

Facility Location in the Phylogenetic Tree Space

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

der Georg-August-Universität Göttingen

im Promotionsprogramm „PhD School of Mathematical Sciences“ (SMS)

der Georg-August University School of Science (GAUSS)

vorgelegt von

Marco Botte

aus Fritzlar

Göttingen, 2019

Betreuungsausschuss

Prof. Dr. Anita Schöbel, seit 1.1.19 Fachbereich Mathematik, Technische Universität Kaiserslautern, vorher Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

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

Mitglieder der Prüfungskommission

Referentin: Prof. Dr. Anita Schöbel, seit 1.1.19 Fachbereich Mathematik, Technische Universität Kaiserslautern, vorher Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Korreferent: Prof. Dr. Stephan Huckemann, Institut für Mathematische Stochastik, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Jun.-Prof. Dr. Anja Fischer, Juniorprofessur Management Science, Technische Universität Dortmund

Prof. Dr. Preda Mihailescu, Mathematisches Institut, Georg-August-Universität Göttingen

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

Prof. Dr. Max Wardetzky, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 28.02.2019

Acknowledgements

First of all I want to thank my supervisor Prof. Anita Schöbel for very fruitful discussions, the constant support throughout the years and for giving me the possibility to work on many different interesting fields of optimization. I also want to thank Prof. Stephan Huckemann for the interesting joint collaboration with international researchers and for introducing interesting points of view on aspects of phylogenetic trees.

Moreover, I thank all members of the work group for the pleasant atmosphere we have in the institute. My special thanks go to Julius for the congenial time in the office and also for proofreading parts of this work.

Last but not least, I wholeheartedly thank my family and Maria for their encouragement and support, especially for always believing in me.

Contents

1	Introduction	1
2	Construction and Properties of the Tree Space	7
2.1	Splits, Compatibility and Orthants	7
2.2	The Geodesic Distance	16
2.2.1	Support and Parametrization of a Geodesic	23
2.3	Modeling the Species Tree Problem in Tree Space	37
3	Location Theory in Tree Space	41
3.1	Introduction to Location Theory	41
3.2	Location Problem Models in Tree Space	43
3.2.1	Split Sets of Optimal Solutions	44
4	Solving Tree Space Location Problems by Transformations to Euclidean Location Problems	49
4.1	Location Problems in \mathcal{T}_3	49
4.2	Adjacent Orthants	51
4.3	Completely Incompatible Orthants	54
4.4	Location Problems with a Fixed Gate Point at the Origin	63
4.4.1	Median Problem	64
4.4.2	Center Problem	66
4.4.3	Fréchet Problem	93
5	Balance Point Algorithms for the Median Problem in Tree Space	99
5.1	Bounds for the Median Problem	100
5.2	The Global Balance Point Heuristic	124
5.2.1	The Balance Point Algorithm	126
5.2.2	Preprocessing	128
5.2.3	Pseudocode Formulation of the Heuristic	129
5.2.4	Experiments	131
5.3	Local Convergence of the BPA	142
5.3.1	The Block-wise Coordinate Descent Method	142
5.3.2	Convergence to a Stationary Point	158
5.3.3	Optimality of Stationary Points and Balance Points	178

5.4	Restrictions of the BPA	187
6	Real Data Example	193
6.1	Data Set	193
6.2	Experiments	195
6.2.1	PPA Experiments	195
6.2.2	BPA Experiments	201
6.3	Summary of Results and Comparison to other Algorithms	207
7	Conclusion	211

1 Introduction

Phylogenetic trees are diagrams with a tree structure that depict the relations of evolutionary history between a certain set of existing species to be investigated. Ever since Ernst Haeckel coined the term *phylogeny* in 1869 as ‘genesis and evolution of a phylum’, where ‘genesis’ translates as ‘origin’ and ‘phylum’ as ‘race’, phylogenetic trees are a part of biologists’ attempts to classify existing species according to their evolutionary history, see Figure 1.1 for an example. This classification is based on shared characteristics and genes.

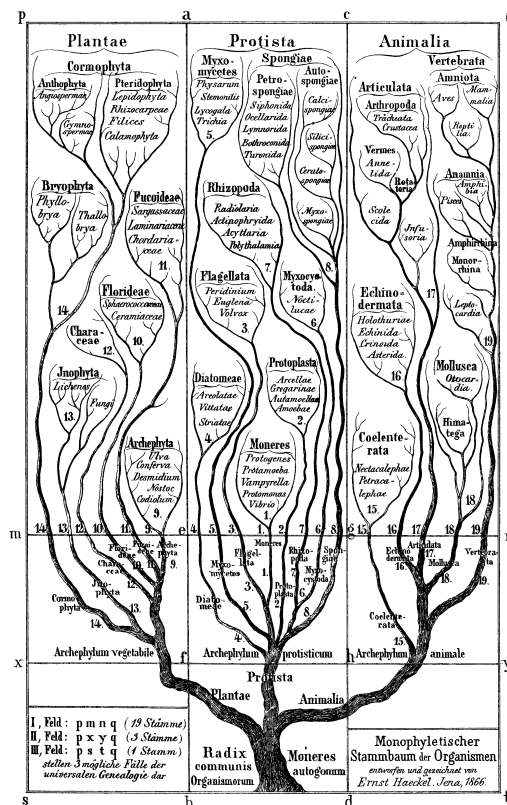


Figure 1.1: Tree of life according to Ernst Haeckel, cf. [Hae].

In contrast to Figure 1.1, a phylogenetic tree nowadays always comes with a *topology*,

which is the tree structure as depicted, but also with positive edge weights, describing the amount of mutations of genes between two nodes of the tree. Phylogenetics is an active field of research because the evolutionary history of a given set of species is often still not agreed upon by biologists. Hence, the so called *species tree*, which is the phylogenetic tree depicting the actual evolution of a given set of species, is usually unknown and is subject to study.

The Species Tree Problem

As there is no straightforward way to obtain the species tree, many methods have been proposed that hypothesize how the species tree could look like. All of these approaches use sets of phylogenetic trees as input in order to give a hypothesis for the species tree, thereby trying to incorporate shared features of the given trees into the species tree.

The phylogenetic trees that are used as data for such methods are called *gene trees* and are obtained as follows: Specific genes that the investigated species share are tagged with a biological marker. To this end one uses gene sequences that are empirically known to be good guesses of the species tree. Then the gene sequences are aligned to be as ‘parsimonious’ as possible. There exist many different approaches to infer trees from these aligned gene sequences, e.g., bayesian methods or bootstrap procedures that were introduced to the field of phylogenetics by [Fel85] marking a milestone in phylogenetic inference. The result of such a method is a phylogenetic tree depicting the relation between the investigated species based on the information of a single gene. These trees are called *gene trees*.

Applying this method, one often obtains different trees for different genes which shows that it is not possible to directly infer the *species tree* from a single gene tree. [Mad97] show a specific case where the *topology*, i.e., the tree structure, of the gene tree does not coincide with the topology of the species tree and [DR06] even show that this need not be the case for the “most likely gene tree” as well. Nevertheless, since it is impossible to directly determine the species tree, one tries to infer as much information from these gene trees as possible in order to develop reasonable hypotheses for the species tree.

Having performed this acquisition of gene trees for several genes, one receives a sample set of phylogenetic gene trees. Inferring the species tree from such a set of gene trees is what is often referred to as the *species tree problem*, or ‘gene tree/species tree problem’, as in [PC97]. One of the earliest works on this field is [AI72], where Adams develops the notion of a *consensus tree* that is supposed to convey the information of the gene trees as good as possible in order to get a reasonable candidate for the species tree. After consensus trees have been introduced, many different methods to find the species tree have been proposed, but they are mostly educated guesses or

heuristic rules. Popular methods include the majority rule consensus tree [MM81] and strict consensus methods [MMN83] that are probably widely used due to their simple computation rules. For an overview of consensus trees and related methods see [Bry03]. Naturally, these heuristic methods are limited in describing the actual species tree, which is also discussed in the literature, see for example [BDS91] and [Nel93] as a reply. Thus, there was an urgent need for more sophisticated techniques which have been developed since then with the hope to better infer the species tree from a given set of gene trees.

A lot of research in the younger history involves mathematical, mostly statistical methods to find the species tree. A large part of this mathematical research is based on [BHV01], since it provides the framework for a new systematic approach to phylogenetic tree related problems. Billera, Holmes and Vogtmann were able to define the metric space \mathcal{T}_n , which contains all possible phylogenetic trees for a fixed set of existing species $\{1, \dots, n\}$. $\{1, \dots, n\}$ are the leaves of the trees in \mathcal{T}_n . In the following we will refer to the space \mathcal{T}_n as *tree space*. \mathcal{T}_n is a metric space of ‘global non-positive curvature’. This directly implies two important features for mathematical modeling: Firstly, it is a metric space containing all possible phylogenetic trees on n leaves, i.e., it is a suitable model space for the species tree problem, since the given data and the solution to the problem are contained in this space. Secondly, there does not only exist a distance measure between any given pair T_1, T_2 of trees in \mathcal{T}_n , which is required by many algorithms, but there even exists a unique path from X to T whose length equals the distance $d(T_1, T_2)$, where d is the metric on \mathcal{T}_n . These are necessary properties to formulate and tackle the species tree problem in \mathcal{T}_n .

In this model setting, natural candidates for the species tree are some sort of ‘average’ or ‘centroid’ trees. The question of finding a *centroid* from a set of sample trees has already been posed in [BHV01] and several possible notions of a centroid of a sample set in a metric space with global non-positive curvature are mentioned. Unfortunately, all given notions of center points actually require computing the geodesic between two trees, or at least the distance of two trees. So even though [BHV01] show that the above-mentioned shortest path between two trees, called geodesic exists and is unique, they were not able to propose a method how to efficiently calculate the distance between two points in \mathcal{T}_n . This shortcoming strongly encouraged follow-up research in this direction. The computation of the distance and the geodesics in \mathcal{T}_n has been extensively studied in [Owe08, Owe11] leading to the milestone of a polynomial time algorithm calculating the distance and the parametrization of the geodesic for two given trees, see [OP11].

Now, having the metric space \mathcal{T}_n and the possibility to efficiently calculate geodesics at hand one is able to use several of the concepts of centroids that [BHV01] introduced to get new hypotheses on the species tree. For example, [MOP15] translate

the algorithm of Sturm that was already mentioned by [BHV01] into the tree space setting and applied it to several data sets. As it turns out, though, the computation of the center points remains an open problem in practice. Sturm’s algorithm theoretically converges to the mean but for high-dimensional and large data sets the method was not able to converge to a pre-specified termination condition since its calculation time was exceeded. So there is still need for improved procedures as the existing methods are of converge slowly (Sturm’s algorithm is sublinear, see [MOP15]) and not tractable for larger instance sizes or even yield the wrong species tree, [MOP15].

As an attempt to make the computation of the mean more efficient, [MOP15] further investigated the structure of \mathcal{T}_n and developed special algorithms, e.g., gradient descent methods that are based on interior point or penalty methods. After all there are still practical and theoretical problems of these approaches, since analytical properties, such as optimality criterions and differentiability, do not exist on specific subsets of \mathcal{T}_n . An up-to-date review on *Fréchet means* in tree space is presented in [BO17], both pointing out the strengths and weaknesses of the concept and giving concrete numeric studies.

