

Remark 8. *Let $\bar{u} = \frac{1}{k}\sum_{i=1}^{k} u_i$. An equivalent expression for $U'(D'D - \frac{1}{n}\nu\nu')U$ in Lemma 7 can be computed as*

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} n_in_j(u_i - u_j)^2 \\
&\quad= \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=1}^{k} n_in_j((u_i - \bar{u}) + (\bar{u} - u_j))^2 \\
&\quad= \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=1}^{k} n_in_j(u_i - \bar{u})^2 + \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=1}^{k} n_in_j(\bar{u} - u_j)^2 + \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=1}^{k} n_in_j(u_i - \bar{u})(\bar{u} - u_j) \\
&\quad= \sum_{i=1}^{k} n_i(u_i - \bar{u})^2 + \frac{1}{2n}\sum_{i=1}^{k} n_i(u_i - \bar{u})\sum_{j=1}^{k} n_j(\bar{u} - u_j) \\
&\quad= \sum_{i=1}^{k} n_i(u_i - \bar{u})^2 + \frac{1}{2n}\left(\sum_{i=1}^{k} n_i(u_i - \bar{u})\right)^2
\end{aligned}$$

If $n_1 = \ldots = n_k = \tilde{n}$, we see directly from the right-hand side that

$$\frac{1}{n} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} n_i n_j (u_i - u_j)^2 = \sum_{i=1}^{k} n_i (u_i - \bar{u})^2 = \tilde{n} \sum_{i=1}^{k} (u_i - \bar{u})^2. \tag{18}$$

The following lemma is well known. It follows by spectral decomposition, see e.g. Kent, Mardia and Bibby (1979), Theorem 3.4.4(b), setting $p = 1$ and $\Sigma = I$.

Lemma 9. *Let $Z \sim \mathcal{N}_n(0, I)$ and let $C \in \mathbb{R}^{n \times n}$ be symmetric and idempotent. Then $Z'CZ \sim \chi^2_r$, where $r = \mathrm{trace}(C) = \mathrm{rank}(C)$.*

APPENDIX C

Location problems with cutoff

Location problems with cutoff

Raoul Müller[*][†][‡] Anita Schöbel[§] Dominic Schuhmacher[‡]

May 3, 2022

Abstract

In this paper we study a generalized version of the Weber problem of finding a point that minimizes the sum of its distances to a finite number of given points. In our setting these distances may be *cut off* at a given value $C > 0$, and we allow for the option of an *empty* solution at a fixed cost C'. We analyze under which circumstances these problems can be reduced to the simpler Weber problem, and also when we definitely have to solve the more complex problem with cutoff.

We furthermore present adaptions of the algorithm of [Drezner et al., 1991, *Transportation Science* 25(3), 183–187] to our setting, which in certain situations are able to substantially reduce computation times as demonstrated in a simulation study. The sensitivity with respect to the cutoff value is also studied, which allows us to provide an algorithm that efficiently solves the problem simultaneously for all $C > 0$.

1 Introduction

For a given finite set $\mathcal{A} \subseteq \mathbb{R}^k$, a metric d on $\mathbb{R}^k$ and some $q \geq 1$, we study the location problem

$$\min_{z \in \mathbb{R}^k} \sum_{a \in \mathcal{A}} \min\{d(a,z)^q, C\}, \tag{1}$$

where $C > 0$ is a cutoff parameter. Additionally we allow the option *not* to choose any location in $\mathbb{R}^k$ at a fixed cost per point in $\mathcal{A}$. Without the cutoff C this problem is known as *Weber* problem, *one-median* problem, *minisum* problem, *Fermat-Torricelli* problem or (generalized) *barycenter* problem. In this paper we call an optimal solution to the problem a *barycenter*. The barycenter problem is among the best studied problems in location theory, see [LNdG20] for recent surveys of existing results and new developments in the field. Many results exist for different metrics and for various extensions. The problem (1) introduces the following two extensions to the classic problem.

The first extension is the cutoff C (as in [DMW91]) which makes the resulting barycenter more robust against outliers. For $q = 1$ and e.g. $d = \ell_1$ this robustness is naturally given, but for other distances, outliers can have a huge effect on the location of the barycenter.

[*] Work supported by DFG RTG 2088.

[†] raoul.mueller@uni-goettingen.de

[‡] Institute for Mathematical Stochastics, University of Göttingen, 37077 Göttingen, Germany.

[§] Technische Universität Kaiserslautern, Faculty for Mathematics, and Fraunhofer Institute for Industrial Mathematics ITWM, 67663 Kaiserslautern, Germany.

A barycenter can be thought of as a *typical* representative of a given set of points. The robustness helps containing this representative property even if outliers are present.
In the second extension we additionally allow the barycenter to be *empty* at a fixed cost which is constant per point in $\mathcal{A}$. This extends the representative property of the barycenter. If the given points we want to represent are so scattered that no single point can represent them, we allow for *no representation.*

Many application of the setting are possible where the cutoff C and the possibility of having an empty barycenter come in naturally. An example is a community which has to decide about building a new waste dump. Anyone can bring their domestic waste for free (but they have the cost of transportation, given by the distance to the waste dump) or have it collected for a fixed cost C. If no dump is built, the community has to pay a fixed fee per person to have their waste collected from the waste dump in a nearby city.

In the paper we investigate the two extensions and compare their solutions to the solutions of the classical problem. We identify cases in which solutions to the classical problem are still optimal for the problem with cutoff and cases in which the empty barycenter is optimal. We also treat the cutoff value C as part of the problem and investigate the sensitivity of an optimal solution w.r.t C.

Algorithmically, the barycenter problem with cutoff has already been studied, see [DMW91], [AHL12] and [Ven20] resulting in an $\mathcal{O}(n^2)$-algorithm for n being the number of existing points in $\mathcal{A}$. We refine this algorithm for the two extensions and experimentally show good computation times.

The remainder of the paper is organized as follows: In the next section we formally introduce the barycenter problem and its two extensions referring to existing literature. In Section 3 we look at some universal properties of the cutoff that will be helpful, when we investigate the relation between the barycenter problem *with* and *without* cutoff in Section 4. Here, we identify cases in which an optimal solution to the classic problem is also optimal for the problem with cutoff. Section 5 looks closer at the problem with *empty* barycenter. We analyze in which cases the empty barycenter is the best solution. In Section 6 we analyze the sensitivity of the barycenter and the objective function value in terms of the cutoff value. Section 7 sketches an application from statistical data analysis, where the barycenter problem with cutoff and empty set occurs as a subproblem when we compute a "typical" point pattern based on a given set of point patterns. In Section 8 we present a simulation study to compare the runtime of the different algorithms. The paper ends with some discussions and outlook to further research.

2 Extensions of the barycenter location problem: cutoff and empty barycenter

From now on we will always assume that we are give a finite set of locations $\mathcal{A} \subseteq \mathbb{R}^k$, $|\mathcal{A}| = n \in \mathbb{N}$. The *diameter of* $\mathcal{A}$

$$\operatorname{diam}(\mathcal{A}) := \max_{a_1, a_2 \in \mathcal{A}} d(a_1, a_2)$$

is the maximum distance between two points of $\mathcal{A}$. For technical reasons we assume that $\operatorname{diam}(\mathbb{R}^k) = \infty$. In this paper we mainly consider *norm-metrics*, i.e., distances

$$d^q(x, y) = (d(x, y))^q = \|y - x\|^q, \;\; x, y \in \mathbb{R}^k, q \geq 1$$

derived from a norm $\|\cdot\|$ (and here in particular the Euclidean norm ℓ_2 and the Manhattan norm ℓ_1), but many results are also true for general metrics d. When we consider ℓ_p norms we allow $p \in [1, \infty]$, so the *maximum norm* is permitted.

(Bar($\mathcal{A}$))**: The barycenter problem**

The classic location problem is to find a point $x \in \mathbb{R}^k$ which minimizes the sum of distances to the given points in $\mathcal{A}$:

$$\mathcal{Z}^* := \min_{x \in \mathbb{R}^k} f(x, \mathcal{A}) := \sum_{a \in \mathcal{A}} d^q(x, a). \tag{Bar($\mathcal{A}$)}$$

We call this problem *barycenter problem* and denote its set of optimal solutions by $\mathcal{X}^*$. If it is clear to which set $\mathcal{A}$ we refer to we may write $f(x)$ for its objective function instead of $f(x, \mathcal{A})$.
(Bar) has already been introduced in the 17th century by Fermat for three points a_1,a_2, and a_3 and for n weighted facilities by Weber in 1909, see, e.g., the survey [DKSW02]. Actual research concerns versions with p facilities [MBHMP07, DBMS15, MP20], barriers [Kla02], obnoxious facility location [DDS18], different types of facilities to be placed [MS98, Sch20], ordered median location problems [NP05, PRC20], location under uncertainty [CdG19], and others, see [LNdG20] and references therein for a recent overview. Here, we consider the following two extensions of (Bar).

(Bar$_C$($\mathcal{A}$))**: The barycenter problem with cutoff**

The first extension we consider is to introduce a cutoff in the distance function: Given a *cutoff value* $C > 0$, we look at the cutoff distance function

$$d^q_C(x, y) := \min\{d^q(x, y), C\}, \;\; x, y \in \mathbb{R}^k, \tag{2}$$

i.e., the distance is not increased any more once it has reached the value C. The corresponding location problem is given as

$$\mathcal{Z}^*_C := \min_{x \in \mathbb{R}^k} f_C(x, \mathcal{A}) := \sum_{a \in \mathcal{A}} \min\{d^q(x, a), C\}. \tag{Bar$_C$($\mathcal{A}$)}$$

It is called *barycenter problem with cutoff.* We denote its set of optimal solutions by $\mathcal{X}^*_C$. Again, if the set $\mathcal{A}$ is known, we may write $f_C(x)$ instead of $f_C(x, \mathcal{A})$. The problem is a special case of the *Weber problem with limited distance* from [DMW91]. The latter problem allows different cutoff values λ_i for each of the existing facilities while in (Bar$_C$) all $\lambda_i = C$. It has also been studied in [AHL12] and [Ven20]. Recently, (Bar$_C$($\mathcal{A}$)) has been investigated within a statistical application, namely for finding barycenters for point patterns, see [MSM20] or Section 7. At the end of Section 3 we present the algorithm of [DMW91] for solving the Weber problem with limited distances.

Related work includes [FAAJ17] where the authors consider a discrete version of a barycenter problem in which they restrict how many existing points have to be within the cutoff distance. The problem is solved by a global optimization algorithm based on a decomposition of the plane into regions for which we know which given points are within the cutoff value C. A reversed approach in which one tries to cover as many points as possible within a given threshold value C and measures only the distance to the non-covered points is investigated in [BJKS15].

Note that the cutoff does not change the properties of the distances. Definiteness, symmetry and triangle inequality are still satisfied.

Lemma 1 ([MSM20])**.** *If d is a metric then d_C is also a metric.*

(Bar$_{C,\alpha}(\mathcal{A})$)**: The empty barycenter as an option.**

The largest distance to the new facility in (Bar$_C(\mathcal{A})$) is bounded by the cutoff value C. In the second extension we go a step further and allow to place no facility (represented as $x = \emptyset$). In this case, each demand point $a \in \mathcal{A}$ has to pay a price of $C' := \alpha \cdot C$ for some given $\alpha > 0$. In order to formulate this setting as location problem, we extend the metric space by the empty set $\emptyset$ for which we define a constant "distance" between $x = \emptyset$ and any other point $y \in \mathbb{R}^k \cup \emptyset$, namely

$$d^q_{C,\alpha}(\emptyset, y) = \begin{cases} \alpha \cdot C & \text{if } y \neq \emptyset \\ 0 & \text{if } y = \emptyset \end{cases}$$

and leave $d^q_{C,\alpha}(x, y) = d^q_C(x, y)$ for all $x, y \neq \emptyset$. The corresponding location problem

$$\mathcal{Z}^*_{C,\alpha} := \min_{x \in \mathbb{R}^k \cup \{\emptyset\}} f_{C,\alpha}(x, \mathcal{A}) := \sum_{a \in \mathcal{A}} d^q_{C,\alpha}(x, a), \tag{Bar$_{C,\alpha}(\mathcal{A})$}$$

is called *barycenter problem with empty set (and cutoff).* We denote its set of optimal solutions by $\mathcal{X}^*_{C,\alpha}$ and call $\xi^* = \emptyset$ the *empty barycenter.* The problem has recently been introduced and motivated in [MSM20] but to the best of our knowledge otherwises not been studied.

Adding the empty barycenter to the metric space with cutoff distance d_C does not change the properties of the metric space if $\alpha \geq \frac{1}{2}$.

Lemma 2. *$M' = (\mathbb{R}^k \cup \emptyset, d_{C,\alpha})$ is a metric space if and only if $\alpha \geq \frac{1}{2}$.*

Proof. The definiteness and the symmetry of the metric $d_{C,\alpha}$ directly hold also for $\emptyset$. The triangle inequality

$$d_{C,\alpha}(x, y) + d_{C,\alpha}(y, z) \geq d_{C,\alpha}(x, z) \tag{3}$$

can be shown by checking all possible cases:

- If $x, y, z \in M$, (3) is satisfied since d_C is a metric.
- If $x = y = z = \emptyset$, or if exactly two of the three points x, y, z are $\emptyset$, (3) follows directly from the definition of $d_{C,\alpha}$.
- For only $x = \emptyset$ the triangle inequality holds since $d_{C,\alpha}(y, z) \geq 0$. The same holds for $z = \emptyset$.

We are left with the case that $y = \emptyset$ and $x, z \in M$. In this case, (3) transfers to

$$\alpha C + \alpha C \geq d_{C,\alpha}(x, z) = d_C(x, z). \tag{4}$$

We have to show two directions:

$\Longrightarrow$ Let (4) hold for all $x, z \in \mathbb{R}^k$. Choose x, z with $d(x, z) > C$, i.e., $d_C(x, z) = C$. Then we receive $2\alpha C \geq C$, i.e., $\alpha \geq \frac{1}{2}$.

$\Longleftarrow$ Let $\alpha \geq \frac{1}{2}$. Then we have that $d_C(x, z) \leq C \leq 2\alpha C$ and (4) is satisfied.

□

Note that the proof also shows that for a strictly increasing metric d (such as ℓ_1 or ℓ_2) without cutoff, $(\mathbb{R}^k \cup \{\emptyset\}, d_{\infty,\alpha})$ never is a metric space since (3) is always violated for $y = \emptyset$ and $d(x, z) > 2\alpha C$. This is the reason why we do not treat location problems with empty set, but without cutoff.

Relations between $(\mathrm{Bar}(\mathcal{A}))$, $(\mathrm{Bar}_C(\mathcal{A}))$, and $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$

We summarize a few observations on the relations between the optimal values of the three problems.

Lemma 3. *We always have*

(i) $\mathcal{Z}^*_{C,\alpha} \leq \mathcal{Z}^*_C \leq \mathcal{Z}^*$

(ii) $\mathcal{Z}^*_C \leq (n-1) \cdot C$

(iii) $\mathcal{Z}^*_{C,\alpha} \leq C \cdot \min\{n-1, n \cdot \alpha\}$.

Proof.

(i) Since $d^q_C(x, y) \leq d^q(x, y)$ we get $f_C(x) \leq f(x)$ for all $x \in \mathbb{R}^k$, hence also $\min_{x \in \mathbb{R}^k} f_C(x) \leq \min_{x \in \mathbb{R}^k} f(x)$ and $\mathcal{Z}^*_C \leq \mathcal{Z}^*$ holds.

Furthermore, the empty barycenter increases the set of feasible solutions, i.e., $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$ is a relaxation of $(\mathrm{Bar}_C(\mathcal{A}))$. We conclude $\mathcal{Z}^*_{C,\alpha} \leq \mathcal{Z}^*_C$.

(ii) Let $a \in \mathcal{A}$. This is a feasible barycenter with objective value of $f_C(a) = \sum_{a' \in \mathcal{A}} d^q_C(a, a') \leq 0 + (n-1) \cdot C$, hence an upper bound on $(\mathrm{Bar}_C(\mathcal{A}))$.

(iii) The empty barycenter is feasible and has an objective value of $f_{C,\alpha} = n \cdot \alpha \cdot C$, hence an upper bound on $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$. Together with (i) and (ii), the result follows.

□

3 Exploiting the local structure of Bar$_C$

We start with some general properties of the cutoff. To this end we need some further notation.

Definition 4. *Let $x \in \mathbb{R}^k$. Then*

- $\text{active}_C(x) := \{a \in \mathcal{A} : d^q(x,a) \leq C\}$ *denotes the* active points *w.r.t x and C, and*
- $\text{const}_C(x) := \mathcal{A} \setminus \text{active}_C(x)$ *denotes the* constant points *w.r.t x and C, i.e., the points whose distances remain locally constant.*

When we know the value of C we just write active(x) *and* const(x)*.*

We can now split the objective function into an active and a constant part,

$$\begin{aligned} f_C(x,\mathcal{A}) &= f_C(x, \text{active}_C(x)) + f_C(x, \text{const}_C(x)) \\ &= \sum_{a \in \text{active}_C(x)} d^q(x,a) + C \cdot |\text{const}_C(x)| \\ &= f(x, \text{active}_C(x)) + C \cdot |\text{const}_C(x)|. \end{aligned} \tag{5}$$

This decomposition gives us a first basic result showing that the barycenter problem with cutoff is equivalent to a problem of type (Bar), but w.r.t a subset of the existing points. The following Lemma is an extension of Lemma 2 of [DMW91], who proved this result for $d \in \{\ell_1, \ell_2\}$, but it is visible that the proof works more generally. For the sake of completeness we present a proof for any metric d and any $q \geq 1$.

Lemma 5. *Let $\xi^* \in \mathcal{X}_C^*$ be an optimal solution to (Bar$_C(\mathcal{A})$). Then the following hold:*

(i) *ξ^* is an optimal solution to* (Bar(active(ξ^*)))*.*

(ii) *All optimal solutions for* (Bar(active(ξ^*))) *are optimal solutions to* (Bar$_C(\mathcal{A})$) *i.e.,* $\mathcal{X}^*(\text{active}(\xi^*)) \subseteq \mathcal{X}_C^*(\mathcal{A})$*.*

Proof.

ad (i) Let $\xi^* \in \mathbb{R}^k$ be a minimizer of $f_C(x,\mathcal{A})$, but assume $f(y, \text{active}(\xi^*)) < f(\xi^*, \text{active}(\xi^*))$ for some $y \in \mathbb{R}^k$. Due to (5) we then receive $f_C(y,\mathcal{A}) < f_C(\xi^*,\mathcal{A})$, a contradiction to the optimality of ξ^*.

ad (ii) For the second statement, take $\eta^* \in \mathcal{X}^*(\text{active}(\xi^*))$. We consider active$(\xi^*)$ and const$(\xi^*) = \mathcal{A} \setminus \text{active}(\xi^*)$ separately:

$$\begin{aligned} f_C(\xi^*, \text{active}(\xi^*)) &= f(\xi^*, \text{active}(\xi^*)) = f(\eta^*, \text{active}(\xi^*)) \;\geq\; f_C(\eta^*, \text{active}(\xi^*)) \\ f_C(\xi^*, \text{const}(\xi^*)) &= \sum_{a \in \text{const}(\xi^*)} C \;\geq\; \sum_{a \in \text{const}(\xi^*)} \min\{d^q(\eta^*, a), C\} \;=\; f_C(\eta^*, \text{const}(\xi^*)) \end{aligned}$$

and together we receive that $f_C(\eta^*, \mathcal{A}) \leq f_C(\xi^*, \mathcal{A})$, hence η^* is also optimal.

□

Algorithm 1: Algorithm for $(\mathrm{Bar}_C(\mathcal{A}))$, based on [DMW91].

```
Input  : Set $\mathcal{A} = \{a_1, \ldots, a_n\}$, cutoff $C > 0$
Output: A barycenter $\xi^*$ of $(\mathrm{Bar}_C(\mathcal{A}))$, objective function value $\mathcal{Z}_C^*$
Set $\xi^* \leftarrow a_1$, $\mathcal{Z}_C^* \leftarrow \infty$;
for $i \leftarrow 1$ to $(n-1)$ do
    Inner Loop;
end
return $\xi^*, \mathcal{Z}_C^*$

Inner Loop:
for $j \leftarrow (i+1)$ to $n$ do
    if $d^q(a_i, a_j) \leq 2^q \cdot C$ then
        Compute the centers $c_1, c_2$ of the balls with radius $C$ that fulfill
          $d^q(c_1, a_i) = d^q(c_1, a_j) = d^q(c_2, a_i) = d^q(c_2, a_j) = C$;
        for $k \leftarrow 1$ to 2 do
            Set $S := \{a \in \mathcal{A} \mid d^q(c_k, a) \leq C\}$;
            Compute for the four sets $S, S \setminus \{a_i\}, S \setminus \{a_j\}, S \setminus \{a_i, a_j\}$ the barycenters
              $\xi_1, \ldots \xi_4$ and the corresponding objective function values $d_1, \ldots d_4$;
            $l \leftarrow \min_{l \in \{1,\ldots,4\}} d_l$;
            if $d_l < \mathcal{Z}_C^*$ then
                $\mathcal{Z}_C^* \leftarrow d_l$, $\xi^* \leftarrow \xi_l$;
            end
        end
    end
end
```

This result is one of the main ideas needed for Algorithm 1 and its improved versions which are described next. We state the approach of [DMW91] for our special case of cut off distances. Note that the versions of [AHL12] and [Ven20] are not relevant for this setting.

The following observations are true in the plane. We know from Lemma 5 that any optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$ is a solution to $(\mathrm{Bar}(A))$ for a subset $A \subseteq \mathcal{A}$. Algorithm 1 uses brute force to calculate the optimal solutions for these subsets. But instead of enumerating all theoretically possible $2^n - 1$ subsets, [DMW91] use the following geometric observation to cut down the number of subsets to look at: Say we have an optimal solution $\xi^* \in \mathcal{X}_C^*$. Set $A := \mathrm{active}(\xi^*)$. A is contained in a (2-dimensional) ball around ξ^* with radius C. This ball can be "moved" so that two points, a_1, a_2 of $\mathcal{A}$ lie on the circumference of the ball and all points of A are still inside. We hence can restrict our search to all balls with radius C that are defined by two points of $\mathcal{A}$ on its circumference. [DMW91] proved that there are at most $\mathcal{O}(n^2)$ of these balls, so we only need to solve $(\mathrm{Bar}(A))$ for $\mathcal{O}(n^2)$ subsets. The arguments of the proof hold for all norm metrics d and all $q \geq 1$, although [DMW91] did not state these cases explicitly.

Theorem 6 ([DMW91])**.** *Let $\mathcal{A} \subseteq \mathbb{R}^2$, let d be a norm-metric and say we can solve $(\mathrm{Bar}(\mathcal{A}))$ in $h(n)$ time. Then Algorithm 1 solves the problem $(\mathrm{Bar}_C(\mathcal{A}))$ in $\mathcal{O}(n^2 \cdot h(n))$ time.*

Proof. The result was proven in [DMW91] for $d \in \{\ell_1, \ell_2\}$. The proof is based on two arguments. First the solution of $(\mathrm{Bar}_C(\mathcal{A}))$ is a solution to (Bar(A)) for some $A \subseteq \mathcal{A}$. And second the number of these subsets A we need to check for the optimal solution is of order n^2. Lemma 5 states the first argument for any metric d and any $q \geq 1$. And the second argument follows directly from the proof of Theorem 1 in [DMW91]. The argument in the proof works for any ball defined by a norm-metric d. A ball that is defined by d^q only differs in its radius from a ball that is defined by d. So the number of candidate subsets A is of order n^2 for any norm-metric d and any $q \geq 1$. □

Remark. *Algorithm 1 is presented only for finite subsets $\mathcal{A}$ of $\mathbb{R}^2$. The method of cutting down the number of $2^n - 1$ theoretically possible subsets of $\mathcal{A}$ to a polynomial number of subsets also works in $\mathbb{R}^k$ for $k \geq 3$. The k-dimensional ball with radius C is uniquely defined by k points that define a $k-1$ dimensional hyperplane. For $k = 2$ we need two points that are not identical, for $k = 3$ we need three points that are not collinear. With the same arguments as for the the 2-dimensional case, the number of candidate sets is bound by $\mathcal{O}(\binom{n}{k}) = \mathcal{O}(n^k)$. Therefore Algorithm 1 can be solved in k dimensions in $\mathcal{O}(n^k \cdot h(n))$ time.*

We additionally suggest the following improvement that is obtained by replacing lines 1-5 by Algorithm 2: Instead of investigating all a_i, a_j with $d(a_i, a_j) \leq 2^q C$ we sort out points a_i for which we can be sure that they will not lead to a solution that improves our current best objective function value. The sorting out is based on the following lemmas.

Lemma 7. *Let $a \in \mathcal{A}$ and $A' := \{a' \in \mathcal{A} \mid d^q(a, a') > 2^q C\}$. Let $x \in \mathbb{R}^k$ s.t. $d^q(x, a) \leq C$. Then*

(i) $A' \subseteq \mathrm{const}(x)$,

(ii) $f_C(x) \geq C \cdot |A'|$.

Proof.

ad (i) Take $a' \in A'$. We know by definition of A' that

$$d^q(a, a') > 2^q C \Rightarrow d(a, a') > 2\sqrt[q]{C}.$$

We also know that

$$d^q(a, x) \leq C \Rightarrow d(a, x) \leq \sqrt[q]{C}.$$

With the triangle inequality we get that

$$\underbrace{d(a, a')}_{>2\sqrt[q]{C}} - \underbrace{d(a, x)}_{\leq \sqrt[q]{C}} \leq d(x, a').$$

The left side of the inequality is $> \sqrt[q]{C}$. Hence $\sqrt[q]{C} < d(x, a') \Rightarrow C < d^q(x, a') \Rightarrow a' \in \mathrm{const}(x)$.

ad (ii) Now we know from (i) that $A' \subseteq \mathrm{const}(x)$. Therefore $f_C(x) \geq C \cdot |\mathrm{const}(x)| \geq C \cdot |A'|$.