Another mathematical line of research that recently evolved is to find a different model space for the phylogenetic problems. To this end, the so-called space of ultrametric trees, equipped with the tropical metric is investigated in [YZZ17, LMY18]. This research needs to be carried out in depth in order to evaluate which advantages and disadvantages this model space has.

Contribution and Structure of the Thesis

As we have seen, there exists a lot of research in the field of phylogenetics including numerous approaches on finding the species tree. Latest mathematical studies mainly concentrate on finding new model spaces or different and faster ways to compute the Fréchet mean to do statistics on the model spaces, but neither can the outcome of these approaches be anticipated nor if the result of these approaches will solve the species tree problem, at least to biologists’ satisfaction.

In this thesis we introduce a different point of view on the species tree problem by incorporating “Location Theory”: We interpret the set of gene trees as *facilities*, as usual for *location problems*. Facility Location problems are optimization problems that are motivated from real-world and economic problems and are usually investigated in Euclidean space. So on one hand this work extends the field of research of Location Theory beyond its usual scope. On the other hand, we aim at exploiting the specific local structure of \mathcal{T}_n to receive some location problems which may be solved within a Euclidean setting. By doing so we then can make use of known algorithms and results from Location Theory in Euclidean space in order to try solving

the original location problems in tree space, thereby connecting these two fields of research.

The thesis is structured as follows (also confer to Figure 1.2). In Chapter 2 we introduce the phylogenetic tree space \mathcal{T}_n as defined in [BHV01]. In Section 2.2 we give a detailed description of the geodesic distance and the parametrization of geodesics in tree space, which are the key tools to work with location problems in \mathcal{T}_n .

In Section 2.3 motivate our location theoretical point of view on the species tree problem in more detail before we build the bridge from the species tree problem to Location Theory in Chapter 3. We introduce three location problems in tree space that yield different notions of ‘centroids’ of a given set of trees as hypotheses for the species tree. We also formulate general results regarding optimal solutions for these three problems in Subsection 3.2.1.

Chapter 4 illustrates first approaches to solve these tree space location problems by providing reformulations and solution methods for special cases.

The central problem of the thesis is tackled in Chapter 5. The goal is to find a median of a given set of points in tree space and we present an algorithm to determine a median that is based on a local Euclidean improvement strategy. This algorithm is called the *Balance Point Algorithm*, which is a heuristic, but for which we prove convergence under certain assumptions in Section 5.3. As the Balance Point Algorithm is a *local* improvement procedure, it is necessary to find a good neighborhood in which the algorithm is started in order to obtain good results. In order to find a good neighborhood, we develop bounds for specific subsets of \mathcal{T}_n in Section 5.1 to determine auspicious subsets of \mathcal{T}_n , where the Balance Point Algorithm may be applied. In Section 5.2 we formulate a heuristic that determines such auspicious subsets in a preprocessing procedure and then applies the Balance Point Algorithm for the remaining subsets.

Finally, we illustrate how the Balance Point Algorithm works on a real data set in Chapter 6 and discuss its results in comparison to other methods before the thesis is summarized in Chapter 7.

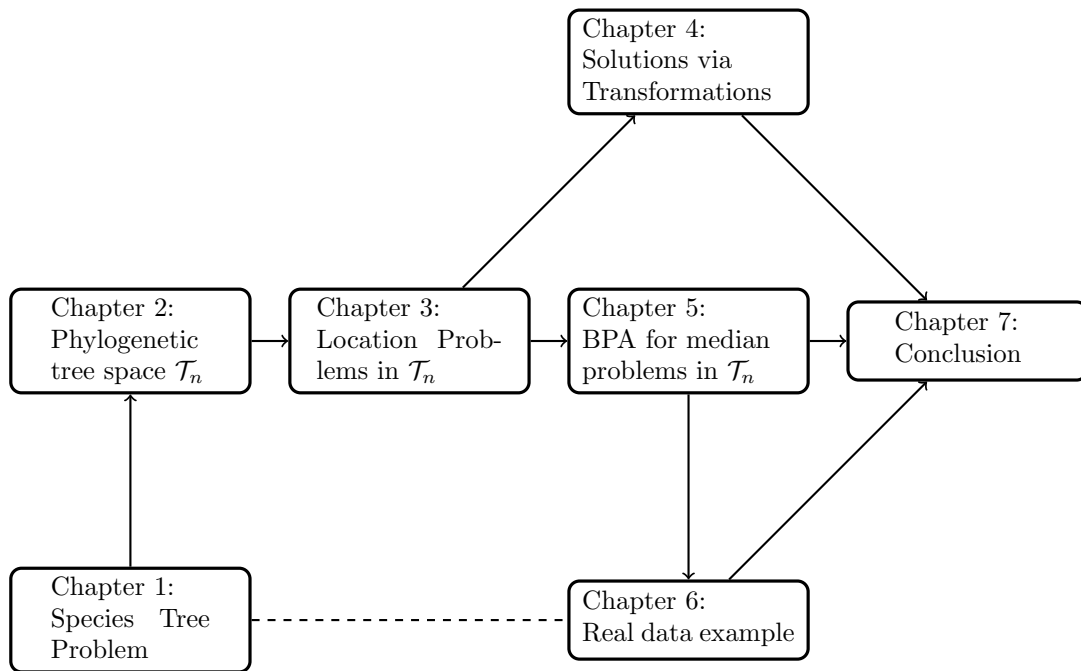


Figure 1.2: Structure of the thesis.

2 Construction and Properties of the Tree Space

As mentioned in the introduction, we will refer to the space of phylogenetic trees \mathcal{T}_n as *tree space*. The construction of this space is given in a meticulous way to ensure that all subtleties of the tree space and important tools required for proofs are introduced carefully. Note, that since this is an introduction of the tree space almost all definitions and results are known from literature and are always indicated with the proper citation.

2.1 Splits, Compatibility and Orthants

We start by giving a concise mathematical definition for the elements of \mathcal{T}_n , the phylogenetic trees. In the following, let n species $\{1, 2, \dots, n\}$ be given whose evolution is to be investigated. These species are leaves of the phylogenetic tree describing their evolution, as they are present today and have no descendants. There is one additional leaf in a phylogenetic tree which is the root node and is indexed by 0. The root node models a known common ancestor of these species.

Due to the problem's nature we are only interested in trees which do not contain any node with degree two. If there exists a node with degree two, it can be removed and its two incident edges may be merged into one, as this node does not convey any information on a bifurcation or multifurcation of species. The edges of a tree can be distinguished into edges which are incident with a leaf and edges which are not. The latter are called *interior edges*. As already mentioned, phylogenetic trees also carry information on how much mutation took place between two nodes, i.e., species in the tree, which is modeled by positive edge weights for all interior edges. Now, with these properties of a phylogenetic tree in mind, we can formally define our desired notion of trees.

Definition 2.1. [BHV01] *A weighted graph $T = (V, E, w)$ is called metric n -tree if*

- $T = (V, E)$ is a tree,
- $\{0, 1, \dots, n\} \subseteq V$ are the only leaves of V and $0 \in V$ is the root node,

- does not have any node of degree two,
- satisfies $w_e > 0$ for all interior edges e of E , i.e., $w \in \mathbb{R}_{>0}^{|E|}$.

In the following the elements of the tree space, the metric n -trees are usually referred to as *trees* in the following. If we want to reference a tree without edge weights, we emphasize this by speaking of the *topology* of the tree or explicitly write $T = (V, E)$, omitting w .

The next result follows immediately from the definition of interior edges and the property of metric n -trees not having nodes of degree two.

Lemma 2.2. [BHV01] *The maximal number of interior edges for metric n -trees is $n - 2$.*

This maximal number of edges is attained when, every non-leaf node has degree three. Then, regarding the tree from the root, 0, being the top, to the leaves, $\{1, \dots, n\}$, at the bottom, each edge from the top splits up to two different edges; this motivates the following name.

Definition 2.3. [BHV01] *A tree with $n - 2$ interior edges is called binary tree.*

Example 2.1.1. *The tree topologies for $n = 4$:*

For $n = 4$ species, all possible topologies of such trees are depicted in figures 2.1 - 2.4. Figure 2.1 contains the star tree as well as four different topologies which all describe that three species have one common ancestor, and one species has developed separately.

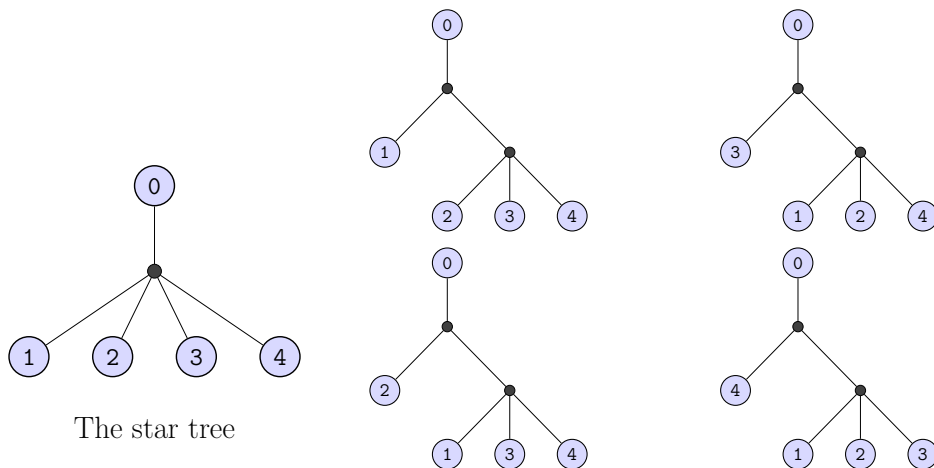


Figure 2.1: Degenerate tree topologies for trees with 4 species.

Figure 2.2 shows the cases of a degenerate tree topology where one pair of species shares a common ancestor that developed after the split of the other two species.

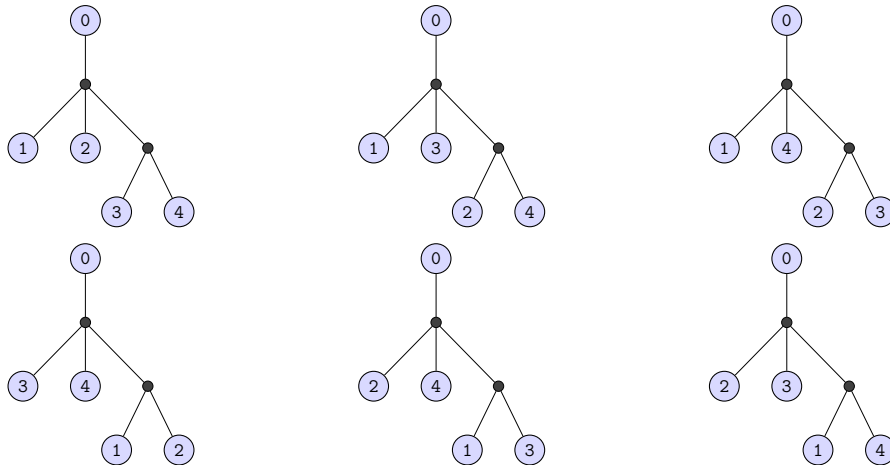


Figure 2.2: Degenerate tree topologies for trees with 4 species.