□

Lemma 8. *Let $a \in \mathcal{A}$, let $z = f_C(x)$ for some $x \in \mathbb{R}^k$. Let $A' := \{a' \in \mathcal{A} \mid d^q(a, a') > 2^q C\}$ like in Lemma 7. If $C \cdot |A'| > z$ then no set $S \ni a$ constructed in Algorithm 1 will lead to an optimal solution $\xi^* \in \mathcal{X}^*_C$.*

Proof. Let $\xi^* \in \mathcal{X}_C^*$ be an optimal solution to $(\text{Bar}_C(\mathcal{A}))$. We know from Lemma 5 that any optimal $\xi^* \in \mathcal{X}_C^*$ is a solution of $(\text{Bar}(\text{active}(\xi^*)))$. Say we have constructed a set S containing $a \in \mathcal{A}$ in line 10 of Algorithm 1. Suppose there is a $\xi \in \mathcal{X}_C^*$, such that $\text{active}(\xi) = S$. We know then that $f_C(\xi) \geq C \cdot |\mathcal{A} \setminus S|$. With Lemma 7(i) and the construction of the set S we know that $\mathcal{A} \setminus S \supseteq A'$. Therefore $f_C(\xi) \geq C \cdot |\mathcal{A} \setminus S| \geq C \cdot |A'| > z \geq \mathcal{Z}_C^*$. Therefore no set S that we construct in Algorithm 1 that contains the point a, nor any of its subsets are the active set of an optimal solution. □

Algorithm 2: Improvement of Algorithm 1

```
   Input  : Set $\mathcal{A} = \{a_1, \dots, a_n\}$, cutoff $C > 0$
   Output: A barycenter $\xi^*$ of $(\text{Bar}_C(\mathcal{A}))$, objective function value $\mathcal{Z}_C^*$
 1 Set $\xi^* \leftarrow a_1$, $\mathcal{Z}_C^* \leftarrow \infty$, $\mathcal{B} \leftarrow \mathcal{A}$;
 2 for $i \leftarrow 1$ to $(n-1)$ do
 3     continue $\leftarrow$ true;
 4     $m \leftarrow |\{a \in \mathcal{B} \mid d^q(a_i, a) \leq 2^q C\}|$;
 5     if $(n-m) \cdot C \geq \mathcal{Z}_C^*$ then
 6         $\mathcal{B} \leftarrow \mathcal{B} \setminus \{a_i\}$;
 7         continue $\leftarrow$ false;
 8     end
 9     if continue then
10         Inner Loop;
11     end
12 end
13 return $\xi^*, \mathcal{Z}_C^*$
```

Theorem 9. *Let* $\mathcal{A} \subseteq \mathbb{R}^2$*, let* d *be a norm-metric and say we can solve* $(\text{Bar}(\mathcal{A}))$ *in* $h(n)$ *time. Then Algorithm 2 solves the problem* $(\text{Bar}_C(\mathcal{A}))$ *in* $\mathcal{O}(n^2 \cdot h(n))$ *time.*

Proof. We have to prove two things: first the runtime and second the correctness.
First: The calculation of m takes $\mathcal{O}(n)$ time and hence does not increase the runtime of the algorithm.
Second: For the correctness we have to prove that although we skip the *Inner Loop* for some a_i we still compute the optimal solution. Lemma 8 implies that we can skip any point a_i if for $A' := \{a' \in \mathcal{A} \mid d^q(a_i, a') > 2^q C\}$ the value $|A'| \cdot C = (n-m) \cdot C$ is larger than the current best objective function value. In addition the proof of Lemma 8 yields that $a_i \notin \text{active}(\xi^*)$ for $\xi^* \in \mathcal{X}_C^*$. It is therefore justified to permanently remove a_i from the candidate set of potentially active points in line 4 of Algorithm 2. □

Later in Section 5, where we solve $(\text{Bar}_{C,\alpha}(\mathcal{A}))$, we can further reduce the computation time with the knowledge that also any point a_i for which $m \leq (1-\alpha) \cdot n$ can also be disregarded, compare Lemma 19.

Other consequences of Lemma 5

Apart from its algorithmic implication, Lemma 5 has several other consequences since it transfers properties that depend on the local structure from $(\text{Bar}(\mathcal{A}))$ to $(\text{Bar}_C(\mathcal{A}))$. This holds for properties which are only based on the metric and on the existing facilities. Such properties then also hold for subsets of the existing facilities, and in particular for active(ξ^*) where ξ^* is an optimal solution of ($\mathrm{Bar}_C(\mathcal{A})$). A first example of such a condition which will be used later in Theorem 17 is the property (conv) for ($\mathrm{Bar}(\mathcal{A})$) that there always exists an optimal solution ξ^* to the barycenter problem which is contained in the convex hull $\mathrm{conv}(\mathcal{A})$ of the existing facilties.

$$\text{For any subset } A \subseteq \mathcal{A} : \mathcal{X}^*(A) \cap \mathrm{conv}(\mathcal{A}) \neq \emptyset \tag{conv}$$

If (conv) holds for ($\mathrm{Bar}(\mathcal{A})$) then it also holds for ($\mathrm{Bar}_C(\mathcal{A})$).

Lemma 10. *If* (conv) *then* $\mathcal{X}_C^*(\mathcal{A}) \cap \mathrm{conv}(\mathcal{A}) \neq \emptyset$.

Proof. Let $\xi^* \in \mathcal{X}_C^*(\mathcal{A})$. From Lemma 5 we know that ξ^* is optimal for Bar(active(ξ^*)). There exists $\eta^* \in \mathcal{X}^*(\text{active}(\xi^*))$ with $\eta^* \in \mathrm{conv}(\text{active}(\xi^*)) \subseteq \mathrm{conv}(\mathcal{A})$. From the second part of Lemma 5 we know that $\eta^* \in \mathcal{X}_C^*(\mathcal{A})$. Together, the result follows. □

Condition (conv) is satisfied for many location problems. We list cases in which it holds below.

- In $\mathbb{R}^2$ (and in $\mathbb{R}^1$) (conv) holds for all distances d dervied from norms [Pla84].
- For $k > 2$, (conv) only holds in general if d is a norm which is linearly equivalent to the ℓ_2-norm [Pla84].
- (conv) holds for $d = \ell_2^2$, since the (unique) optimal solution of ($\mathrm{Bar}(\mathcal{A})$) is in this case the coordinate-wise mean of the points in $\mathcal{A}$ which is always contained in $\mathrm{conv}(\mathcal{A})$.

There are many other examples of conditions which can be transferred form ($\mathrm{Bar}(\mathcal{A})$) to ($\mathrm{Bar}_C(\mathcal{A})$). Among them are:

- There exists a finite candidate set for ($\mathrm{Bar}_C(\mathcal{A})$) if d is derived from a polyhedral norm. This candidate set can be found by using the intersection points of the fundamental directions.
- For problems ($\mathrm{Bar}_C(\mathcal{A})$) with restricted set R all optimal solutions are either optimal solutions for the unrestricted problem or are contained in the boundary of R.

Above we stated the property (conv), which will enable us to make a connection between ($\mathrm{Bar}(\mathcal{A})$) and ($\mathrm{Bar}_C(\mathcal{A})$) in Theorem 17. We now state a weaker assumption that also allows for a connection between ($\mathrm{Bar}(\mathcal{A})$) and ($\mathrm{Bar}_C(\mathcal{A})$).

$$\begin{aligned}&\text{There exists a ball } B = B(x,r) \text{ such that:}\\ &\text{for all } A \subseteq \mathcal{A} : \mathcal{X}^*(A) \cap B \neq \emptyset\end{aligned} \tag{B}$$

Lemma 11. *Condition* (B) *implies that* $\mathcal{A} \subseteq B$.

Proof. For any point $a \in \mathcal{A}$ the singleton $\{a\}$ is a subset of $\mathcal{A}$. The optimal solution ξ^* of ($\mathrm{Bar}(\{a\})$) is $\xi^* = a$. Therefore B must contain all points $a \in \mathcal{A}$. □

The set $\mathcal{A}$ is finite. That means there are only finitely many subsets $A \subseteq \mathcal{A}$ and a ball B that fulfills (B) always exists. We consider the smallest one.

Definition 12. *We define a* ball *with center* $x \in \mathbb{R}^k$ *and radius* $r > 0$ *by* $B := B(x, r) := \{y \in \mathbb{R}^k \mid d(x, y) \leq r\}$. *We denote by*
$B_0 = B(x_0, r_0)$ *a* smallest ball *(in terms of radius) that fulfills* (B)*.*

We can now make a connection between the optimal objective function values $\mathcal{Z}^*$ and $\mathcal{Z}_C^*$.

Theorem 13. *If* $2r_0 \leq \sqrt[q]{C}$*, then* $\mathcal{Z}^* = \mathcal{Z}_C^*$

Proof. Take an optimal solution ξ^* of $(\mathrm{Bar}_C(\mathcal{A}))$ inside B_0 and an optimal solution η^* of $(\mathrm{Bar}(\mathcal{A}))$ inside B_0. Both solutions must exist due to (B). We prove that $\mathcal{Z}^* = f(\eta^*) = f_C(\xi^*) = \mathcal{Z}_C^*$.
We know that $\xi^* \in B_0$. That means for every $a \in \mathcal{A}$ that $d(\xi^*, a) \leq d(\xi^*, x_0) + d(x_0, a) \leq 2r_0$. Then $d^q(\xi^*, a) \leq (2r_0)^q \leq C$ and hence $f_C(\xi^*) = f(\xi^*)$. Analogously we get that $f(\eta^*) = f_C(\eta^*)$. From the optimality of both ξ^* and η^* it follows that $f_C(\xi^*) \leq f_C(\eta^*)$ and $f(\eta^*) \leq f(\xi^*)$ and therefore $f(\xi^*) = f_C(\xi^*) \leq f_C(\eta^*) = f(\eta^*) \leq f(\xi^*)$, i.e. $f(\xi^*) = f(\eta^*)$. □

We will see in the next section in Lemma 14 that $\mathcal{Z}^* = \mathcal{Z}_C^*$ also implies that $\mathcal{X}^* \subseteq \mathcal{X}_C^*$.

4 Comparing Bar$_C$ and Bar

In this section we have a closer look at the barycenter problem with cutoff in comparison to the barycenter problem without cutoff. In general, problem $(\mathrm{Bar}(\mathcal{A}))$ has an easier structure than problem $(\mathrm{Bar}_C(\mathcal{A}))$. While $(\mathrm{Bar}(\mathcal{A}))$ is a convex problem for every norm-metric d, the cutoff destroys convexity and can, e.g., lead to non-connected optimal solution sets. In the following we identify conditions under which solving $(\mathrm{Bar}(\mathcal{A}))$ gives us the objective function value of $(\mathrm{Bar}_C(\mathcal{A}))$ or even an optimal solution of the latter.

(a) (Bar) has the same objective function value as (Bar_C), i.e. $\mathcal{Z}^* = \mathcal{Z}_C^*$.

(b) Any solution to (Bar) is a solution to (Bar_C), i.e., $\mathcal{X}^* \subseteq \mathcal{X}_C^*$.

If the second condition holds then it is sufficient to solve (Bar). We first show that condition (b) already follows from (a) (but not vice versa), so either condition is useful.

Lemma 14. *If condition* (a) *holds then* (b) *holds as well.*

Proof. Let $\xi^* \in \mathcal{X}^*$ be an optimal solution to $(\mathrm{Bar}(\mathcal{A}))$ and $\eta^* \in \mathcal{X}_C^*$ an optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$. From condition (a) we know that $f(\xi^*) = f_C(\eta^*)$. In order to show that $\xi^* \in \mathcal{X}_C^*$, we compute

$$f_C(\xi^*) \leq f(\xi^*) = f_C(\eta^*) \leq f_C(\xi^*).$$

Consequently, $f_C(\xi^*) = f_C(\eta^*)$ and hence $\xi^* \in \mathcal{X}_C^*$. □

With this result we know, that as soon as condition (a) holds, we can solve the problem (Bar) and automatically get a solution to (Bar_C).
The implication $(b) \Rightarrow (a)$ is not true in general, as a simple one-dimensional example shows:

Example 1 (Counterexample to (b) $\Rightarrow$ (a))**.** *Let 4 points in $\mathbb{R}$ be given, $a_1 = a_2 = 0$, $a_3 = C + \varepsilon$ and $a_4 = -C - \varepsilon$ and let $q = 1$. The solution to both problems,* ($\mathrm{Bar}(\mathcal{A})$) *and* ($\mathrm{Bar}_C(\mathcal{A})$) *is $\mathcal{X}^* = \mathcal{X}^*_C = \{0\}$, but $\mathcal{Z}^* = 2(C + \varepsilon) > 2C = \mathcal{Z}^*_C$.*

We now show that for a set $\mathcal{A}$ with a large diameter, condition (a) is not met. To this end, we use that for two points, a barycenter is given by their arithmetic mean.

Lemma 15. *For two points $a, b \in \mathbb{R}^k$, for any ℓ_p-norm and for all $q \geq 1$, a minimizer of $\|a - x\|^q + \|b - x\|^q$ is $x = \frac{a+b}{2}$.*

Proof. For ℓ_p-norms this can be treated as a one-dimensional problem, since the optimal solutions are on the line between a and b. W.l.o.g say $a = 0, b = 1$. Every other case follows by scaling. The resulting objective function is $f(x) = x^q + (1 - x)^q$ whose minimium is attained at $\bar{x} = \frac{1}{2}$. □

The next theorem identifies cases in which condition (a) does not hold; i.e., cases in which the objective function value of ($\mathrm{Bar}_C(\mathcal{A})$) is strictly smaller than that of ($\mathrm{Bar}(\mathcal{A})$).

Theorem 16. ($\mathrm{Bar}_C(\mathcal{A})$) *has a strictly smaller objective function value than* ($\mathrm{Bar}(\mathcal{A})$) *in the following two cases:*

(i) $\mathrm{diam}(\mathcal{A}) > 2C$, $q = 1$ *and d is a metric,*

(ii) $\mathrm{diam}(\mathcal{A}) > 2\sqrt[q]{C}$, $q > 1$ *and d is derived from an ℓ_p-norm.*

Proof. Since $\mathcal{A}$ is finite there exist two points $a, b \in \mathcal{A}$ such that $d(a, b) = \mathrm{diam}(\mathcal{A}) > 2\sqrt[q]{C}$.

ad (i): for any point $x \in \mathbb{R}^k$ the triangle inequality directly gives $d(x, a) + d(x, b) \geq d(a, b) > 2C$.

ad (ii): we use Lemma 15, namely that a minimizer of $\|a - x\|^q + \|b - x\|^q$ is given by $\bar{x} = \frac{a+b}{2}$. We receive that for any point $x \in \mathbb{R}^k$:

$$\begin{aligned} d^q(a, x) + d^q(x, b) &= \|a - x\|^q + \|x - b\|^q \geq \|a - \bar{x}\|^q + \|\bar{x} - b\|^q \\ &= 2^{-(q-1)}\|a - b\|^q > 2^{-(q-1)} \cdot 2^q C = 2C. \end{aligned}$$

In both cases, at least one of the distances $d^q(a, x)$ or $d^q(x, b)$ is larger than the cutoff C for any $x \in \mathbb{R}^k$. This holds especially for a barycenter $\xi^* \in \mathcal{X}^*$. Therefore

$$f(\xi^*) = \sum_{a \in \mathcal{A}} d^q(\xi^*, a) > \sum_{a \in \mathcal{A}} \min\{d^q(\xi^*, a), C\} = f_C(\xi^*).$$

Let $\eta^* \in \mathcal{X}^*_C$. We know that $f_C(\xi^*) \geq f_C(\eta^*)$. Hence,

$$\mathcal{Z}^* = f(\xi^*) > f_C(\xi^*) \geq f_C(\eta^*) = \mathcal{Z}^*_C.$$

□

The next theorem identifies a setting in which condition (a) and hence also condition (b) hold, i.e., in which ($\mathrm{Bar}(\mathcal{A})$) can be used to obtain an optimal solution to ($\mathrm{Bar}_C(\mathcal{A})$).

Theorem 17. *Consider a location problem* ($\mathrm{Bar}(\mathcal{A})$) *which satisfies property* (conv). *If $\mathrm{diam}(\mathcal{A}) \leq \sqrt[q]{C}$, then $\mathcal{Z}^* = \mathcal{Z}^*_C$ and $\mathcal{X}^* \subseteq \mathcal{X}^*_C$.*

Proof. Let $\xi^* \in \mathcal{X}^*$ be an optimal solution to $(\mathrm{Bar}(\mathcal{A}))$ and $\eta^* \in \mathcal{X}^*_C$ be an optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$. By (conv) and Lemma 10 we may choose both, ξ^* and $\eta^* \in \mathrm{conv}(\mathcal{A})$.
For any point $x \in \mathrm{conv}(\mathcal{A})$ and for all $a \in \mathcal{A}$ we have that $d(x,a) \leq \mathrm{diam}(\mathcal{A}) \leq \sqrt[q]{C}$, therefore $d^q(x,a) \leq C$ and thus $f(x) = f_C(x)$. In particular, we receive

$$\begin{aligned} f(\xi^*) &= f_C(\xi^*) \\ f_C(\eta^*) &= f(\eta^*). \end{aligned}$$

Hence we obtain

$$f_C(\xi^*) = f(\xi^*) \leq f(\eta^*) = f_C(\eta^*) \leq f_C(\xi^*),$$

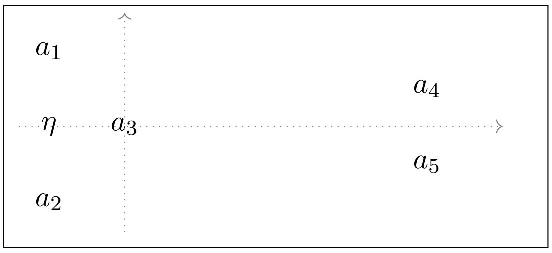
i.e., $\mathcal{Z}^* = f(\xi^*) = f_C(\eta^*) = \mathcal{Z}^*_C$. By Lemma 14, we also get $\mathcal{X}^* \subseteq \mathcal{X}^*_C$. □

We hence know that $\mathcal{Z}^* = \mathcal{Z}^*_C$ if the diameter of the set $\mathcal{A}$ is smaller or equal to $\sqrt[q]{C}$ and that $\mathcal{Z}^* > \mathcal{Z}^*_C$ if the diameter is greater than $2\sqrt[q]{C}$. The following examples demonstrate that for the remaining cases, $\sqrt[q]{C} < \mathrm{diam}(\mathcal{A}) \leq 2\sqrt[q]{C}$ everything may happen.

Example 2. *[Example where (a) holds for $q = 1$ and $C < \mathrm{diam}(\mathcal{A}) \leq 2C$] Let $n+2$ points in $\mathbb{R}$ be given, $n \geq 2$. Say $a_1 = \ldots = a_n = 0, a_{n+1} = C - \varepsilon, a_{n+2} = -(C-\varepsilon)$. The diameter of this set is $\mathrm{diam}(\mathcal{A}) = d(a_{n+1}, a_{n+2}) = 2C - 2\varepsilon$. Now $\mathcal{X}^* = \mathcal{X}^*_C = \{0\}$ and $\mathcal{Z}^* = \mathcal{Z}^*_C$.*

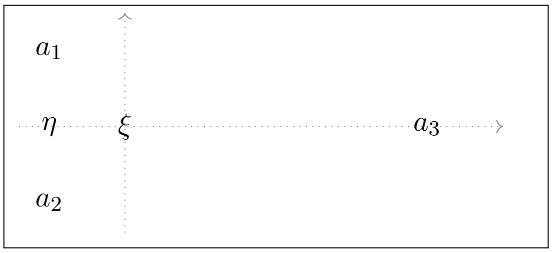
In this case the solution of (Bar) is also a solution of (Bar_C). But there are simple examples, where $\mathcal{X}^* \nsubseteq \mathcal{X}^*_C$.

Example 3. *[Example where (b) does not hold for $q = 1$, $d = \ell_1$ and $C < \mathrm{diam}(\mathcal{A}) \leq 2C$]*



Let 5 points a_1 to a_5 in $\mathbb{R}^2$ be given with coordinates $a_1 = (-\frac{\varepsilon}{2}/\frac{\varepsilon}{2}), a_2 = (-\frac{\varepsilon}{2}/-\frac{\varepsilon}{2}), a_3 = (0/0), a_4 = (C-\frac{\varepsilon}{4}/\frac{\varepsilon}{4}), a_5 = (C-\frac{\varepsilon}{4}/-\frac{\varepsilon}{4})$ for some $\varepsilon > 0$. The placement is pictured above. We get the following distances: $d(a_3,a_4) = d(a_3,a_5) = C$, $d(a_1,a_3) = d(a_2,a_3) = \varepsilon$. The diameter of this set is $d(a_1,a_5) = d(a_2,a_4) = d(a_1,a_3) + d(a_3,a_5) = C + \varepsilon$. The optimal solution of $(\mathrm{Bar}(\mathcal{A}))$ is a_3 with $\mathcal{Z}^* = 2C + 2\varepsilon = f_C(a_3)$. For $\eta = (-\frac{\varepsilon}{2}/0)$ we get $f_C(\eta) = \frac{3\varepsilon}{2} + 2C < f_C(a_3)$. Thus $a_3 \in \mathcal{X}^*$ but $a_3 \notin \mathcal{X}^*_C$.

Example 4. *[Example where (b) does not hold for $q = 1$, $d = \ell_2$ and $C < \mathrm{diam}(\mathcal{A}) \leq 2C$]*



Let 3 *points a_1 to a_3 in $\mathbb{R}^2$ be given with coordinates $a_1 = (-\frac{C}{2}/\frac{\sqrt{3}C}{2}), a_2 = (-\frac{C}{2}/-\frac{\sqrt{3}C}{2}), a_3 = (C/0)$. They form an equilateral triangle with sidelength $\sqrt{3}C$. The placement is sketched above. The diameter of this set is equal to the length of one side of the triangle which is larger than C but smaller than $2C$. The optimal solution to* $(\mathrm{Bar}(\mathcal{A}))$ *is $\xi = (0/0)$ with $\mathcal{Z}^* = 3C = f_C(\xi)$. But for $\eta = (-\frac{C}{2}/0)$ we get $f_C(\eta) = (\sqrt{3}+1)C < f_C(\xi)$. Therefore $\xi \in \mathcal{X}^*$ but $\xi \notin \mathcal{X}^*_C$. (Optimal solutions to* $(\mathrm{Bar}_C(\mathcal{A}))$ *would be each of the points a_1 to a_3 with $\mathcal{Z}^*_C = 2C$.)*

Example 5. *[Example where (b) is not true for $q = 2$, $d = \ell_2$ and $\sqrt{C} < \mathrm{diam}(\mathcal{A}) \leq 2\sqrt{C}$] Take the same situation as in Example 4, but for simplicity set $C = 1$. The diameter of this set is equal to the length of one side of the triangle which is $\sqrt[2]{3} \approx 1.732$ and therefore smaller than 2. The optimal solution to* $(\mathrm{Bar}(\mathcal{A}))$ *is $\xi = (0/0)$ with $\mathcal{Z}^* = 3 = f_C(\xi)$. But for $\eta = \left(-\frac{1}{2}/0\right)$ we get $f_C(\eta) = \left(2 \cdot \left(\frac{\sqrt{3}}{2}\right)^2 + 1\right) = \frac{5}{2} < f_C(\xi)$. Therefore $\xi \in \mathcal{X}^*$ but $\xi \notin \mathcal{X}^*_C$.*

We remark that in the last three examples above we have $f(\xi^*) = f_C(\xi^*)$ for the (respective) optimal solution ξ^* to $(\mathrm{Bar}(\mathcal{A}))$ but still $\xi^* \notin \mathcal{X}^*_C$, i.e., this solution is not optimal for $(\mathrm{Bar}_C(\mathcal{A}))$.

In the following table we summarize the results for metrics d and $q \geq 1$:

	$\mathrm{diam}(\mathcal{A}) \leq \sqrt[q]{C}$	$\sqrt[q]{C} < \mathrm{diam}(\mathcal{A}) \leq 2\sqrt[q]{C}$	$\mathrm{diam}(\mathcal{A}) > 2\sqrt[q]{C}$
(a) $\mathcal{Z}^* = \mathcal{Z}^*_C$	holds if (conv), see Thm 17	may or may not hold, see Examples 2 to 5	never for $q = 1$, never for $q > 1$ for ℓ_p-norms, see Thm 16
(b) $\mathcal{X}^* \subseteq \mathcal{X}^*_C$	holds if (conv), follows from Lem 14	may or may not hold, see Examples 2 to 5	may or may not hold, see Examples 1 and 3

Furthermore, we have seen in Theorem 13 that (a) and (b) are always true if $2r_0 \leq \sqrt[q]{C}$, where r_0 is the radius of a smallest ball B such that $\mathcal{X}^*(A) \cap B \neq \emptyset$ for all $A \subseteq \mathcal{A}$.

For a very small cutoff C relative to the distances between the points of $\mathcal{A}$ and for a large cutoff compared to the diameter of $\mathcal{A}$ we can say something about the optimal solutions to $(\mathrm{Bar}_C(\mathcal{A}))$:

Lemma 18. *Let ξ^* be an optimal solution to* $(\mathrm{Bar}_C(\mathcal{A}))$.

(i) If $C < \frac{1}{2^q} \min\limits_{a_1 \neq a_2 \in \mathcal{A}} d^q(a_1, a_2)$ we have $\xi^* \in \mathcal{A}$ and $|\mathrm{active}(\xi^*)| = 1$.

(ii) If $\sqrt[q]{C} \geq 2 \cdot \mathrm{diam}(\mathcal{A})$ we have $|\mathrm{active}(\xi^*)| = n$, *implying $\mathcal{Z}^* = \mathcal{Z}^*_C$ and $\mathcal{X}^* = \mathcal{X}^*_C$.*

Proof. (i) We show that there is no better barycenter than a point $a \in \mathcal{A}$. The cutoff is smaller than the shortest distance between two points of $\mathcal{A}$. Therefore for any $a \in \mathcal{A}$: $f_C(a, \mathcal{A}) = (n-1) \cdot C$.
Suppose $|\mathrm{active}_C(a)| = 1$ and there is a point $\xi \in \mathbb{R}^k : f_C(\xi, \mathcal{A}) < (n-1) \cdot C$, then $|\mathrm{active}(\xi)| \geq 2$. Take two different points $a_1, a_2 \in \mathrm{active}(\xi)$. Then $d(a_1, \xi) + d(\xi, a_2) \geq d(a_1, a_2) > 2\sqrt[q]{C}$. Therefore one of the distances $d(a_1, \xi), d(\xi, a_2)$ is larger than $\sqrt[q]{C}$ and thus one of the distances $d^q(a_1, \xi), d^q(\xi, a_2)$ is larger than C. This contradicts the assumption that both points are in $\mathrm{active}(\xi)$.

(ii) We prove that if $|\text{active}(\xi^*)| < n$, then ξ^* is not optimal for $(\text{Bar}_C(\mathcal{A}))$. Suppose $|\text{active}(\xi^*)| < n$. Then there is a point $a_1 \in \mathcal{A}$ such that $d^q(\xi^*, a_1) > C \geq 2^q \cdot \text{diam}(\mathcal{A})^q$ and therefore

$$d(\xi^*, a_1) > 2 \cdot \text{diam}(\mathcal{A}).$$

Thus for any $a \in \mathcal{A}$

$$d(\xi^*, a) \geq \underbrace{d(\xi^*, a_1)}_{>2\,\text{diam}(\mathcal{A})} - \underbrace{d(a_1, a)}_{\leq \text{diam}(\mathcal{A})} > \text{diam}(\mathcal{A}).$$

We know now that for all $a \in \mathcal{A}$: $d^q(\xi^*, a) \geq \text{diam}(\mathcal{A})^q$. Therefore $f_C(\xi^*) \geq n \cdot \text{diam}(\mathcal{A})^q > (n-1) \cdot \text{diam}(\mathcal{A})^q \geq f_C(a_1)$, which means ξ^* is not optimal for $(\text{Bar}_C(\mathcal{A}))$.

Thus under the conditions of *(ii)* we do have $|\text{active}(\xi^*)| = n$. So we know that ξ^* is a barycenter of all points in $\mathcal{A}$ and therefore an optimal solution to $(\text{Bar}(\mathcal{A}))$ with $\mathcal{Z}_C^* = f_C(\xi^*) = f(\xi^*) = \mathcal{Z}^*$. Since $\xi^* \in \mathcal{X}_C^*$ was arbitrary, we obtain $\mathcal{X}_C^* \subseteq \mathcal{X}^*$ and Lemma 14 implies $\mathcal{X}^* \subseteq \mathcal{X}_C^*$, hence $\mathcal{X}^* = \mathcal{X}_C^*$.