Figure 2.3 shows the three binary tree topologies in which the species have developed in pairs:

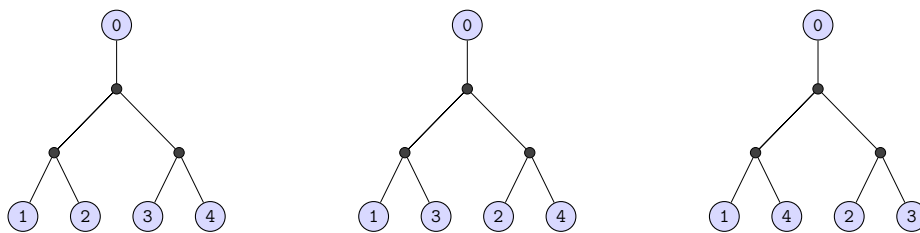


Figure 2.3: 3 of the $15 = (2 \cdot 4 - 3)!! = 5 \cdot 3 \cdot 1$ binary tree topologies for trees with 4 species.

Figure 2.4 shows the remaining twelve binary tree topologies where one species has developed separately, while three of them have one common ancestor from which another species developed separately.

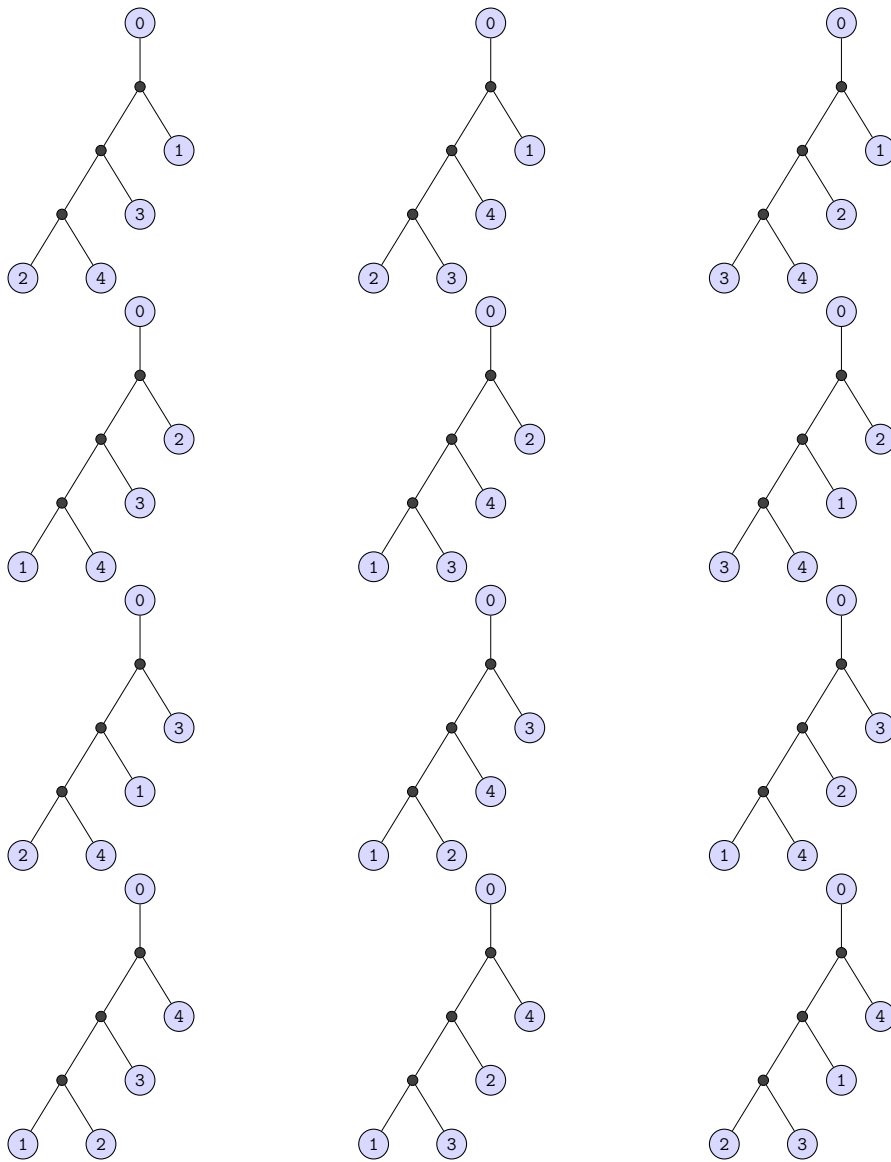


Figure 2.4: 12 of the $15 = (2 \cdot 4 - 3)!! = 5 \cdot 3 \cdot 1$ binary tree topologies for trees with 4 species.

This enumeration of trees is in fact exhausting, i.e., all different tree topologies for $n = 4$ have been depicted up to different planar embeddings of the graphs. Handling phylogenetic trees via tree topologies (V, E) is very cumbersome, which is the reason for the introduction of splits of trees in the next section.

Splits

The figures above have show that there are many different topologies for trees. Our goal now is to describe tree topologies in an easy way, as sets of *splits*. Note that the number of interior nodes (non-leaf nodes) and the number of edges may differ from tree to tree. Thus, a representation using nodes or edges does not seem appropriate. Instead, trees will be described using partitions of its set of leaves that are induced by the *interior edges*, which are edges of the tree that are not incident to any leaf. Partitions are induced when removing edges, which is a well-known observation from graph theory:

A tree $T = (V, E)$ with vertices V and edges E gets disconnected and divided into two trees if we remove any of its edges $e \in E$. More precisely, when removing any edge $e \in E$ the tree T is split into two connected components $(V_1(e), E_1(e))$ and $(V_2(e), E_2(e))$ with $V_1(e) \cup V_2(e) = V$ and $E_1(e) \cup \{e\} \cup E_2(e) = E$.

In particular this yields a partition of the leaves $\{0, 1, \dots, n\} \subseteq V$ of the tree into two disjoint sets $A(e) = V_1(e) \cap \{0, 1, \dots, n\}$ and $A(e)^c = V_2(e) \cap \{0, 1, \dots, n\} = \{0, 1, \dots, n\} \setminus A(e)$. It is easily observed that removing an edge e incident to a leaf $i \in \{0, 1, \dots, n\}$ of the tree results in the partition $\{i\}$ and $\{0, 1, \dots, n\} \setminus \{i\}$, and that these partitions can be obtained for any tree $T \in \mathcal{T}_n$. We can hence neglect these trivial partitions when describing tree topologies. The other partitions are formalized as follows:

Definition 2.4. [Owe11] A split $s = (A|A^c)$ is a partition of $\{0, 1, \dots, n\}$ into two sets A, A^c such that $|A| \geq 2$ and $|A^c| \geq 2$. The set of all splits of the set $\{0, 1, \dots, n\}$ is denoted as \mathcal{S} .

Note that $(A|A^c)$ and $(A^c|A)$ are considered to be the same split. Due to the symmetry of the union $A \cup A^c = A^c \cup A$ they generate the same partition of $\{0, 1, \dots, n\}$. We use the convention that the set which does not contain 0 is always denoted as A , i.e., $0 \in A^c$. Even though a split is already uniquely defined by choosing A , it will help to also write down A^c when working with splits later on.

The most important thing about splits is that they represent interior edges of a tree without the need of interior nodes to define them. That makes it more convenient to represent trees and also to compare how *similar* trees are by just checking which common splits they possess.

Example 2.1.2. Figure 2.5 illustrates how splits arise when removing an interior edge. It is important to note that this justifies to use the words *interior edge* or *split* interchangeably, since there is a one-to-one correspondence between these notions. This fact is stressed again in Lemma 2.7

As shown in the example, one receives splits by removing interior edges e of a tree. Doing this for all interior edges of the tree one receives a set of splits that, as we

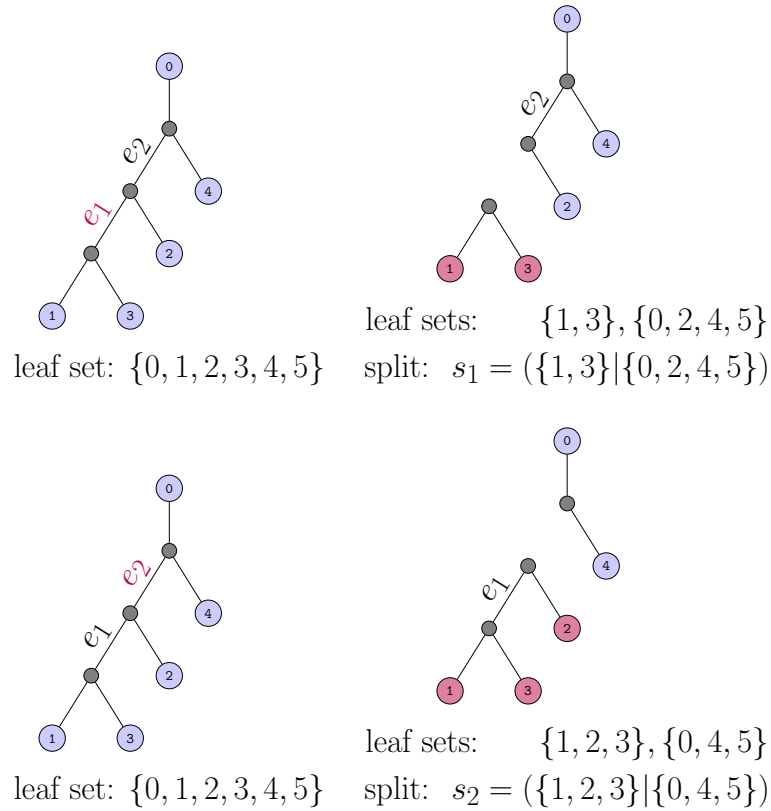


Figure 2.5: An interior edge of a metric n -tree induces a split on $\{0, 1, \dots, n\}$.

will see later, already completely defines the *topology* (V, E) of a tree (V, E, w) . The notion of *topology* is frequently used in the same sense as ‘the combinatorial structure of the tree’ in the literature (see, e.g., [OP11]). Our definition here is a little more specific.

Definition 2.5. For a tree T its induced set of splits, the topology of T , is defined as

$$\text{Split}(T) := \{(A(e) | A(e)^c) : e \text{ is an interior edge of } T\} \subseteq \mathcal{S}.$$

As mentioned earlier we use splits as they offer a convenient and consistent way to describe the topology of a tree by a set of splits. Nonetheless, not all sets of splits $S \subset \mathcal{S}$ yield a tree topology since some pairs of partitions can not be present in the same tree:

Example 2.1.3. Assume $n = 4$ and consider two splits on $\{0, 1, 2, 3, 4\}$,

$$s_1 = (\{1, 2\} | \{0, 3, 4\}) \quad \text{and} \quad s_2 = (\{1, 3\} | \{0, 2, 4\}).$$

There cannot exist a tree whose interior edges yield both of these splits as there cannot be two edges in the same tree stating that $\{1, 2\}$ are direct neighbors and that $\{1, 3\}$ are direct neighbors. More precisely, $\{1, 2\}$ induces a subtree, where 1 is contained in the subtree $\{1, 2\}$ which does not contain 3. $\{1, 3\}$, however induces the subtree $\{1, 3\}$ which does not contain 2, which is a contradiction.

The concept of *compatibility* describes whether two splits may be contained in the same tree and even more, if a set of splits yields a tree topology:

Definition 2.6. [BHV01] Two splits $(A|A^c)$ and $(B|B^c)$ are compatible if $A \subseteq B$ or $A \subseteq B^c$, $B \subseteq A$ or $B^c \subseteq A$. A set of splits $S \subset \mathcal{S}$ is called compatible if every pair $s_i, s_j \in \mathcal{S}$ is compatible.

The definition of splits and compatibilities allows for the following nice representation of tree topologies, which implies the one-to-one correspondence between splits and interior edges that we mentioned in Example 2.1.2.

Lemma 2.7 ([BHV01],[Vog07]). There exists a metric n -tree T with topology $S = \text{Split}(T)$ if and only if all pairs of splits $\{s_1, s_2\} \subset S$ are compatible.

As an example for this, the unweighted tree $T = (V, E)$ or tree topology depicted in Figure 2.5 has two interior edges which result in the two splits $(\{1, 3\}|\{0, 2, 4\})$ and $(\{1, 2, 3\}|\{0, 4\})$. These, in turn, uniquely define the tree topology of T .

Lemma 2.7 in particular implies that two splits are compatible if and only if they can exist in the same tree.

Moreover, it justifies why we called the set of splits the *topology* of a tree in Definition 2.5 and why this fits the terminology chosen in Definition 2.1: The topology of a tree (in the sense of Definition 2.1) $T = (V, E, w)$ is uniquely characterized by its set of splits $\text{Split}(T)$.

An immediate consequence of the one-to-one correspondance and Lemma 2.2 is the following corollary.

Corollary 2.8. The maximal number of compatible splits is $n - 2$.

Embedding into \mathbb{R}_+^N

Now that we have defined splits and compatibilities, we can make use of the representation of trees by split sets to embed the trees into a Euclidean space. To this end, we construct vectors of the length of all splits; easy combinatorics show that there are $N := |\mathcal{S}| = 2^n - n - 2$ splits on $\{0, \dots, n\}$. As the non-negative orthant of \mathbb{R}^N , \mathbb{R}_+^N , is defined inconsistently throughout literature, we specify

$$\mathbb{R}_+^N = \{x \in \mathbb{R}^N : x_i \geq 0 \forall i = 1, \dots, N\}.$$

Definition 2.9. Let the splits in \mathcal{S} be given in a fixed order according to their indices, say, $\mathcal{S} = \{s_1, \dots, s_N\}$. Then we can describe every metric n -tree T by a vector $t \in \mathbb{R}_+^N$ by defining

$$t_i := \begin{cases} w_e > 0 & \text{if } s_i = (A(e)|(A(e)^c) \in \text{Split}(T) \\ 0 & \text{if } s_i \notin \text{Split}(T), \end{cases} \quad (2.1)$$

for $i = 1, \dots, N$.

Let \mathcal{T}_n denote the set of all metric n -trees. Then the mapping

$$\chi : \mathcal{T}_n \rightarrow \mathbb{R}_+^N \quad \chi(T) = t,$$

with t defined as above is called (canonical) embedding of \mathcal{T}_n .

Note that the term ‘embedding’ implicitly requires injectiveness of the mapping. That χ is injective is easy to see: When $\chi(T_1) = t = \chi(T_2)$, then it follows that T_1 and T_2 have the same set of splits, the ones that correspond to the non-zero components of t . Moreover, they have the same weights t_j for these splits, so $T_1 = T_2$. We illustrate the embedding of \mathcal{T}_n into \mathbb{R}_+^N using metric 4-trees.

Example 2.1.4. For $n = 4$ the $N = 2^n - n - 2 = 2^4 - 6 = 10$ possible splits on $\{0, 1, 2, 3, 4\}$ are

$$\begin{aligned} s_1 &= (\{1, 2\}|\{0, 3, 4\}) & s_2 &= (\{1, 3\}|\{0, 2, 4\}) \\ s_3 &= (\{1, 4\}|\{0, 2, 3\}) & s_4 &= (\{2, 3\}|\{0, 1, 4\}) \\ s_5 &= (\{2, 4\}|\{0, 1, 3\}) & s_6 &= (\{3, 4\}|\{0, 1, 2\}) \\ s_7 &= (\{1, 2, 3\}|\{0, 4\}) & s_8 &= (\{1, 2, 4\}|\{0, 3\}) \\ s_9 &= (\{1, 3, 4\}|\{0, 2\}) & s_{10} &= (\{2, 3, 4\}|\{0, 1\}) \end{aligned}$$

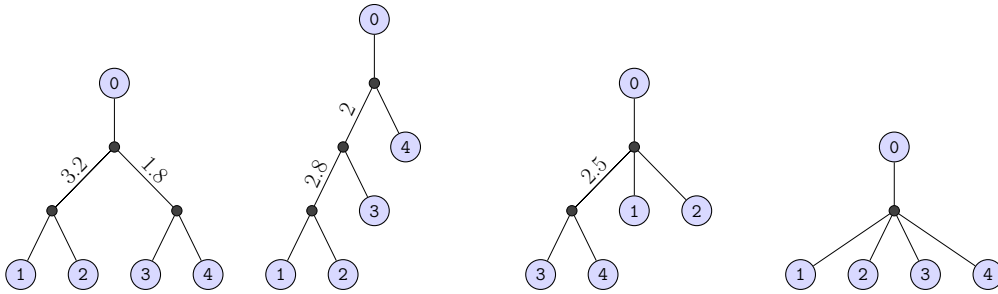


Figure 2.6: Four trees of \mathcal{T}_4 .

The four metric 4–trees depicted in Figure 2.6 are hence given as

$$t^1 = \begin{pmatrix} 3.2 \\ 0 \\ 0 \\ 0 \\ 1.8 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, t^2 = \begin{pmatrix} 2.8 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, t^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2.5 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, t^4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The tree T^4 corresponding to $t^4 = 0 \in \mathbb{R}_+^{10}$ is called the star tree and is depicted on the rightmost in Figure 2.6. Note that not every vector in \mathbb{R}_+^{10} has a pre-image. Consider for example $t_5 = (2\ 3\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^t$. This would translate to the split s_1 having a weight of 2 and s_2 having a weight of 3. But s_1 and s_2 are incompatible, so there exists no tree that can contain both of these splits. This implies that there is no pre-image for t_5 .

With this embedding, each metric n –tree T yields a vector $t \in \mathbb{R}_+^N$, but, as explained in the example, not every $x \in \mathbb{R}_+^N$ represents a metric n –tree, since the embedding is not surjective. We now describe which $x \in \mathbb{R}_+^N$ are representatives of trees metric n –trees, which follows directly from Lemma 2.7 and the definition of the embedding. Recall, that it is crucial to first choose a fixed order of the splits in \mathcal{S} and then to embed all trees into \mathbb{R}_+^N with respect to to this order.

Theorem 2.10. *Let $x \in \mathbb{R}_+^N$. Then x represents a tree if and only if all $s_i, s_j \in \mathcal{S}$ with $x_i \neq 0$ and $x_j \neq 0$ are compatible splits.*

After having introduced the concept of splits and defined the embedding we will now incorporate two possibilities to uniquely identify trees $T = (V, E, w)$, without using the (V, E, w) notation. The first is to specify a tuple $T = ((s_1, \dots, s_k), (w_1, \dots, w_k))$ of a vector of splits (s_1, \dots, s_k) that yields a compatible set of splits $\{s_1, \dots, s_k\} \subset \mathcal{S}$ and the corresponding positive edge weights w_i for $s_i, i = 1, \dots, k$, that are induced by the translation of an interior edge to a split. The second possibility is to identify a tree $T \in \mathcal{T}_n$ via the embedding, i.e., by its corresponding vector $t \in \mathbb{R}_+^N$, such that $\chi^{-1}(t) = T$.

Since we have these two possibilities we will always reference trees $T \in \mathcal{T}_n$ with a capital letter and they consist of a tuple of splits and lengths, whereas the respective lower-case letter t denotes the embedding of T into \mathbb{R}_+^N .

2.2 The Geodesic Distance

In the previous section we have introduced convenient ways of referencing trees and have seen that compatible pairs of splits yield tree topologies. Having all possible sets S of compatible splits we thus get all tree topologies. Considering all possible weight vectors for these compatible split sets we are able to construct all metric n -trees. The tree space \mathcal{T}_n is the collection of all these trees. Next, we define a metric on the set of \mathcal{T}_n to make it a metric space.

This is achieved in two steps: First, we define a distance between two trees that are contained in a common maximal *orthant*, a specific region of the tree space that we define in Definition 2.11. After we know how to measure distance in single orthants we extend the distance to the whole space by investigating how one traverses from one orthant into another.

We start by introducing orthants. The notion of orthants has already been mentioned in [BHV01]; here we define it in a slightly different manner, as regions of the tree space instead of Euclidean orthants.

Definition 2.11. *Let $S \subseteq \mathcal{S}$ be a set of compatible splits in \mathcal{T}_n . Then*

$$\mathcal{O}(S) := \{T \in \mathcal{T}_n : \text{Split}(T) \subseteq S\}$$

is called the orthant of S . Moreover, for $T \in \mathcal{T}_n$ let $\mathcal{O}(T) = \mathcal{O}(\text{Split}(T))$.

In case that S is a set of pairwise compatible splits with maximum cardinality of $n - 2$, we call $\mathcal{O}(S)$ a maximal orthant.

By definition $\mathcal{O}(S_1) \subseteq \mathcal{O}(S_2)$ holds for $S_1 \subseteq S_2$, so any set of splits S with $|S| < n - 2$ is contained in some maximal orthant S' , i.e., $S \subset S'$, $|S'| = n - 2$. Also note, that a tree T whose set of splits $\text{Split}(T)$ is not maximal belongs to several maximal orthants, which is illustrated in Figure 2.7. The most extreme example for this is the star tree $0 \in \mathcal{T}_n$. Since $\text{Split}(0) = \emptyset$, the star tree is contained in *every* orthant of \mathcal{T}_n , since $\text{Split}(0) = \emptyset \subset S$ for all $S \subset \mathcal{S}$.

Before defining the distance in single orthants we refer to a very important result. The number of orthants of \mathcal{T}_n is exponential n :

Theorem 2.12 ([BHV01]). *There exist $(2n - 3)!! = (2n - 3) \cdot (2n - 5) \cdot \dots \cdot 3$ maximal orthants in \mathcal{T}_n .*

With the definition of orthants we now define the distance between trees that are contained in a common orthant. To this end, let $\mathcal{O} = \mathcal{O}(S)$ be an orthant that contains T and X , i.e., $\text{Split}(T) \subset S$ and $\text{Split}(X) \subset S$. [BHV01] define the distance for two such trees to be the Euclidean norm of their weight vectors. Now using the embedding into \mathbb{R}_+^N , i.e., using $t, x \in \mathbb{R}_+^N$ instead of T, X respectively, their distance is given by

$$d(X, T) := \|t - x\|_2.$$

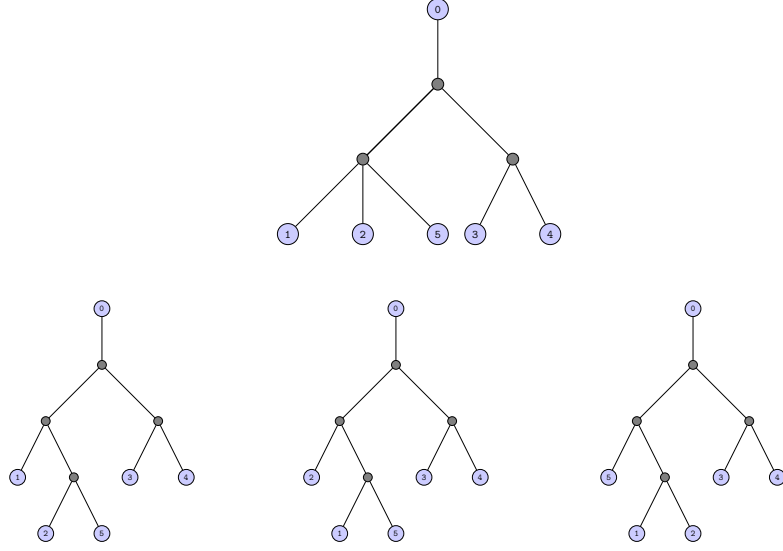


Figure 2.7: A degenerate tree topology with splits $S = \{s_1, s_2\}$ (top) and the three tree topologies with maximal split sets S_1, S_2, S_3 that contain S .