□

5 Comparing $\textbf{Bar}_C$ with $\textbf{Bar}_{C,\alpha}$

From an applied point of view it might be interesting to consider the empty barycenter as a valid solution. The barycenter of a set of points is representative for said set. Having no barycenter can then be interpreted as "the points are so widely spread, that no single point represents them".
After solving $(\text{Bar}_C(\mathcal{A}))$ it is easy to check if the empty barycenter is a better solution. But it would save computation time if we knew before the calculations that the empty barycenter *must* be better. In this section we compare $(\text{Bar}_C(\mathcal{A}))$ with $(\text{Bar}_{C,\alpha}(\mathcal{A}))$ and work out criteria under which we know that either the empty barycenter is the optimal solution to $(\text{Bar}_{C,\alpha}(\mathcal{A}))$ or that the empty barycenter cannot be the optimal solution.

If the empty barycenter is not the best solution to $(\text{Bar}_{C,\alpha}(\mathcal{A}))$ we know that the points of $\mathcal{A}$ must contain a cluster which has a certain *density*. That means that there must exist a subset $A \subseteq \mathcal{A}$ with $\text{diam}(A) \leq 2\sqrt[q]{C}$ containing at least $(1-\alpha) \cdot n$ points:

Lemma 19. *The empty barycenter is an optimal solution if there is no ball B with radius $\sqrt[q]{C}$ that contains more than $(1-\alpha) \cdot n$ points.*

Proof. Suppose such a ball does not exist. Let $\xi^* \in \mathbb{R}^k$ be an optimal solution to $(\text{Bar}_{C,\alpha}(\mathcal{A}))$. We know for all $a \in \text{active}(\xi^*)$ that $d(\xi^*, a) \leq \sqrt[q]{C}$. Therefore there exists a ball B with radius $\sqrt[q]{C}$ that contains all points of $\text{active}(\xi^*)$ and no points of $\text{const}(\xi^*)$. Since, by assumption, B can not contain more than $(1-\alpha) \cdot n$ points, we know that $\text{active}(\xi^*)$ does not contain more than $(1-\alpha) \cdot n$ points, i.e., $|\text{active}(\xi^*)| \leq (1-\alpha) \cdot n$. But then $|\text{const}(\xi^*)| \geq n - (1-\alpha) \cdot n = \alpha \cdot n$. So the points in $\text{const}(\xi^*)$ alone contribute at least $\alpha \cdot C \cdot n = f_{C,\alpha}(\emptyset, \mathcal{A})$ to $f_{C,\alpha}(\xi^*, \mathcal{A})$, hence $f_{C,\alpha}(\xi^*, \mathcal{A}) \geq \alpha \cdot C \cdot n = f_{C,\alpha}(\emptyset, \mathcal{A})$. If $\emptyset \notin \mathcal{X}_{C,\alpha}^*$, then $f_{C,\alpha}(\xi^*, \mathcal{A}) < \alpha \cdot C \cdot n$, which contradicts the optimality of ξ^*. □

In Lemma 19 we could argue with active(ξ^*) alone. If $|\text{active}(\xi^*)| \leq (1-\alpha)\cdot n$, then the empty barycenter is an optimal solution. But we do not know ξ^* and therefore active(ξ^*) before solving ($\text{Bar}_C(\mathcal{A})$). Checking if such a ball exists might in general be computationally more easy than solving ($\text{Bar}_C(\mathcal{A})$). E.g. for data in $\mathbb{R}^2$ and the Euclidean distance, i.e. $d = \ell_2$, $q = 1$, it can be checked in $\mathcal{O}(n^2)$ time if such a ball exists, see [CL86].

We further improve Algorithm 2 by using the empty barycenter as an upper bound on the optimal solution to ($\text{Bar}_{C,\alpha}(\mathcal{A})$). For the empty barycenter we know directly the value $f_{C,\alpha}(\emptyset, \mathcal{A}) = n\cdot\alpha\cdot C$ and initialize the algorithm with this value as current best solution.

Algorithm 3: Second improvement of Algorithm 1

```
Input  : Set $\mathcal{A} = \{a_1,\dots,a_n\}$, cutoff $C > 0$, $\alpha > 0$
Output: A barycenter $\xi^*$ of ($\text{Bar}_{C,\alpha}(\mathcal{A})$), objective function value $\mathcal{Z}^*_{C,\alpha}$
Set $\xi^* \leftarrow \emptyset$, $\mathcal{Z}^*_{C,\alpha} \leftarrow n\cdot\alpha\cdot C$, $\mathcal{B} \leftarrow \mathcal{A}$;
for $i \leftarrow 1$ to $(n-1)$ do
    $continue \leftarrow true$;
    $m \leftarrow |\{a \in \mathcal{B} \mid d^q(a_i, a) \leq 2^q C\}|$;
    if $(n-m)\cdot C \geq \mathcal{Z}^*_{C,\alpha}$ then
        $\mathcal{B} \leftarrow \mathcal{B} \setminus \{a_i\}$;
        $continue \leftarrow false$;
    end
    if $continue$ then
        Inner Loop;
    end
end
return $\xi^*$, $\mathcal{Z}^*_{C,\alpha}$
```

Theorem 20. *Let $\mathcal{A} \subseteq \mathbb{R}^2$, let d be a norm-metric and say we can solve* ($\text{Bar}(\mathcal{A})$) *in $h(n)$ time. Then Algorithm 3 solves the problem* ($\text{Bar}_C(\mathcal{A})$) *in $\mathcal{O}(n^2\cdot h(n))$ time.*

Proof. The runtime and the correctness of the algorithm follow directly from the proof of Theorem 9. Formally, the only difference is that Algorithm 3 is initialized with the empty barycenter as the current best solution. □

Compared with Algorithm 2 we replace the initial ξ^* in the declaration from $\xi^* \leftarrow a_1$ with $\xi^* \leftarrow \emptyset$. This is of course only better, if $f_{C,\alpha}(\emptyset,\mathcal{A}) \leq f_C(a_1,\mathcal{A})$, which implies $\alpha \leq \frac{n-1}{n}$, compare Lemma 3. We will see in the following Lemma that for α close enough to 1, the empty barycenter cannot be an optimal solution to ($\text{Bar}_{C,\alpha}(\mathcal{A})$):

Lemma 21. *If $\alpha > \frac{n-1}{n}$ then the empty barycenter is never an optimal solution.*

Proof. Referring to Lemma 3 we compare $f_{C,\alpha}(\emptyset,\mathcal{A})$, with $\alpha > \frac{n-1}{n}$, to the upper bound of $\mathcal{Z}^*_C$: $f_{C,\alpha}(\emptyset,\mathcal{A}) = \alpha\cdot n\cdot C > \frac{n-1}{n}\cdot n\cdot C = (n-1)\cdot C \geq \mathcal{Z}^*_C$.
Hence a point $\xi^* \in \mathbb{R}^k$ exists, such that $f_C(\xi^*,\mathcal{A}) < f_{C,\alpha}(\emptyset,\mathcal{A})$. □

To determine if the empty barycenter is a better solution than any solution in $\mathbb{R}^k$ before solving ($\text{Bar}_C(\mathcal{A})$), we can look at the pairwise distances between the points of $\mathcal{A}$.

If the points of $\mathcal{A}$ are close to each other compared to C, it is more likely that the cost of an empty barycenter exceeds the cost of a solution in $\mathbb{R}^k$. If on the other hand the points are far apart, it is more likely that the empty barycenter is optimal. We define the *mean pairwise distance* between points of $\mathcal{A}$ and study its relation to the optimal solution of $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$ w.r.t n, α and C.

Definition 22. *Let* $\mathcal{A} = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}^k$. *We define the* mean pairwise distance

$$\mathrm{mpd}(\mathcal{A}) := \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_C^q(a_i, a_j) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_C^q(a_i, a_j).$$

The following statements are immediately clear.

Lemma 23. *We always have*

- $0 \leq \mathrm{mpd}(\mathcal{A}) \leq C$,
- $0 \leq \mathrm{mpd}(\mathcal{A}) \leq \mathrm{diam}(\mathcal{A})^q$.

The mean pairwise distance can be computed in $\mathcal{O}(n^2)$ time. If it is 'small' compared to the cutoff C and α and n, we know that the empty barycenter can again not be an optimal solution to $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$. Let us hence study $\mathrm{mpd}_C(\mathcal{A}) := \frac{1}{C}\,\mathrm{mpd}(\mathcal{A}) \in [0,1]$ as percentage of C. We can strengthen Lemma 21 as follows:

Lemma 24. *If* $\alpha > \mathrm{mpd}_C(\mathcal{A}) \cdot \frac{n-1}{n}$, *then for at least one point* $a \in \mathcal{A} : f_{C,\alpha}(a, \mathcal{A}) < f_{C,\alpha}(\emptyset, \mathcal{A})$, *i.e. the empty barycenter is never an optimal solution.*

Proof. Suppose such a point a does not exist. We show that then $\mathrm{mpd}_C(\mathcal{A}) \cdot \frac{n-1}{n} > \alpha$: For any $i \in \{1, \ldots, n\}$: $\sum_{j=1}^{n} d_C^q(a_i, a_j) = f_{C,\alpha}(a_i, \mathcal{A}) > f_{C,\alpha}(\emptyset, \mathcal{A}) = \alpha \cdot C \cdot n$. The mean pairwise distance then is

$$\begin{aligned} \mathrm{mpd}(\mathcal{A}) &= \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_C^q(a_i, a_j) \\ &> \frac{1}{n(n-1)} \sum_{i=1}^{n} \alpha \cdot C \cdot n = \frac{1}{n(n-1)} \cdot \alpha \cdot C \cdot n^2 = \alpha \cdot C \cdot \frac{n}{n-1}, \end{aligned}$$

hence $\mathrm{mpd}_C(\mathcal{A}) \cdot \frac{n-1}{n} > \alpha$. □

We can directly transfer this result to the diameter $\mathrm{diam}(\mathcal{A})$ which we used in Section 4, since the mean pairwise distance is never larger than the diameter raised to the power q.

Corollary 25. *If* $\alpha > \frac{\mathrm{diam}(\mathcal{A})^q}{C} \cdot \frac{n-1}{n}$ *then the empty barycenter is never an optimal solution.*

Proof. From $\alpha > \frac{\mathrm{diam}(\mathcal{A})^q}{C} \cdot \frac{n-1}{n}$ it follows that $\alpha > \mathrm{mpd}_C \cdot \frac{n-1}{n}$. With Lemma 24 we know that then the emtpy barycenter is not an optimal solution to $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$. □

We can reformulate Lemma 24 and Corollary 25 to get a condition for the diameter and the mean pairwise distance for the empty barycenter not being optimal: $\mathrm{mpd}(\mathcal{A}) < \alpha \cdot C \cdot \frac{n}{n-1}$ or $\mathrm{diam}(\mathcal{A})^q < \alpha \cdot C \cdot \frac{n}{n-1}$ then $\emptyset$ is not optimal.
On the other hand we show that for small α the empty barycenter is always an optimal solution to $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$:

Lemma 26. *If* $\alpha \leq \min\left\{\frac{1}{2^q}\frac{\operatorname{diam}(\mathcal{A})^q}{n \cdot C}, \frac{1}{n}\right\}$ *then the empty barycenter is an optimal solution to* $(\operatorname{Bar}_{C,\alpha}(\mathcal{A}))$*.*

Proof. Let $\xi^* \in \mathcal{X}_C^*$ be an optimal solution to $(\operatorname{Bar}_C(\mathcal{A}))$. We know from the triangle inequality that there exists a point $a \in \mathcal{A}$ such that $d(\xi^*, a) \geq \frac{1}{2}\operatorname{diam}(\mathcal{A})$. Therefore $d_C^q(\xi^*, a) \geq \min\{\frac{1}{2^q}\operatorname{diam}(\mathcal{A})^q, C\}$ and hence $\mathcal{Z}_C^* \geq \min\{\frac{1}{2^q}\operatorname{diam}(\mathcal{A})^q, C\}$. Then

$$\begin{aligned} f_{C,\alpha}(\emptyset, \mathcal{A}) = n \cdot \alpha \cdot C &\leq \min\left\{\frac{1}{2^q}\operatorname{diam}(\mathcal{A})^q, C\right\} \\ \Leftrightarrow \alpha &\leq \min\left\{\frac{1}{2^q}\frac{\operatorname{diam}(\mathcal{A})^q}{n \cdot C}, \frac{1}{n}\right\} \end{aligned}$$

□

Remark. *When* $\operatorname{diam}(\mathcal{A}) \geq 2C$ *then* $\min\left\{\frac{1}{2^q}\frac{\operatorname{diam}(\mathcal{A})^q}{n \cdot C}, \frac{1}{n}\right\} = \frac{1}{n}$*. And thus for* $\alpha \leq \frac{1}{n}$ *and* $\operatorname{diam}(\mathcal{A}) \geq 2C$ *the empty barycenter is an optimal solution to* $(\operatorname{Bar}_{C,\alpha}(\mathcal{A}))$*.*

The following lemma and example show that for larger $\operatorname{mpd}(\mathcal{A})$ both cases, $\emptyset \in \mathcal{X}_{C,\alpha}^*$ or $\emptyset \notin \mathcal{X}_{C,\alpha}^*$, are possible.