Note, that all components of t and x that do not correspond to splits in $\text{Split}(T) \subset \mathcal{O}$ or $\text{Split}(X) \subset \mathcal{O}$ are 0 by definition of the embedding χ , Definition 2.9.

Before we continue to define the distance in \mathcal{T}_n for two trees who are not contained in a common orthant, we want to emphasize the local Euclidean structure, that is implied by the definition of distance for two trees in an orthant.

Since $d(T, X) = \|t - x\|_2$ for two trees T, X that are in the same orthant, a single maximal orthant with $n - 2$ splits is isometric to the non-negative Euclidean orthant \mathbb{R}_+^{n-2} :

Definition 2.13. *Let a maximal orthant $\mathcal{O} = \mathcal{O}(S) \subset \mathcal{T}_n$ with $S = \{s_{i_1}, \dots, s_{i_{n-2}}\}$ be given. Then the embedding $\psi_{\mathcal{O}}$ of orthant \mathcal{O} into \mathbb{R}_+^{n-2} is given by the map onto the length vectors, i.e., for a tree $T = ((s_{i_1}, \dots, s_{i_{n-2}}), (w_1, \dots, w_{n-2})) \in \mathcal{O}$, define*

$$\psi_{\mathcal{O}}(T) = \begin{pmatrix} w_1 \\ \vdots \\ w_{n-2} \end{pmatrix}$$

and for non-binary trees $T = ((s_{j_1}, \dots, s_{j_l}), (w_1, \dots, w_l)) \in \mathcal{O}(S)$, i.e., when $l < n - 2$ and $\text{Split}(T) = \{s_{j_1}, \dots, s_{j_l}\} \subsetneq S$, define $\psi_{\mathcal{O}}(T)$ coordinate-wise,

$$(\psi_{\mathcal{O}}(T))_m = \begin{cases} w_l & \text{if } s_{j_l} = s_{i_m} \\ 0 & \text{else} \end{cases} \quad \text{for } m = 1, \dots, n - 2.$$

Recall that we have also defined an embedding χ of \mathcal{T}_n into \mathbb{R}_+^N in Definition 2.9. The orthant embedding $\psi(\mathcal{O})$ is simply the projection of $\chi(T)$ to the components of \mathbb{R}_+^N that correspond to splits in \mathcal{O} .

Now, that we have defined the distance between trees that are contained in a common orthant and understood the local Euclidean structure of the orthants, the second step is to *connect* the orthants and measure distance for trees that are not contained in a common orthant via paths through several connected orthants.

Assume now that T and X are not both contained in some orthant \mathcal{O} . This implies, that there exists some split s_1 in $\text{Split}(X)$ that is incompatible with a split s_2 in $\text{Split}(T)$, as we could otherwise take the compatible split set $\bar{S} := \text{Split}(T) \cup \text{Split}(X)$ and $T, X \in \mathcal{O}(\bar{S})$ would show that T and X are contained in a common orthant.

Here, it does not make sense to define the distance by the Euclidean norm of the embedded vectors $\|t - x\|_2$: For each point $y = (1 - \lambda)x + \lambda t$, $\lambda \in (0, 1)$ on the Euclidean shortest path, i.e., the line segment connecting t and x in \mathbb{R}_+^N , we would have that $y_i > 0$ whenever one of x_i and t_i is greater than 0 for some $i \in \{1, \dots, N\}$. Hence the coordinates corresponding to the incompatible splits s_1 and s_2 are both positive for y and Theorem 2.10 implies that $y \in \mathbb{R}_+^N$ is not a representative of a tree, i.e., $y \notin \chi(\mathcal{T}_n)$. This is easier to understand when looking at the example for the embedding of \mathcal{T}_3 into \mathbb{R}_+^3 in Example 2.2.1.

Example 2.2.1. Consider \mathcal{T}_3 . A maximal split set in \mathcal{T}_3 has $n - 2 = 3 - 2 = 1$ splits and there exist $N = 2^3 - 3 - 2 = 3$ splits in \mathcal{T}_3 , compare p. 13. We take two trees with a different topology:

$$T_1 = ((\{2, 3\}|\{0, 1\}), (2)) \quad T_2 = ((\{1, 3\}|\{0, 2\}), (1))$$

Figure 2.8 depicts the embedding of \mathcal{T}_3 into \mathbb{R}^3 , which is exactly the non-negative axes that are drawn as solid lines. The Euclidean shortest path between T_1 and T_2 is the dotted line. Clearly, the line is not contained in $\chi(\mathcal{T}_3)$, so it yields a distance between the trees, but not a corresponding path through tree space. In general, the Euclidean distance of the embedded vectors is the extrinsic distance when used for the subset of trees, as it is the distance that is derived from the ambient space \mathbb{R}_+^N , i.e., the space into which \mathcal{T}_n is embedded. This distance does not have any desirable properties for the space of trees, which is why the distance needs to be defined differently.

What we actually want is a shortest path P from X to T in the tree space, i.e., $P \subseteq \mathcal{T}_n$ that traverses several orthants to connect X and T . To find such a path we need to be able to transform one tree topology into a different tree topology. This is done as follows: By definition, two trees T_1, T_2 with split sets $\text{Split}(T_1), \text{Split}(T_2)$ are in the orthant $\mathcal{O}(S)$, when $\text{Split}(T_1) \cup \text{Split}(T_2) = S$ is a compatible set of splits. That means that we can search for a path from X to T by removing some splits of X to get some tree T' , such that $\text{Split}(T') \subset \text{Split}(X)$ and then add some splits of

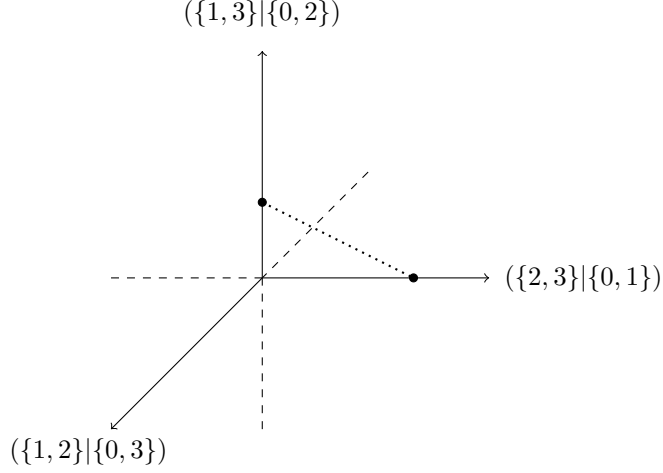


Figure 2.8: Embedding of \mathcal{T}_3 into \mathbb{R}^3 and a Euclidean shortest path in \mathbb{R}^3 connecting two trees of \mathcal{T}_3 .

T to T' and repeat this procedure until we are only left with splits of X , hence the topology of X . Thereby, it is important to remove and add splits in such a way that two successive trees on the way from X to T are contained in a common orthant and we can thus use the Euclidean distance for these pairs of trees. In this way, we have *connected* X and T by a path through tree space.

Now that we have sketched the idea, we formalize the length of paths between two trees, as it has already been described in [BHV01].

Definition 2.14. *Let (T_1, T_2, \dots, T_k) be a sequence of trees in \mathcal{T}_n such that for each $i = 1, \dots, k-1$ there exists an orthant \mathcal{O}_i with $T_i, T_{i+1} \in \mathcal{O}_i$. Then define the path $P = (T_1, \dots, T_k) \subset \mathcal{T}_n$ to be the path that contains all trees T_i and for each T_i, T_{i+1} , $i = 1, \dots, k-1$ it contains all trees $\chi^{-1}(\lambda t_i + (1-\lambda)t_{i+1})$ for $\lambda \in (0, 1)$.*

$P = (T_1, T_2, \dots, T_k)$ is called a path from T_1 to T_k and its length is defined as

$$L(T_1, T_2, \dots, T_k) := \sum_{i=1}^{k-1} \|t_{i+1} - t_i\|.$$

It is crucial to understand that for each $i = 1, \dots, k-1$, T_i and T_{i+1} are contained in a common orthant \mathcal{O}_i and the distance there is defined as the Euclidean distance of their weight vectors. Note that then $y_\lambda = \lambda t_i + (1-\lambda)t_{i+1} \in \mathbb{R}_+^N$ always represents a tree, i.e. $y_\lambda \in \chi(\mathcal{T}_n)$. This holds, as only splits that are in the common orthant \mathcal{O}_i may have non-zero components in the embedding. More precisely, $y_\lambda \in \chi(\mathcal{O}_i) \subset \chi(\mathcal{T}_n)$. As the measure in an orthant is the Euclidean distance, it follows that the

shortest path between their weight vectors is just the straight line which is mapped back to \mathcal{T}_n by χ^{-1} .

We illustrate the concept of orthants and paths in \mathcal{T}_4 . First, we state all maximal orthants and then we show an example of a path and a geodesic.

Example 2.2.2. *In \mathcal{T}_4 it can be easily checked that no more than two splits can be pairwise compatible. The maximal orthants in \mathcal{T}_4 hence all contain two compatible splits. There exist 15 maximal orthants, namely*

$$\begin{aligned}
\mathcal{O}_1 &= \{(\{1, 2\}|\{0, 3, 4\}), (\{3, 4\}|\{0, 1, 2\})\} = \{s_1, s_6\} \\
\mathcal{O}_2 &= \{(\{1, 3\}|\{0, 2, 4\}), (\{2, 4\}|\{0, 1, 3\})\} = \{s_2, s_5\} \\
\mathcal{O}_3 &= \{(\{1, 4\}|\{0, 2, 3\}), (\{2, 3\}|\{0, 1, 4\})\} = \{s_3, s_4\} \\
\mathcal{O}_4 &= \{(\{1, 2\}|\{0, 3, 4\}), (\{1, 2, 3\}|\{0, 4\})\} = \{s_1, s_7\} \\
\mathcal{O}_5 &= \{(\{1, 2\}|\{0, 3, 4\}), (\{1, 2, 4\}|\{0, 3\})\} = \{s_1, s_8\} \\
\mathcal{O}_6 &= \{(\{1, 3\}|\{0, 2, 4\}), (\{1, 2, 3\}|\{0, 4\})\} = \{s_2, s_7\} \\
\mathcal{O}_7 &= \{(\{1, 3\}|\{0, 2, 4\}), (\{1, 3, 4\}|\{0, 2\})\} = \{s_2, s_9\} \\
\mathcal{O}_8 &= \{(\{1, 4\}|\{0, 2, 3\}), (\{1, 2, 4\}|\{0, 3\})\} = \{s_3, s_8\} \\
\mathcal{O}_9 &= \{(\{1, 4\}|\{0, 2, 3\}), (\{1, 3, 4\}|\{0, 2\})\} = \{s_3, s_9\} \\
\mathcal{O}_{10} &= \{(\{2, 3\}|\{0, 1, 4\}), (\{1, 2, 3\}|\{0, 4\})\} = \{s_4, s_7\} \\
\mathcal{O}_{11} &= \{(\{2, 3\}|\{0, 1, 4\}), (\{2, 3, 4\}|\{0, 1\})\} = \{s_4, s_{10}\} \\
\mathcal{O}_{12} &= \{(\{2, 4\}|\{0, 1, 3\}), (\{1, 2, 4\}|\{0, 3\})\} = \{s_5, s_8\} \\
\mathcal{O}_{13} &= \{(\{2, 4\}|\{0, 1, 3\}), (\{2, 3, 4\}|\{0, 1\})\} = \{s_5, s_{10}\} \\
\mathcal{O}_{14} &= \{(\{3, 4\}|\{0, 1, 2\}), (\{1, 3, 4\}|\{0, 2\})\} = \{s_6, s_9\} \\
\mathcal{O}_{15} &= \{(\{3, 4\}|\{0, 1, 2\}), (\{2, 3, 4\}|\{0, 1\})\} = \{s_6, s_{10}\}
\end{aligned}$$

The structure of \mathcal{T}_4 can be represented by a graph $G = (\mathcal{S}, \mathcal{O})$ whose node set \mathcal{S} contains the ten splits $\{s_1, \dots, s_{10}\}$ of \mathcal{T}_4 and whose edges connect two splits (s_i, s_j) if and only if s_i and s_j are compatible. Consequently, the edges belong to the maximal sets of compatible splits in \mathcal{T}_4 and which are the orthants of \mathcal{T}_4 , see Figure 2.9. Note, that this compatibility graph actually is the Petersen graph.

Now consider the trees $X = (\{s_2, s_9\}, (2, 2))$ and $T = (\{s_1, s_6\}, (1, 4))$. Then Figure 2.10 depicts two different paths from X to T :

$$P_1 = (X, T_2, T_3, T) \quad P_2 = (X, T_1, T)$$

Then, calculating the Euclidean distance between all pairs of trees on the path, we receive

$$\begin{aligned}
L(P_1) &= \|t - t_2\|_2 + \|t_2 - t_3\|_2 + \|t_3 - x\|_2 \\
&= \sqrt{2^2 + (2-1)^2} + \sqrt{2^2 + 1^2} + \sqrt{(4-2)^2 + 1^2} \approx 6.71, \\
L(P_2) &= \|t - t_1\|_2 + \|t_1 - x\|_2 = \|t\|_2 + \|x\|_2 = \sqrt{8} + \sqrt{17} \approx 6.83.
\end{aligned}$$

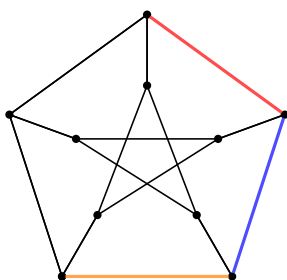


Figure 2.9: Compatibility Graph for \mathcal{T}_4 . The three colored orthants are used for the example for the paths below.

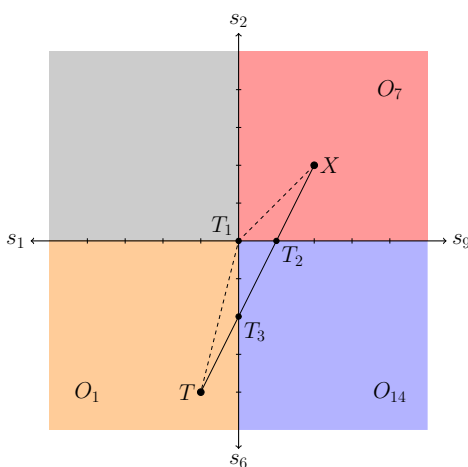


Figure 2.10: Example for two paths from X to T in a diagram of a part of \mathcal{T}_4 . The shortest path connecting X and T is the solid path.

In Definition 2.15, the distance between two trees X and T is now simply defined as the length of the *geodesic*, i.e., a path in tree space from X to T with minimal length. Before that, we provide a little bit of background knowledge about the term *geodesic* in mathematics and how it relates to tree space.

The notion of geodesics originates from differential geometry and is used to generalize the notion of straight lines to curved spaces and is, e.g., used for Riemannian manifolds. Geodesics have later been generalized to metric spaces, thereby introducing the notion of a geodesic metric space. As a matter of fact Theorem 2.16 will show that \mathcal{T}_n is a geodesic metric space, which is why this shortest path is called geodesic. Moreover, as geodesics were introduced as a generalization of straight lines it is in-

interesting to note that the geodesics here are basically concatenations of straight lines in single orthants and that the tree space actually has global non-positive curvature (see [BHV01]).

Definition 2.15. [BHV01] *Let $T, X \in \mathcal{T}_n$. The geodesic distance between X and T is given by*

$$d(X, T) = \inf \{L(T_1, \dots, T_k) : (T_1, \dots, T_k) \text{ is a path from } X \text{ to } T \text{ in } \mathcal{T}_n\}.$$

A path $P = (T_1, \dots, T_k)$ from X to T , $X = T_1, T = T_k$ that attains this minimal distance, i.e., $L(P) = d(T, X)$ is called geodesic.

Plugging in the length of a path from Definition 2.14 we get

$$d(X, T) = \inf \left\{ \sum_{i=1}^k \|t_{i+1} - t_i\|_2 : (T_1, \dots, T_k) \text{ is a path from } X \text{ to } T \text{ in } \mathcal{T}_n \right\}.$$

Naturally, every path from X to T yields an upper bound on the geodesic distance. A special path that connects any two trees in \mathcal{T}_n is the path through the star tree $0 \in \mathcal{T}_n$ and it is called *cone path*. Formally, the cone path for $X, T \in \mathcal{T}_n$ is $(X, 0, T)$ and its length is $\|t\|_2 + \|x\|_2$.

Due to the triangle inequality of $\|\cdot\|_2$ that we may apply in every orthant \mathcal{O} , we can replace a sequence

$$(T_1, \dots, T_i, T_{i+1}, \dots, T_j, \dots, T_l)$$

of trees where T_i, T_{i+1}, \dots, T_j are contained in the same orthant \mathcal{O} by

$$(T_1, \dots, T_i, T_j, \dots, T_l)$$

and will thereby not increase the length of the path.

By the definition of geodesics it is not obvious that the infimum is always attained. The existence of a unique geodesic that attains this infimum connecting any pair of trees in \mathcal{T}_n has been shown in [BHV01] by applying a result of [Gro87] concerning metric spaces of global non-positive curvature, also called *CAT(0) spaces*. As we do not need curvature in the following, but only need it to get existence and uniqueness of geodesics we refer the interested reader to the appendix, Chapter 7.

Theorem 2.16 ([BHV01]). *\mathcal{T}_n is a CAT(0) space. In particular, the geodesic distance d is a metric on \mathcal{T}_n .*

For CAT(0) spaces, existence and uniqueness of geodesics are guaranteed:

Theorem 2.17 ([BHV01]). *Let $T, X \in \mathcal{T}_n$. There exists a unique geodesic $\Gamma(X, T)$ from X to T .*

As a final note concerning curvature we remark that a Hadamard space is defined as a complete CAT(0) space. Since \mathcal{T}_n is complete, this makes it a Hadamard space, as mentioned in Chapter 7, or compare [Bac14b] for more details on the subject. The fact that \mathcal{T}_n is a Hadamard space will be of use later.

2.2.1 Support and Parametrization of a Geodesic

Now that the existence and uniqueness has been established, we want to actually find the geodesic between X and T , which we denote with $\Gamma(X, T)$.

The key to finding and parametrizing the geodesic from a tree X to a tree T is the concept of the *support sequence* or the *support* of a geodesic, that describes in which order edges of the tree X are removed and edges of T are added. With this concept it is possible to apply combinatorial methods to find the geodesic.

[BHV01] prove that splits that are contained in both X and T are contained in every tree along the geodesic. Hence, we do not need to consider these splits when searching for the topologies of trees along the geodesic. The same result holds for splits that are contained in one tree and are compatible with all splits of the other tree. These two types of splits are merged in the following definition.

Definition 2.18. *Given two trees $T, X \in \mathcal{T}_n$. A split s is called double compatible if s is compatible with all splits in $\text{Split}(X)$ as well as with all splits in $\text{Split}(T)$. Let $C = C(T, X)$ denote the set of double compatible splits for two trees.*

Next we define a partition of the split sets that describes in which order we remove splits from X and add splits from T on the path from X to T .

Definition 2.19. [OP11] *With the notation introduced above, a support sequence from $X \in \mathcal{T}_n$ to $T \in \mathcal{T}_n$ is a pair $(\mathcal{A}, \mathcal{B}) = ((A_1, \dots, A_k), (B_1, \dots, B_k))$ with $A_i \subset \text{Split}(X)$, $B_i \subset \text{Split}(T)$ such that*

$$\begin{aligned} A_1 \cup \dots \cup A_k &= \text{Split}(X) \setminus C \\ B_1 \cup \dots \cup B_k &= \text{Split}(T) \setminus C \end{aligned}$$

is a partition, i.e., a disjoint union.

In order to define the *support* of a geodesic we introduce the following notation. Let $T = ((s_1, \dots, s_l), (w_1, \dots, w_l)) \in \mathcal{T}_n$. Then for a subset $A \subseteq \text{Split}(T)$ define

$$\|A\|_2 = \sqrt{\sum_{s_i \in A} w_{s_i}}.$$

Moreover, we also write w_s^T, w_s^X to indicate the weight of splits s in tree T or tree X respectively for a split $s \in \text{Split}(T) \cap \text{Split}(X)$.

Definition 2.20. [OP11] Let $T, X \in \mathcal{T}_n$ and let $\mathcal{A} = (A_1, \dots, A_k), \mathcal{B} = (B_1, \dots, B_k)$ be a support sequence from X to T . Then $(\mathcal{A}, \mathcal{B})$ is called a support of the geodesic from X to T if it satisfies the following three properties:

(P1) For each $i > j$, A_i and B_j are compatible sets of splits.

(P2)

$$\frac{\|A_1\|}{\|B_1\|} \leq \dots \leq \frac{\|A_k\|}{\|B_k\|}$$

(P3) For $i = 1, \dots, k$ there is no non-trivial partition (C_1, C_2) of A_i and (D_1, D_2) of B_i such that C_2 is compatible with D_1 and $\frac{\|C_1\|}{\|D_1\|} < \frac{\|C_2\|}{\|D_2\|}$.

In the following we describe what the three properties (P1), (P2) and (P3) actually mean.

From the definition of the geodesic distance it can be seen that the geodesic consists of “straight lines”, also called *legs*, one for each orthant that it traverses. This follows from its property of being the shortest path connecting two trees together with the definition of distance in the orthants, being the Euclidean distance.

Now, knowing that the geodesic is, roughly speaking, a piecewise Euclidean path that may only bend when traversing a boundary of an orthant, we can use this knowledge to interpret (P1), (P2) and (P3), where we stick to the notation of Definition 2.20.