Lemma 27. *Let* $\alpha = \frac{1}{2}, q = 1$*. Then for any* $\varepsilon > 0$ *there exists a set* $\mathcal{A} \subseteq \mathbb{R}^k$ *such that* $\operatorname{mpd}(\mathcal{A}) = \frac{1}{2} \cdot C + \varepsilon$*, but where* $f_{C,\frac{1}{2}}(\emptyset, \mathcal{A}) < f_{C,\frac{1}{2}}(x, \mathcal{A})$ *for any* $x \in \mathbb{R}^k$*.*

Proof. We construct a set $\mathcal{A} = \{a_1, \ldots, a_{2n}\} \subseteq \mathbb{R}$, where $n > \frac{C}{4\varepsilon} + \frac{1}{2}$ and set $\delta := \frac{2n(2n-1)\varepsilon - nC}{4(n-1)}$. The points have the coordinates $a_1 = -\delta, a_2 = \ldots = a_n = 0, a_{n+1} = \ldots = a_{2n-1} = C, a_{2n} = C + \delta$. Then $\operatorname{mpd}(\mathcal{A}) = \frac{n^2C + 2(n-1)\delta}{n(2n-1)} = \frac{nC}{2n-1} + \frac{2(n-1)\delta}{n(2n-1)} = \frac{C}{2} + \frac{C}{2(2n-1)} + \frac{2(n-1)\delta}{n(2n-1)}$.With the specified δ we have $\operatorname{mpd}(\mathcal{A}) = \frac{C}{2} + \varepsilon$ and since $n > \frac{C}{4\varepsilon} + \frac{1}{2}$ we have $\delta > 0$.
An optimal solution of $(\operatorname{Bar}_C(\mathcal{A}))$ is any of the points $a_2, \ldots a_{2n-1}$. The optimal objective function value $\mathcal{Z}_C^* = f_{C,\frac{1}{2}}(a_2, \mathcal{A}) = nC + \delta$ is larger than $nC = f_{C,\frac{1}{2}}(\emptyset, \mathcal{A})$. So the empty barycenter is a better solution than the best solution in $\mathbb{R}$.

□

We have seen in Lemma 24 that for $\operatorname{mpd}(\mathcal{A}) < \alpha \cdot C \cdot \frac{n}{n-1}$ the empty barycenter is not optimal. Yet, Lemma 27 proves that the mpd can be arbitrarily close to $\alpha \cdot C$ and still the empty barycenter *is* an optimal solution. The following example proves on the other hand that there exist sets with an mpd arbitrarily close to C, for which the empty barycenter is *not* an optimal solution.

Example 6. *[Example for large* $\operatorname{mpd}(\mathcal{A})$ *where* $\emptyset \notin \mathcal{X}_{C,\alpha}^*$ *for* $\alpha \geq \frac{1}{2}$*.]*
Let two points in $\mathbb{R}$ *be given,* $a_1 = 0, a_2 = \sqrt[q]{C - \varepsilon}$ *for some* $\varepsilon > 0$*. Now* $f_{C,\alpha}(a_1, \mathcal{A}) = C - \varepsilon < C \leq 2 \cdot \alpha \cdot C = f_{C,\alpha}(\emptyset, \mathcal{A})$*.*

Note that the situation of this example is the same if we replace the mpd by the *minimum* or *median* distance between points, since both values are $C - \varepsilon$.
We finally summarize our findings. We know that the empty barycenter *is* an optimal solution to $(\operatorname{Bar}_{C,\alpha}(\mathcal{A}))$ if

- there is no ball B with radius $\sqrt[q]{C}$ that contains at least $(1-\alpha) \cdot n$ points, see Lemma 19,

- $\alpha \leq \min\left\{\frac{1}{2^q}\frac{\operatorname{diam}(\mathcal{A})^q}{n \cdot C}, \frac{1}{n}\right\}$, see Lemma 26.

On the other hand we know that the empty barycenter *is not* an optimal solution if

- $\alpha > \frac{n-1}{n}$, see Lemma 21,
- $\alpha > \operatorname{mpd}_C(\mathcal{A}) \cdot \frac{n-1}{n}$ or $\alpha > \frac{\operatorname{diam}(\mathcal{A})^q}{C} \cdot \frac{n-1}{n}$, see Lemma 24 and the subsequent Corollary.

In the other cases we do not know in advance if the empty barycenter is optimal, cf. Lemma 27 and Example 6.

6 Sensitivity analysis w.r.t C

So far we have assumed that the cutoff value C is a priori specified, but there is a wide range of scenarios where this is not the case.
In the application described in Section 7 we often (but not always) know the order of magnitude of a reasonable cutoff C due to the physical reality of the data, but this typically still leaves a large interval of possible choices which may lead to very different outcomes.
In a more direct location problem setting the actual C may be determined by another player trying to maximize her profit based on knowledge of the entire function $g = [C \mapsto \min_\xi f_C(\xi, \mathcal{A})]$ (or $[C \mapsto \min_\xi f_{C,\alpha}(\xi, \mathcal{A})]$, where we still assume $\alpha > 0$ to be fixed). Taking up the waste dump example from the introduction, it may be that in the decision process the local transportation company is asked for the price C, at which it would offer to transport domestic waste to the dump. A profit maximizing choice of C depends on detailed knowledge of the function g.
This function is what we study in the present section.

Definition 28. *Let $x \in \mathbb{R}^k$ be a fixed point. Define the two functions*

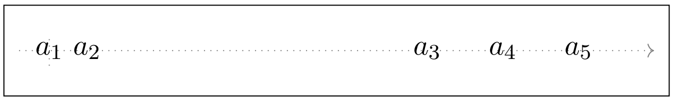
(i) $g_x : \mathbb{R}_+ \to \mathbb{R}, C \mapsto f_C(x, \mathcal{A})$

(ii) $g : \mathbb{R}_+ \to \mathbb{R}, C \mapsto f_C(\mathcal{A}) := \min_{\xi \in \mathbb{R}^k} f_C(\xi, \mathcal{A})$.

The function g_x maps the cutoff C to the objective function value $f_C(x, \mathcal{A})$, while the function g maps the cutoff to the optimal *objective function value $\mathcal{Z}_C^*$, compare with* $(\operatorname{Bar}_C(\mathcal{A}))$.

We study how the barycenter $\xi^* = \xi_C^*$ and the values of g_x and g change with changing C.

Example 7. *[Example for discontinuity of the optimal solution to* $(\operatorname{Bar}_C(\mathcal{A}))$ *w.r.t C]*



Let 5 points in $\mathbb{R}$ be given, $a_1 = 0, a_2 = 0.5, a_3 = 5, a_4 = 6, a_5 = 7$. Consider $q = 1$ and $d = |\cdot|$. The location of the points is sketched above. The following gives a complete description of $\mathcal{X}_C^*$ for various C. Note that at the boundaries of the ranges the union of the barycenters in the lower and the higher range are in $\mathcal{X}_C^*$. For $0 \leq C \leq 0.5$ any of the points a_1 to a_5 is optimal for $(\operatorname{Bar}_C(\mathcal{A}))$. *For $0.5 \leq C \leq 1.5$ any point on the interval $[a_1, a_2]$ is optimal, for $1.5 \leq C \leq 5.25$ a_4 is optimal, for $5.25 \leq C$ a_3, the barycenter of the problem without cutoff, is optimal.*

As seen in the example, the optimal solution set for $(\mathrm{Bar}_C(\mathcal{A}))$ can change abruptly in C. On the other hand we will see that g, i.e. the objective function $C \mapsto f_C(\mathcal{A})$ is quite well-behaved and can be computed efficiently.

Lemma 29. *Let $x \in \mathbb{R}^k$ be some fixed point, $\mathcal{A} = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}^k$. Sort the points increasingly by their distance to x and define $d_i := d^q(x, a_i)$, so that $d_1 \leq d_2 \leq \ldots \leq d_n$. Then the function g_x is*

(i) of the form

$$g_x(C) = \sum_{i=1}^{j} d_i + (n-j) \cdot C, \quad \textit{with } j = |\mathrm{active}_C(x)| \tag{6}$$

g_x is therefore piecewise linear with kinks in d_i.

(ii) continuous,

(iii) non-decreasing,

(iv) concave.

Proof. (i) W.l.o.g we assume that $x \neq a_i$ for all $i \in \{1, 2, \ldots, n\}$. Otherwise we eliminate the first j points from our list, where $j := \max\{i \in \{1, \ldots, n\} \mid d_i = 0\}$, use $n' = n - j$ in the proof and re-enumerate $a_{j+1}, \ldots, a_n$ to $a_1, \ldots, a_{n'}$.

By definition

$$g_x(C) = f_C(x, \mathcal{A}) = \sum_{i=1}^{n} \min\{d_i, C\}.$$

We can split this sum into sums over indices of active and constant points, as defined in Section 3. Let $m = m_C = |\mathrm{active}_C(x)|$. Then $d_1 \leq \ldots \leq d_{m_C} \leq C$ and $C < d_{m_C+1} \leq \ldots \leq d_n$ and

$$g_x(C) = \sum_{i=1}^{n} \min\{d_i, C\} = \sum_{i=1}^{m_C} d_i + \sum_{i=m_C+1}^{n} C = \sum_{i=1}^{m_C} d_i + (n - m_C) \cdot C.$$

For any $j \in \{1, \ldots, n\}$ the sum over the d_i is constant for $d_j \leq C < d_{j+1}$ so g_x is piecewise linear. The slope of the j-th line segments is $(n - j)$.

(ii) We know that g_x is piecewise linear with kinks in d_i. On these line segments, i.e. $d_i < C < d_{i+1}$, the function is continuous. We have to check for the kinks of g_x, i.e. $C = d_j$ for some j, if the two line segments for $C = d_j - \varepsilon$ and $C = d_j + \varepsilon$, $\varepsilon > 0$, intersect at $g_x(d_j)$.

Take $j \in \{1, \ldots, n\}$. For $d_{j-1} < C < d_j$: $g_x(C) = \sum_{i=1}^{j-1} d_i + \sum_{i=j}^{n} C$. For $j = 1$ the first sum is 0. For $d_j < C < d_{j+1}$: $g_x(C) = \sum_{i=1}^{j} d_i + \sum_{i=j+1}^{n} C$. For $j = n$ the second sum is 0.
For $C = d_j$: $\sum_{i=1}^{j-1} d_i + \sum_{i=j}^{n} C = \sum_{i=1}^{j-1} d_i + d_j + \sum_{i=j+1}^{n} C = \sum_{i=1}^{j} d_i + \sum_{i=j+1}^{n} C$. So the two line segments intersect and therefore is g_x continuous.

(iii) Take two cutoffs $C_1 < C_2$. Since $C_1 < C_2$ we also have $\min\{d_i, C_1\} \leq \min\{d_i, C_2\}$ for all $i \in \{1, \ldots, n\}$. Now $g_x(C_1) = \sum_{i=1}^{n} \min\{d_i, C_1\} \leq \sum_{i=1}^{n} \min\{d_i, C_2\} = g_x(C_2)$, so g_x is non-decreasing.

(iv) With larger C the cardinality m_C of $\mathrm{active}_C(x)$ increases. Therefore the slope $(n - m_C)$ of the line segments decreases with growing C, so g_x is also concave. □

We can now extend the results for g_x which hold for all $x \in \mathbb{R}^k$ to the function g.

Theorem 30. *The function g is continuous, non-decreasing and concave.*

Proof. For calculating the function g we need to solve $(\mathrm{Bar}_C(\mathcal{A}))$ for every C. For a fixed C an optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$ is a solution of $(\mathrm{Bar}(\mathcal{A}))$ for some subset $A \subseteq \mathcal{A}$, compare Lemma 5.
Since $\mathcal{A}$ is finite, there are only finitely many subsets of $\mathcal{A}$. There is therefore only a finite set S of candidates for an optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$. The function g is the minimum of the functions g_ξ, i.e. $g(C) = \min_{\xi \in S} g_\xi(C)$.
The minimum of finitely many continuous functions is continuous. The same holds for the properties "non-decreasing" and "concave". □

Recall the piecewise linear form of the function g_x. The function g as minimium of finitely many piecewise linear functions is then itself piecewise linear. The slopes of the line segments are given by the cardinalities of the sets $\mathrm{active}_C(\xi^*)$ for optimal solutions $\xi^* = \xi^*_C \in \mathcal{X}^*_C$ for the different C. Those slopes are integers between $n-1$ and 0. Since g is continuous and concave the function consists of at most n linear pieces. The slope only changes at the kinks of the function g. And only there does the cardinality of the set $\mathrm{active}(\xi^*)$ change. Let us say we have kinks at $C_1 < C_2$ and no kinks in between. Let $C \in (C_1, C_2)$. Any solution $\xi^*_C \in \mathcal{X}^*_C$ defines the same function $g_{\xi^*_C}$ on the interval (C_1, C_2). The function g has its next kink at C_2, so C_2 is the smallest value greater than C_1 where any of the functions $g_{\xi^*_C}$ can have a kink. It follows that any solution $\xi^*_C \in \mathcal{X}^*_C$ is optimal on the whole interval $[C_1, C_2]$.
But that means if we find all values C at which g has a kink, and a corresponding optimal solution ξ^* for each of those C, we have an optimal solution to $(\mathrm{Bar}_C(\mathcal{A}))$ and the value of g for any $0 \leq C < \infty$.

We describe in Algorithm 4 how we can calculate these optimal solutions ξ^* and values of g in at most $n-1$ steps, by finding the different line segments. The function $\texttt{bar}(C, \mathcal{A})$ calculates an optimal solution ξ^* of $(\mathrm{Bar}_C(\mathcal{A}))$ and the corresponding value $\mathcal{Z}^*_C$ for a given cutoff C. $\texttt{bar}(\infty, \mathcal{A})$ returns an optimal solution to $(\mathrm{Bar}(\mathcal{A}))$.
Let the set $\mathcal{S} \subseteq \{0, \ldots, n-1\}$ contain the *slopes* of the line segments that we have already found and the set $\mathfrak{O} \subseteq \{0, \ldots, n-1\} \setminus \mathcal{S}$ the slopes of the segments that we still might find, the *open* slopes. Each of the lines $l_{n-1}, \ldots, l_0$ is defined by a point $(C, \mathcal{Z}^*_C)$ and the slope i, which is indicated by its index. The algorithm will calculate (up to) n different lines. These lines are tangents for the function g. By calculating the intersection points of lines l_i and l_{i+1} we get the kinks of g and thereby the complete function $g(C)$, which is a combination of segments of the lines $l_{n-1}, \ldots, l_0$.

Theorem 31. *Say we can solve $(\mathrm{Bar}_C(\mathcal{A}))$ in $h_C(n)$ time. Then Algorithm 4 computes the function g in $\mathcal{O}(n \cdot h_C(n))$ time.*

Proof. We have to prove two things: first the runtime and second the correctness.
First: The function $\texttt{bar}(C, \mathcal{A})$ is called at most $n-1$ times. Once in line 2 and once in every iteration of the while-loop, lines 4 to 18. The while-loop is called at most $n-2$ times. The computation of the (at most) $n-1$ intersection points in line 19 to 22 is done in $\mathcal{O}(n)$ time. Together we have a runtime of $\mathcal{O}(n \cdot h_C(n) + n) = \mathcal{O}(n \cdot h_C(n))$.
Second: We know by Lemma 29 that g starts in $(0,0)$ with slope $n-1$ and will eventually get constant, taking the value $g(C) = \mathcal{Z}^*_0 = \texttt{bar}(\infty, \mathcal{A})$. So the lines l_0 and l_{n-1} defined in

```
Algorithm 4: Calculate $f_C(\mathcal{A})$ for all $C$
   Input  : The set $\mathcal{A} = \{a_1, \dots, a_n\}$
   Output: The function $g(C)$ and up to $n-1$ barycenters corresponding to the
           different values of $C$.
1  Set $\mathcal{S} \leftarrow \{0, n-1\}$, $\mathfrak{D} \leftarrow \{1, \dots, n-2\}$;
2  $(\xi_0^*, \mathcal{Z}_0^*) \leftarrow$ bar$(\infty, \mathcal{A})$, $(\xi_{n-1}^*, \mathcal{Z}_{n-1}^*) \leftarrow (a_1, 0)$;
3  Define the lines $l_0$ by the point $(0, \mathcal{Z}_0^*)$ and the slope 0 and $l_{n-1}$ by the point $(0,0)$
    and the slope $n-1$ ;
4  while $\mathfrak{D} \neq \emptyset$ do
5      Take smallest index $o$ from $\mathfrak{D}$;
6      Take the largest $i \in \mathcal{S} : i < o$ and the smallest $j \in \mathcal{S} : o < j$;
7      Calculate the intersection point $(C, y)$ of lines $l_i$ and $l_j$;
8      $(\xi_C^*, \mathcal{Z}_C^*) \leftarrow$ bar$(C, \mathcal{A})$ ;
9      if $\mathcal{Z}_C^* = y$ then
10         $\mathfrak{D} \leftarrow \mathfrak{D} \setminus \{i+1, i+2, \dots, j-1\}$;
11     end
12     else
13         Set $m \leftarrow |\text{active}(\xi_C^*)|$, $\xi_{n-m}^* \leftarrow \xi_C^*$, $\mathcal{Z}_{n-m}^* \leftarrow \mathcal{Z}_C^*$;
14         Define $l_{n-m}$ by the point $(C, \mathcal{Z}_{n-m}^*)$ and slope $n-m$;
15         $\mathfrak{D} \leftarrow \mathfrak{D} \setminus \{n-m\}$;
16         $\mathcal{S} \leftarrow \mathcal{S} \cup \{n-m\}$;
17     end
18 end
19 for $i \in \mathcal{S} \setminus \{n-1\}$ do
20     $j \leftarrow \min\{k \in \mathcal{S} \mid k > i\}$;
21     Calculate the intersection point $(C_i, g(C_i))$ of lines $l_i$ and $l_j$;
22 end
23 $(C_{n-1}, g(C_{n-1})) \leftarrow (0,0)$;
24 return $\{(C_i, g(C_i)) \mid i \in \mathcal{S}\}$, $\{\xi_i^* | i \in \mathcal{S}\}$
```

line 3 contain the outermost line segments of g and bound g above due to its concavity. We have to prove that the algorithm finds all line segments in between. Suppose we have two line segments l_i and l_j. Now we compute the intersection of those lines in line 7. We know that this intersection point (C, y) must lie on or above g by concavity. We now calculate the value $\mathcal{Z}_C^*$ for this C and get a point $(C, \mathcal{Z}_C^*)$ that is on g. There are two possibilities:

- Either $y = \mathcal{Z}_C^*$, which means the intersection point of the lines *is* already on g. But that means that the function g can have no line segments with slopes between i and j, so all these values are removed from the set $\mathfrak{D}$ in line 10.

- Or we found a point on g that is below the intersection point (C, y). For the found barycenter ξ^* we look at the cardinality m of active(ξ^*). The slope of the line l_{n-m}, which is a tangent on g, is given by the *non*-active points. So the next line l_{n-m} is given by the slope $n-m$ and the point $(C, \mathcal{Z}_c^*)$. This line is saved and we add $n-m$ to $\mathcal{S}$ and delete $n-m$ from $\mathfrak{D}$.

So with every iteration of lines 4 to 18 we either find out that l_i and l_j intersect on g or we find one new line segment that is a tangent for g. The function g is uniquely defined by these tangents. □

Note that Algorithm 4 can be parallelized. When the intersection of two lines is calculated in line 7, the problem can then be split into a subproblem to the left of this point and to the right of this point.

Remark. *Having computed g, the function $g^{(\alpha)} : \mathbb{R}_+ \to \mathbb{R}$, $C \mapsto \min_{\xi \in \mathbb{R}^k \cup \{\emptyset\}} f_{C,\alpha}(\xi, \mathcal{A})$ is easily derived, since $g^{(\alpha)}(C) = \min\{g(C), \alpha \cdot n \cdot C\}$. Provided that $\alpha < \frac{n-1}{n}$, which means $\alpha \cdot n$ is smaller than the initial slope of g, we obtain from the concavity of g and the fact that it must eventually be constant, that there is exactly one $C_0 > 0$ where the graph of g intersects with the linear function $[C \mapsto \alpha \cdot n \cdot C]$. We then have $g^{(\alpha)}(C) = \alpha \cdot n \cdot C$ to the left of C_0 and $g^{(\alpha)}(C) = g(C)$ to the right of C_0. If $\alpha \geq \frac{n-1}{n}$, we have $g^{(\alpha)} = g$ everywhere.*

7 Applications

The original motivation for investigating ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) comes from [MSM20], where two of the current authors studied barycenters of finite collections of point patterns for their use as summary statistics. We briefly describe here the relevant details, because we think that the involved concepts and their algorithmic implications may well be of interest in the context of location problems where e.g. an optimal supply chain is to be maintained to a number of companies that each have several branch offices.

For the present purpose we define a point pattern as a finite subset of $\mathbb{R}^k$ and denote the set of all such patterns by $\mathfrak{N}$. Then for given point patterns $\xi_1, \ldots, \xi_m \in \mathfrak{N}$, a barycenter is any minimizer of the Fréchet functional

$$F(\zeta) = \sum_{j=1}^{m} \tau(\xi_j, \zeta)^q \tag{7}$$

over $\zeta \in \mathfrak{N}$. Here τ is the transport-transform (TT) metric on $\mathfrak{N}$ introduced in [MSM20]. Basically, $\tau(\xi_j, \zeta)^q$ is the minimal "cost" of matching a subset of ξ_j and a subset of ζ, where each pairing of a point $x \in \xi_j$ and a point $z \in \zeta$ incurs a cost of $d(x,z)^q$ and each unmatched point of either pattern incurs a cost of $\frac{1}{2}C$.

If we consider point patterns as discrete measures by identifying $\xi = \{x_1, \ldots, x_n\}$ with $\sum_{i=1}^{n} \delta_{x_i}$ for pairwise distinct x_i, we can re-interpret the TT metric as a special case of an unbalanced Wasserstein metric, see [CPSV18] for the definition of the latter or [MSM20], Remark 3, for the full argument.

Intuitively, a barycenter can be thought of as a "typical" representative, in a sense an "average point pattern" that reflects common properties of the data point patterns. In [MSM20] barycenters were applied to point patterns of crime locations in two cities, with the goal of detecting systematic differences over the years or between different seasons. Another goal might be for planning the efficient deployment of police officers according to the time of the day (or year) and maybe other side constraints (predictive policing).

[BP21] prove that the computation of a sparse Wasserstein barycenter is $\mathcal{NP}$-hard for three point patterns with the same number of points in $\mathbb{R}^2$ and $q = 2$. In the authors' setting the barycenter can be a more general discrete finite measure (not necessarily with unit weights), but their sparseness condition limits the number of support points. There does not seem to be a direct theoretical result for our problem (7), but based on the current state of theoretical and applied research, we assume that this problem is insolvable for all practical purposes. Therefore [MSM20] proposed a heuristic algorithm based on an equivalent form of the TT metric: First fill up the point patterns $\xi_1, \ldots, \xi_m$ so that they all have the same cardinality n, say, by adding points at a single "virtual" location $\aleph \notin \mathbb{R}^k$ at distance $(\frac{1}{2}C)^{1/q}$ apart from any locations in $\mathbb{R}^k$. For $\xi_j = \{x_{j1}, \ldots, x_{jn}\}$ and $\zeta = \{z_1, \ldots, z_n\}$ (multisets since they may include $\aleph$ several times), we may then express the metric τ equivalently as

$$\tau(\xi_j, \zeta)^q = \min_{\pi \in S_n} \sum_{i=1}^{n} d'(x_{ji}, z_{\pi(i)})^q, \tag{8}$$

where S_n denotes the set of permutations on $\{1, \ldots, n\}$ and

$$d'(x,z)^q = \begin{cases} \min\bigl\{d(x,z)^q, C\bigr\} & \text{if } x, z \in \mathbb{R}^k; \\ \frac{1}{2}C & \text{if } \aleph \in \{x,z\},\ x \neq z; \\ 0 & \text{if } x = z = \aleph; \end{cases} \tag{9}$$

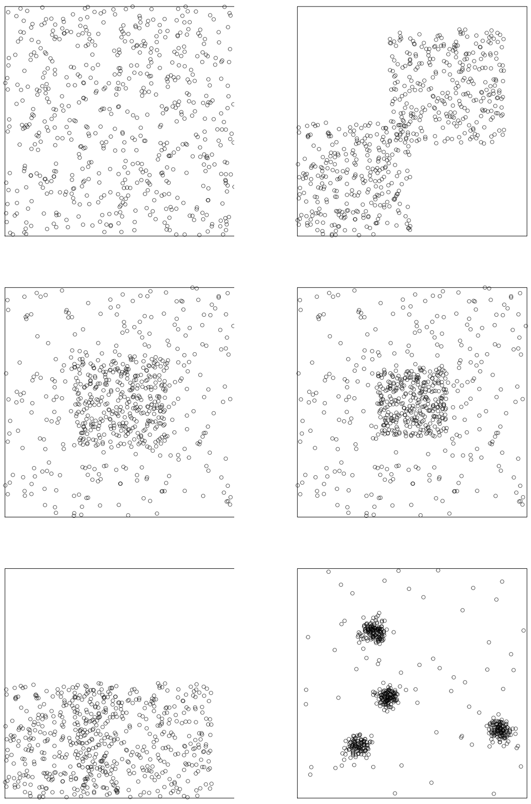
see [MSM20], Theorem 1. We may then find a local optimum of the Fréchet functional (7) by alternating between forming pairwise disjoint clusters of the form $\mathcal{C} = \{x_{1,i_1}, \ldots, x_{m,i_m}\}$, $i_1, \ldots, i_m \in \{1, \ldots, n\}$, including exactly one (maybe virtual) point from each data pattern via optimal matching, and computing suitable "centers" for each such cluster $\mathcal{C}$ by minimizing

$$f(z) = \sum_{j=1}^{m} d'(x_{j,i_j}, z)^q \tag{10}$$

over $z \in \mathbb{R}^k \cup \{\aleph\}$. In the algorithm of [MSM20] this minimization was only performed approximately, using some crude but fast heuristics. However, except for the fact that $x_{j,i_j} = \aleph$ may hold for individual j, the minimization (10) corresponds to problem $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$ with $\alpha = \frac{1}{2}$. Noting that the contribution from $x_{j,i_j} = \aleph$ is constant as long as $z \in \mathbb{R}^k$, we may therefore use a slightly adapted version of Algorithm 3 to compute the centers exactly.

8 Simulation study

For comparing Drezners algorithm with the two improvements Algorithm 2 and Algorithm 3 , we created six test scenarios of point patterns inside the unit square and compared runtimes and solutions of the algorithm. For scenarios (1) to (5) we chose rectangles and generated the coordinates of the points inside each rectangle uniformly at random, independently of one another. In scenario (2) to (5) we combined two of those rectangles. The number of points in every rectangle follows a Poisson distribution with parameters chosen in such a way that the expected number of points is 600 in each scenario. The scenarios are (from left to right, top to bottom):



1. one unit square
2. two squares with edge length 0.5 that overlap in a square of size 0.1×0.1, half of the points in each square
3. one small square with edge length 0.4 inside the unit square, half of the points in each square
4. one small square with edge length 0.3 inside the unit square, half of the points in each square
5. two rectangles overlapping on one strip of width 0.2. Height for both rectangles is 0.5, width 0.5 and 0.6, half of the points in each rectangle
6. 4 small clusters with background noise. The clusters are two-dimensional Gaussians with $\sigma = 0.025$. The cluster centers are uniformly drawn for each pattern individually. The expected number of points in the clusters is 90%, an expected number of 10% are uniformly drawn in the unit square.

We ran a simulation study with 100 patterns from each scenario to compare the three algorithms. The results are in Table 1. As expected, all three algorithms find the exact solution every time.

In these calculations we set $\alpha = 0.5$, so the cost of an empty barycenter is $C \cdot \frac{|\mathcal{A}|}{2}$. We have $d = \ell_2$ and $q = 2$. For the computation we used the publicly available R-package ttbary, see [MS21].