A support sequence as in Definition 2.19 generally defines a sequence of orthants that a path traverses through. In case of the geodesic from X to T , the first orthant on the geodesic is the orthant of X , $\mathcal{O}(X)$, so the trees on the straight line, the leg, in this orthant have the splits $\text{Split}(X)$. To get to the orthant of T we have to remove the splits of $X \setminus C$ and add splits of T ; that is the general idea of the partitions of $\text{Split}(X) \setminus C$ into A_1, \dots, A_k and $\text{Split}(T) \setminus C$ into B_1, \dots, B_k . We want to successively delete splits of X , starting with A_1 , then A_2 and so on and add splits of T starting at B_1 , then B_2 and so forth.

(P1) ensures that this procedure of removing and adding splits actually yields an orthant of \mathcal{T}_n , i.e., that $B_1 \cup \dots \cup B_i \cup A_{i+1} \cup \dots \cup A_k$ is a compatible set of splits. Now, while (P1) cares about “feasibility”, (P2) and (P3) together ensure that the split weights along the path through these orthants are chosen optimally.

Theorem 2.21. [OP11] Let $T, X \in \mathcal{T}_n$, $(\mathcal{A}, \mathcal{B})$ be the support for the geodesic from X to T and C be the set of double compatible splits. Define the path $P(\mathcal{A}, \mathcal{B}) = \{\gamma(\lambda) : \lambda \in [0, 1]\}$ via a parametrization $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ that consists of several

legs P_l , which are different sections of the path that are straight lines in different orthants:

$$P_l = \begin{cases} \left[\gamma(\lambda) : \frac{\lambda}{1-\lambda} < \frac{\|A_{l+1}\|_2}{\|B_{l+1}\|_2} \right] & l = 0 \\ \left[\gamma(\lambda) : \frac{\|A_l\|_2}{\|B_l\|_2} \leq \frac{\lambda}{1-\lambda} < \frac{\|A_{l+1}\|_2}{\|B_{l+1}\|_2} \right] & l = 1, \dots, k-1 \\ \left[\gamma(\lambda) : \frac{\|A_k\|_2}{\|B_k\|_2} \leq \frac{\lambda}{1-\lambda} \right] & l = k \end{cases}$$

The trees $\gamma(\lambda)$ on each leg P_l , i.e., for $\lambda \in \left[\frac{\|A_l\|_2}{\|A_l\|_2 + \|B_l\|_2}, \frac{\|A_{l+1}\|_2}{\|A_{l+1}\|_2 + \|B_{l+1}\|_2} \right)$ have the set of splits $\text{Split}(\gamma(\lambda)) = B_1 \cup \dots \cup B_l \cup A_{l+1} \cup \dots \cup A_k \cup C$ with lengths

$$w_s(\lambda) = \begin{cases} \frac{(1-\lambda)\|A_j\|_2 - \lambda\|B_j\|_2}{\|A_j\|_2} \cdot w_s^X, & s \in A_j, \\ \frac{\lambda\|B_j\|_2 - (1-\lambda)\|A_j\|_2}{\|B_j\|_2} \cdot w_s^T, & s \in B_j, \\ (1-\lambda)w_s^X + \lambda w_s^T, & s \in C. \end{cases}$$

The path $P(\mathcal{A}, \mathcal{B})$ is then the geodesic from X to T , i.e., $P(\mathcal{A}, \mathcal{B}) = \Gamma(T, X)$ with parametrization $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ and the length of the path is

$$d(T, X) = \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2}.$$

Theorem 2.17 has already established that there always exists a unique geodesic between two trees $T, X \in \mathcal{T}_n$, but Theorem 2.21 gives us a parametrization of this path. Despite the uniqueness of the geodesic and the dependence of the parametrization on the support, the support of a geodesic need not be unique. At first glance this seems to be a contradiction, but the following example illustrates, that, if a geodesic has more than one support they all yield the same parametrization.

Example 2.2.3. Consider the following four splits for $n = 4$:

$$\begin{aligned} s_1 &= (\{1, 2\} | \{0, 3, 4\}), & s_2 &= (\{1, 2, 3\} | \{0, 4\}), \\ s_3 &= (\{1, 3\} | \{0, 2, 4\}), & s_4 &= (\{1, 3, 4\} | \{0, 2\}) \end{aligned}$$

Now, let $T = ((s_1, s_2), (2, 2))$ and $X = ((s_3, s_4), (1.5, 1.5))$. The situation is depicted in Figure 2.11, where the geodesic from X to T , $\Gamma(X, T)$, is indicated as the solid black line. In $\mathcal{O}(X)$, all geodesics $\Gamma(Y, T)$ with Y on the upper side of the line emanating from T through the origin (the blue region) have support

$$(\mathcal{A}, \mathcal{B}) = ((\{s_3, s_4\}), (\{s_1, s_2\})),$$

whereas for all trees Y below this line (the red region) the support for $\Gamma(Y, T)$ is

$$(\mathcal{A}', \mathcal{B}') = ((A'_1, A'_2), (B'_1, B'_2)) = ((\{s_4\}, \{s_3\}), (\{s_2\}, \{s_1\})).$$

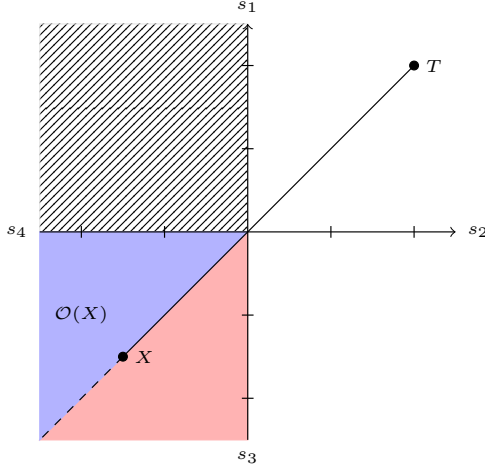


Figure 2.11: The geodesic from X to T does not have a unique support.

For X or any other tree on the ray from 0 through X , both supports satisfy **(P1)**, **(P2)** and **(P3)** and yield the same parametrization. We first show that **(P1)**, **(P2)** and **(P3)** hold for both supports:

We start with $(\mathcal{A}, \mathcal{B})$: For the ‘trivial’ support, **(P1)** and **(P2)** are satisfied automatically as there is only one pair A_1, B_1 . For **(P3)** we need to check if there exists a partition C_1, C_2, D_1, D_2 for $A_1 = \{s_3, s_4\}$ and $B_1 = \{s_1, s_2\}$ such that C_2 and D_1 are compatible and the ratio requirement holds. The only possibility to partition them such that C_2 and D_1 are compatible is to set $C_1 = \{s_4\}, C_2 = \{s_3\}$ and $D_1 = \{s_2\}, D_2 = \{s_1\}$. Then the ratio is

$$\frac{\|C_1\|_2}{\|C_2\|_2} = \frac{1.5}{1.5} = \frac{2}{2} = \frac{\|D_1\|_2}{\|D_2\|_2},$$

so $(\mathcal{A}, \mathcal{B})$ satisfies **(P3)** as well and is a support for $\Gamma(X, T)$.

Now, for $(\mathcal{A}', \mathcal{B}')$ **(P1)** holds because $B'_1 = \{s_2\}$ is compatible with $A'_2 = \{s_3\}$. We calculate the ratio sequence

$$\frac{\|A'_1\|_2}{\|B'_1\|_2} = \frac{1.5}{2} = \frac{1.5}{2} = \frac{\|A'_2\|_2}{\|B'_2\|_2}$$

and conclude that **(P2)** holds as well. **(P3)** holds trivially because there is no non-trivial partition of the sets A'_i, B'_i because they only have one element each.

Note, that **(P3)** holds with equality for the trivial support $(\mathcal{A}, \mathcal{B})$ and that for the other support $(\mathcal{A}', \mathcal{B}')$, **(P2)** only holds with equality; this is the origin of the non-uniqueness.

As we have seen, supports need not be unique. However, a support of a geodesic yields a parametrization for the geodesic and geodesics in the tree space are unique. This implies that the parametrizations that we get from these supports have to coincide and in the following we explain why this has to hold.

For support $(\mathcal{A}', \mathcal{B}')$ the geodesic yields three legs through the orthants

$$\mathcal{O}(X), \mathcal{O}(\{s_2, s_3\}), \mathcal{O}(T).$$

However, there exists no $\lambda \in [0, 1]$ such that a tree is contained in leg P_1 , since λ would have to satisfy

$$\frac{\|A'_1\|_2}{\|B'_1\|_2} = 1 \leq \frac{\lambda}{1-\lambda} < 1 = \frac{\|A'_2\|_2}{\|B'_2\|_2}.$$

So the only two legs containing trees are the ones in $\mathcal{O}(X)$ and $\mathcal{O}(T)$, just as for the trivial support. However, it remains to show that the lengths of the splits coincide. We start with the first leg in $\mathcal{O}(X)$. The lengths of the splits in $A'_1 \cup A'_2$ are given by

$$\frac{(1-\lambda)\|A'_j\|_2 - \lambda\|B'_j\|_2}{\|A'_j\|_2} \cdot w_s^X = (1-\lambda)w_s^X - \lambda \frac{\|B'_j\|_2}{\|A'_j\|_2} w_s^X$$

for $j = 1, 2$. We have seen that the sets A_1, B_1 can be partitioned into A'_1, A'_2, B'_1, B'_2 such that $\frac{\|A'_1\|_2}{\|B'_1\|_2} = \frac{\|A'_2\|_2}{\|B'_2\|_2}$ holds. Then, using $\|A_1\|_2^2 = \|A'_1\|_2^2 + \|A'_2\|_2^2$ and $\|B_1\|_2^2 = \|B'_1\|_2^2 + \|B'_2\|_2^2$ we get

$$\begin{aligned} \frac{\|A'_1\|_2^2}{\|B'_1\|_2^2} &= \frac{\|A'_2\|_2^2}{\|B'_2\|_2^2} \Leftrightarrow \|A'_2\|_2^2 \cdot \|B'_1\|_2^2 = \|A'_1\|_2^2 \cdot \|B'_2\|_2^2 \\ \Leftrightarrow (\|A'_1\|_2^2 + \|A'_2\|_2^2) \cdot \|B'_1\|_2^2 &= (\|B'_1\|_2^2 + \|B'_2\|_2^2) \cdot \|A'_1\|_2^2 \\ \Leftrightarrow \frac{\|B'_1\|_2^2}{(\|B'_1\|_2^2 + \|B'_2\|_2^2)} &= \frac{\|A'_1\|_2^2}{(\|A'_1\|_2^2 + \|A'_2\|_2^2)} \Leftrightarrow \frac{\|A_1\|_2^2}{\|B_1\|_2^2} = \frac{\|A'_1\|_2^2}{\|B'_1\|_2^2}. \end{aligned}$$

With this it follows that for all λ it holds that

$$(1-\lambda)w_s^X - \lambda \frac{\|B'_j\|_2}{\|A'_j\|_2} w_s^X = (1-\lambda)w_s^X - \lambda \frac{\|B_1\|_2}{\|A_1\|_2} w_s^X$$

for $j = 1, 2$ and thus also

$$\frac{(1-\lambda)\|A'_j\|_2 - \lambda\|B'_j\|_2}{\|A'_j\|_2} \cdot w_s^X = \frac{(1-\lambda)\|A_1\|_2 - \lambda\|B_1\|_2}{\|A_1\|_2} \cdot w_s^X$$

for $j = 1, 2$. The same calculation can be done for all $s \in B'_1 \cup B'_2 = B_1$ and one gets

$$\frac{\lambda\|B'_j\|_2 - (1-\lambda)\|A'_j\|_2}{\|B'_j\|_2} \cdot w_s^T = \frac{\lambda\|B_1\|_2 - (1-\lambda)\|A_1\|_2}{\|B_1\|_2} \cdot w_s^T$$

for $j = 1, 2$. Thus for a fixed $\lambda \in [0, 1]$ the lengths of all splits are the same for the two parametrizations.