The runtime depends highly on the point pairs from which barycenter candidates are calculated. The larger the cutoff C, the more point pairs are taken into account and barycenter candidates have to be checked. Therefore the runtime gets higher with larger cutoff. For Algorithm 1 for the smallest cutoff $C = 0.01$ the runtime is about $4 - 12$ seconds for the 100 patterns combined. In Scenario 6 the runtime for $C = 0.01$ is about 81 seconds for the 100 runs, because even with the small cutoff due to the small clusters many barycenter candidates have to be calculated. For all the scenarios the runtime goes up to $1709 - 2574$ seconds for $C = 0.3$. We also counted how many barycenter candidates had to be calculated by this algorithm in total for each scenario. In Table 1 we compare the runtime of the two improved Algorithms 2 and 3 to the 'original' runtime and compare how many barycenter candidates could be skipped by the improved algorithms.

The column 'skipped points' presents for each scenario and cutoff the *relative* number of barycenter candidates that were skipped by this algorithm.

The values correspond to Algorithm 2/Algorithm 3/Algorithm 1 (first and second improvement and original algorithm). For example in Scenario 1, $C = 0.01$ the 0.450/1.000/0 means that Algorithm 2 was able to skip 45% of the barycenter candidates, Algorithm 3 skipped 100% of the barycenter candidates, and of course Algorithm 1 skipped nothing.

Similarly the column 'time' presents for each scenario and cutoff the *relative* time the algorithms took for the 100 point patterns compared to the runtime of the original Algorithm 1. For example in Scenario 1, $C = 0.01$ the 0.702/0.252/1 means that Algorithm 2 was about 30% faster and Algorithm 3 was about 75% faster than Algorithm 1.

We can clearly see the connection between the amount of skipped barycenter candidates and the amount of time that is saved. The first improvement, Algorithm 2, is almost never slower than the original algorithm and can for smaller cutoffs C save up to 30% of the runtime. The second improved version, Algorithm 3, is much faster than the other two. The empty barycenter is in these scenarios for small cutoffs always the optimal solution. For cutoffs up to $C = 0.1$ almost all point pairs can be skipped a priori. In scenario 1 even for $C = 0.2$ the runtime is below 1% of the runtime of the original algorithm.

	Scenario 1		Scenario 2		Scenario 3	
$C =$	skipped points	time	skipped points	time	skipped points	time
0.01	0.450/1.000/0	0.702/0.252/1	0.383/1.000/0	0.703/0.144/1	0.289/1.000/0	0.778/0.133/1
0.02	0.273/1.000/0	0.773/0.082/1	0.128/1.000/0	0.911/0.044/1	0.105/1.000/0	0.911/0.038/1
0.03	0.094/1.000/0	0.940/0.040/1	0.027/1.000/0	0.970/0.020/1	0.087/1.000/0	0.917/0.018/1
0.05	0.006/1.000/0	0.999/0.015/1	0.004/1.000/0	0.992/0.008/1	0.081/1.000/0	0.922/0.007/1
0.075	0.001/1.000/0	1.005/0.007/1	0.001/1.000/0	1.001/0.004/1	0.073/1.000/0	0.931/0.003/1
0.1	0.000/1.000/0	1.005/0.004/1	0.000/1.000/0	1.001/0.002/1	0.064/1.000/0	0.938/0.002/1
0.2	0.000/0.995/0	1.002/0.006/1	0.000/0.762/0	1.000/0.239/1	0.049/0.287/0	0.953/0.716/1
0.3	0.000/0.263/0	1.000/0.738/1	0.000/0.009/0	1.001/0.991/1	0.018/0.041/0	0.983/0.960/1
	Scenario 4		Scenario 5		Scenario 6	
$C =$	skipped points	time	skipped points	time	skipped points	time
0.01	0.189/1.000/0	0.854/0.095/1	0.397/1.000/0	0.683/0.125/1	0.056/1.000/0	0.954/0.012/1
0.02	0.071/1.000/0	0.952/0.026/1	0.137/1.000/0	0.886/0.036/1	0.023/1.000/0	0.979/0.004/1
0.03	0.070/1.000/0	0.945/0.012/1	0.044/1.000/0	0.963/0.017/1	0.012/1.000/0	0.990/0.002/1
0.05	0.072/1.000/0	0.933/0.005/1	0.008/1.000/0	0.993/0.006/1	0.015/1.000/0	0.986/0.002/1
0.075	0.076/1.000/0	0.926/0.002/1	0.004/1.000/0	0.999/0.003/1	0.080/0.987/0	0.920/0.014/1
0.1	0.079/0.981/0	0.925/0.021/1	0.002/1.000/0	0.998/0.002/1	0.181/0.953/0	0.820/0.048/1
0.2	0.116/0.231/0	0.887/0.772/1	0.000/0.171/0	1.000/0.831/1	0.127/0.528/0	0.875/0.473/1
0.3	0.028/0.032/0	0.972/0.968/1	0.000/0.000/0	1.001/1.000/1	0.042/0.132/0	0.960/0.870/1

Table 1: Comparison of the runtime of the two improved Algorithms 2 and 3 to the 'original' runtime of Algorithm 1 and how many barycenter candidates were skipped by the improved algorithms. The column 'skipped points' presents for each scenario and cutoff the *relative* number of barycenter candidates that were skipped by this algorithm. The values correspond to Algorithm 2/Algorithm 3/Algorithm 1 (first and second improvement and original algorithm). Similarly the column 'time' presents for each scenario and cutoff the *relative* time the algorithms took for the 100 point patterns compared to the runtime of the original Algorithm 1.

8.1 Consequences for the barycenter algorithm of [MSM20]

As mentioned in Section 7 problem 10, that stems from [MSM20], is identical to $(\mathrm{Bar}_{C,\alpha}(\mathcal{A}))$ with $\alpha = \frac{1}{2}$. In the algorithm of [MSM20] this problem was solved by a fast heuristic: Starting with a point $x \in \mathbb{R}^k$ we calculate $\mathrm{active}(x)$, solve $\mathrm{Bar}(\mathrm{active}(x))$ with optimal solution ξ^* and set $x \leftarrow \xi^*$. The heuristic uses the idea that is proven in Lemma 5, that the optimal solution of $(\mathrm{Bar}_C(\mathcal{A}))$ must be an optimal solution of (Bar) for some subset of $\mathcal{A}$. With this heuristic the objective function value cannot increase, since the distances to active(x) are optimized and the distances to const(x) cannot increase by definition of const(x).

An implementation of the original algorithm of [MSM20] can be found in the publicly available R package `ttbary`, [MS21]. We implemented Algorithm 3 in the algorithm of [MSM20] to replace the heuristic. In a simulation study we compared the implementation in [MS21] with our version in which the heuristic is replaced with Algorithm 3.
It turned out that doing the exact calculation instead of the heuristic for solving problem 10 does not improve the algorithm of [MSM20] in general. In the algorithm the size of $\mathcal{A}$ in ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) depends on the number of point patterns. The set $\mathcal{A}$ consists of exactly one point (including $\aleph$, see Section 7) of every pattern. We compared the runtime and the resulting objective function value (cost) of the computed pseudo-barycenters. For the three 'groupsizes' of 20, 50 and 100 point patterns per group we created 600 groups each. Both algorithms had the same input for each of the 1800 groups. In our tests about half of the costs with the exact solutions of problem 10 were smaller and half of the costs were larger compared to the heuristic. At the same time the runtime for the algorithm with the exact subroutine for 10 is about 3.5, 13.5 or 42 times larger for the groupsizes of 20, 50 and 100 respectively. Since the heuristic is a lot faster and does not yield a worse solution we recommend to stay with the original version of the algorithm as it is presented in [MSM20].

9 Discussions and outlook

In this paper we presented the problems ($\mathrm{Bar}(\mathcal{A})$), also known as the Weber problem, and the extension ($\mathrm{Bar}_C(\mathcal{A})$), which is related to a problem studied by [DMW91]. Additionally we introduced the new barycenter problem ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$), where we extend the classic problem by the option to have an empty solution. In Sections 4 and 5 we investigated under which conditions an optimal solution of ($\mathrm{Bar}(\mathcal{A})$) is also an optimal solution to ($\mathrm{Bar}_C(\mathcal{A})$) or ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$). We also investigated under which conditions an optimal solution to ($\mathrm{Bar}(\mathcal{A})$) cannot be an optimal solution to ($\mathrm{Bar}_C(\mathcal{A})$) or ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$). Most results are based solely on the geometric structure of the dataset, like the diameter of the set or the mean pairwise distance between its points. The summaries of the results can be found in Table 4 and the statements thereafter and at the end of Section 5.

For the average problem we typically do not know if we can reduce ($\mathrm{Bar}_C(\mathcal{A})$) or ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) to ($\mathrm{Bar}(\mathcal{A})$). We presented two improvements of the algorithm introduced by [DMW91] to solve ($\mathrm{Bar}_C(\mathcal{A})$) and ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) more efficiently. We furthermore gave an algorithm for solving ($\mathrm{Bar}_C(\mathcal{A})$) simultaneously for *all* $C \geq 0$ by solving $\mathcal{O}(n)$ problems of type ($\mathrm{Bar}_C(\mathcal{A})$) for specified values of C.

For future research it might be interesting to generalize the improved algorithms to the original problem stated by [DMW91], who allowed different cutoffs for every point.
Another interesting topic is to find new criteria to determine beforehand if solving ($\mathrm{Bar}(\mathcal{A})$) is sufficient. Another algorithmic idea is to split the original problem into subproblems that can be solved independently, where one optimal solution of the subproblems is guaranteed to be the optimal solution of the original problem. One could also study how ($\mathrm{Bar}_C(\mathcal{A})$) simplifies for special cases like the ℓ_1-metric, where we can optimize separately over the k dimensions. These findings could help to solve the problems ($\mathrm{Bar}_C(\mathcal{A})$) and ($\mathrm{Bar}_{C,\alpha}(\mathcal{A})$) faster in the future.

References

[AHL12] D. Aloise, P. Hansen, and L. Liberti. An improved column generation algorithm for minimum sum-of-squares clustering. *Mathematical Programming*, 131(1-2):195–220, 2012.

[BJKS15] J. Brimberg, H. Juel, M.-C. Körner, and A. Schöbel. On models for continuous facility location with partial coverage. *Journal of the Operational Research Society*, 66(1):33–43, 2015.

[BP21] S. Borgwardt and S. Patterson. On the computational complexity of finding a sparse Wasserstein barycenter. *Journal of Combinatorial Optimization*, 41(3):736–761, 2021.

[CdG19] I. Correia and F. Saldanha da Gama. Facility location under uncertainty. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 8, pages 185–213. Springer, 2019.

[CL86] B. M. Chazelle and D. T. Lee. On a circle placement problem. *Computing*, 36(1):1–16, 1986.

[CPSV18] L. Chizat, G. Peyré, B. Schmitzer, and F. Vialard. Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609, 2018.

[DBMS15] Z. Drezner, J. Brimberg, N. Mladenović, and S. Salhi. New heuristic algorithms for solving the planar p-median problem. *Computers and Operations Research*, 62(C):296–304, 2015.

[DDS18] T. Drezner, Z. Drezner, and A. Schöbel. The Weber obnoxious facility location model: A big arc small arc approach. *Computers and Operations Research*, 98:240–250, 2018.

[DKSW02] Z. Drezner, K. Klamroth, A. Schöbel, and G. Wesolowsky. The Weber problem. In Z. Drezner and H.W. Hamacher, editors, *Facility Location - Applications and Theory*, chapter 1, pages 1–36. Springer, 2002.

[DMW91] Z. Drezner, A. Mehrez, and G. O. Wesolowsky. The facility location problem with limited distances. *Transportation Science*, 25(3):183–187, 1991.

[FAAJ17] I. F. Fernandes, D. Aloise, D. J. Aloise, and T. P. Jeronimo. A polynomial-time algorithm for the discrete facility location problem with limited distances and capacity constraints. *Brazilian Journal of Operations & Production Management*, 14(2):136–144, 2017.

[Kla02] K. Klamroth. *Single-Facility Location Problems with Barriers*. Springer Series on Operations Research. Springer, 2002.

[LNdG20] G. Laporte, S. Nickel, and F. Saldanha da Gama, editors. *Location Science*. Springer International Publishing, Switzerland, 2020.

[MBHMP07] N. Mladenović, J. Brimberg, P. Hansen, and J.A. Moreno-Pérez. The p-median problem: a survey of metaheuristic approaches. *European Journal of Operational Research*, 179(3):927–939, 2007.

[MP20] A. Marin and M. Pelegrin. p-median problems. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 2, pages 25–50. Springer, 2020.

[MS98] H. Martini and A. Schöbel. Median hyperplanes in normed spaces — a survey. *Discrete Applied Mathematics*, 89:181–195, 1998.

[MS21] R. Müller and D. Schuhmacher. *ttbary: Barycenter Methods for Spatial Point Patterns*, 2021. R package version 0.2-0. `https://CRAN.R-project.org/package=ttbary`.

[MSM20] R. Müller, D. Schuhmacher, and J. Mateu. Metrics and barycenters for point pattern data. *Statistics and Computing*, 30(4):953–972, 2020.

[NP05] S. Nickel and J. Puerto. *Location Theory: A unified approach.* Springer, 2005.

[Pla84] F. Plastria. Localization in single facility location. *European Journal of Operational Research*, 18(2):215–219, 1984.

[PRC20] J. Puerto and A.M. Rodriguez-Chia. Ordered median location problems. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 10, pages 261–302. Springer, 2020.

[Sch20] A. Schöbel. Locating dimensional facilities in a continuous space. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 7, pages 143–184. Springer, 2020.

[Ven20] P. Venkateshan. A note on “The facility location problem with limited distances”. *Transportation Science*, 54(6):1439–1445, 2020.