This sketches the reason that the choice of the support (in case of non-uniqueness) does not influence the parametrization: as the ratios in **(P2)** and **(P3)** hold with equality we have that the scalars for w_s^X, w_s^T are the same and the additional legs given by one of the supports are actually empty.

At this point we mention an important aspect concerning the definition of the geodesic and its parametrization. We have introduced the geodesic in a similar fashion as in [OP11], who introduced the notions of supports and developed the polynomial time algorithm to calculate the geodesic and the parametrization above. Here, the *geodesic is defined as the unique shortest path between two trees*, i.e., the path that attains the infimum in Definition 2.15. In literature about geodesic metric spaces, however, a geodesic on a CAT(0) space (\mathcal{X}, \tilde{d}) is usually defined as a parametrized curve $\tilde{\gamma} : [0, 1] \rightarrow \mathcal{X}$ that has the *constant speed property*, i.e., satisfies

$$\tilde{d}(\tilde{\gamma}(\lambda), \tilde{\gamma}(\lambda')) = |\lambda' - \lambda| \cdot \tilde{d}(\tilde{\gamma}(0), \tilde{\gamma}(1)) \quad \forall \lambda, \lambda' \in [0, 1]. \quad (2.2)$$

This in particular implies that the length of the geodesic equals the distance of its start and end point $\tilde{\gamma}(0), \tilde{\gamma}(1)$, but in comparison to the definition in [OP11], (2.2) holds for the parametrization.

As the tree space is a CAT(0) space, it is clear that a parametrization satisfying (2.2) exists. Moreover, since geodesics in CAT(0) spaces are unique it is also clear that the parametrization of the geodesic in Theorem 2.21 in fact yields the correct path. Nevertheless, it does not hold that any parametrization for the path of the geodesic also satisfies (2.2)! In our opinion it is not immediately clear from the definition that the parametrization from Theorem 2.21 satisfies the constant speed property (2.2). Even though some papers implicitly use that the parametrization satisfies this property, and we assume that the authors were aware that this property holds, we have not found a proof for it and hence present it in the following.

Before we are able to prove the statement we prove several basic facts about the geodesic and its parametrization from [Owe11], that are known, but for which we could not find a proof as well.

Theorem 2.22. For $i = 1, \dots, k$ define $\lambda_i = \frac{\|A_i\|_2}{\|A_i\|_2 + \|B_i\|_2}$ and $T_i = \gamma(\lambda_i)$. As usual, let $t_i = \chi(T_i)$ be the embedded vectors for $i = 1, \dots, k$.

$$(F1) \text{ Split}(T_i) = B_1 \cup \dots \cup B_{i-1} \cup A_{i+1} \cup \dots \cup A_k \cup C.$$

$$(F2) \ d(X, T) = \|x - t_1\|_2 + \sum_{i=1}^{k-1} \|t_{i+1} - t_i\|_2 + \|t - t_k\|_2.$$

$$(F3) \ d(X, T) = d(X, Z) + d(Z, T) \Leftrightarrow Z \in \Gamma(X, T)$$

(F4) If $Y_1 \neq Y_2 \in \Gamma(X, T)$ then $\Gamma(Y_1, Y_2) \subset \Gamma(X, T)$.

Proof. (F1): Note that $\lambda_i = \frac{\|A_i\|_2}{\|A_i\|_2 + \|B_i\|_2} \Leftrightarrow \frac{\lambda_i}{1-\lambda_i} = \frac{\|A_i\|_2}{\|B_i\|_2}$. Then it holds, by definition of the parametrization, that $T_i = \gamma(\lambda_i) \in P_i$ and $\text{Split}(T_i) = B_1 \cup \dots \cup B_i \cup A_{i+1} \cup \dots \cup A_k \cup C$. Now we take $s \in B_i$ and calculate the split length for the parameter λ_i . As the denominator

$$\begin{aligned} \lambda_i \|B_i\|_2 - (1 - \lambda_i) \|A_i\|_2 &= \frac{\|A_i\|_2}{\|A_i\|_2 + \|B_i\|_2} \cdot \|B_i\|_2 + \frac{\|A_i\|_2}{\|A_i\|_2 + \|B_i\|_2} \cdot \|A_i\|_2 - \|A_i\|_2 \\ &= \frac{\|A_i\|_2 (\|A_i\|_2 + \|B_i\|_2)}{\|A_i\|_2 + \|B_i\|_2} - \|A_i\|_2 = 0, \end{aligned}$$

the split length of all splits in B_i are zero. This means, when being precise, $s \notin B_i$. Thus, the split set of $\text{Split}(T_i)$ is $B_1 \cup \dots \cup B_{i-1} \cup A_{i+1} \cup A_k \cup C$, as the other denominators are non-zero.

(F2): Follows from (F1). The T_i are by construction of the parametrization, exactly the points of the geodesic from X to T that are contained in the intersection of two legs, i.e., are contained in the intersection of two orthants. The length of the legs is given by the Euclidean norms of the endpoints of the legs, i.e., $\|t_{i+1} - t_i\|_2$, by definition of the geodesic distance for two trees in a common orthant. Then $d(X, T) = \sum_{i=1}^k \|t_{i+1} - t_i\|_2$ follows, as the length of the geodesic equals the sum of the lengths of its legs, compare Definition 2.15.

(F3): “ \Leftarrow ”: If Z is one of the T_i , then it follows directly from (F2). Otherwise, Z is on the leg between some T_i and T_{i+1} , and as the geodesic is locally the Euclidean line segment, $z = (1 - \lambda)t_i + \lambda t_{i+1}$. Then $\|t_{i+1} - t_i\|_2 = \|t_{i+1} - z\|_2 + \|z - t_i\|_2$. “ \Rightarrow ”: If $Z \notin \Gamma(X, T)$, then the concatenation of the geodesics $\Gamma(X, Z)$ and $\Gamma(Z, T)$ would yield an alternative path from X to T than $\Gamma(X, T)$ with the same length, contradicting the uniqueness of geodesics.

(F4): We first show that $d(X, T) = d(X, Y_1) + d(Y_1, Y_2) + d(Y_2, T)$ holds. Note, that when $Y_1 \neq Y_2 \in \Gamma(X, T)$, then $d(X, Y_1) \neq d(X, Y_2)$, as we would otherwise get two geodesics from X to Y_1 and X to Y_2 of the same length, and this would either contradict the uniqueness of the geodesic from X to T or imply that one of the $Y_i \notin \Gamma(X, T)$. Hence we may assume w.l.o.g. that $d(X, Y_1) < d(X, Y_2)$. In the proof of (F3) we have shown that, when $Z \in \Gamma(X, T)$, then $\Gamma(X, T)$ is the concatenation of $\Gamma(X, Z)$ and $\Gamma(Z, T)$, which we informally write as $\Gamma(X, T) = \Gamma(X, Z) + \Gamma(Z, T)$ for the sake of brevity. Since $Y_1 \in \Gamma(X, T)$ we get $\Gamma(X, T) = \Gamma(X, Y_1) + \Gamma(Y_1, T)$. Now, because $d(X, Y_2) > d(X, Y_1)$, $Y_2 \notin \Gamma(X, Y_1)$, thus $Y_2 \in \Gamma(Y_1, T)$. Applying the statement for Y_2 we get $\Gamma(Y_1, T) = \Gamma(Y_1, Y_2) + \Gamma(Y_2, T)$. This concludes the proof, as $\Gamma(Y_1, Y_2) \subset \Gamma(Y_1, T) \subset \Gamma(X, T)$. \square

With these properties at hand we are able to prove that the constant speed property holds for the parametrization from Theorem 2.21.

Theorem 2.23. *Let $T, X \in \mathcal{T}_n$ and $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ the parametrization of the geodesic from X to T as in Theorem 2.21. Then γ satisfies (2.2).*

Proof. We start by proving the theorem for specific values of $\lambda, \lambda' \in (0, 1)$ and gradually generalize the values of λ, λ' .

First of all, let $(\mathcal{A}, \mathcal{B}) = ((A_1, \dots, A_k), (B_1, \dots, B_k))$ be a support for the geodesic from X to T with the set C of double compatible splits. Define the parameters for the 'bending points',

$$\lambda_i = \frac{\|A_i\|_2}{\|A_i\|_2 + \|B_i\|_2},$$

as well as $T_i := \gamma(\lambda_i)$ for $i = 1, \dots, k$. Moreover, set $\lambda_0 = 0, T_0 = X$ and $\lambda_{k+1} = 1, T_{k+1} = T$. Then (F2) from Theorem 2.22 implies that $d(X, T) = \sum_{i=1}^{k+1} \|t_i - t_{i-1}\|_2$. We start by proving that for $i = 0, \dots, k-1$ it holds that

$$d(\gamma(\lambda_{i+1}), \gamma(\lambda_i)) = |\lambda_{i+1} - \lambda_i| d(\gamma(0), \gamma(1)) = d(X, T) \quad (2.3)$$

In the following we omit the absolute value, as $\lambda_{i+1} \geq \lambda_i$ by properties of the support. Now T_i and T_{i+1} are contained in the same orthant and due to the parametrization of the geodesic from X to T , this orthant contains the splits $S_i := B_1 \cup \dots \cup B_i \cup A_{i+1} \cup \dots \cup A_k \cup C$. So when omitting the common zeroes in the embedded vectors, their distance is only the Euclidean distance of their weight vectors. In formulae

$$d(T_i, T_{i+1}) = \|t_{i+1} - t_i\|_2 = \sqrt{\sum_{j=1}^i \sum_{s \in B_j} (w_s^{T_i} - w_s^{T_{i+1}})^2 + \sum_{j=i+1}^k \sum_{s \in A_j} (w_s^{T_i} - w_s^{T_{i+1}})^2 + \sum_{s \in C} (w_s^{T_i} - w_s^{T_{i+1}})^2}. \quad (2.4)$$

Using the parametrization of γ we can compute the length of the splits S_i in both trees, which are $w_s^{T_i}$ and $w_s^{T_{i+1}}$ for $s \in S_i$. Therefore we simply plug in λ_i and λ_{i+1} in the parametrization and calculate the split lengths for the three cases $s \in A_j$, $s \in B_j$ or $s \in C$.

Consider $s \in A_j$ for $j \geq i+1$. Then

$$w_s^{T_i} = \frac{(1 - \lambda_i)\|A_j\|_2 - \lambda_i\|B_j\|_2}{\|A_j\|_2} \cdot w_s^X$$

$$w_s^{T_{i+1}} = \frac{(1 - \lambda_{i+1})\|A_j\|_2 - \lambda_{i+1}\|B_j\|_2}{\|A_j\|_2} \cdot w_s^X$$

and subtracting yields

$$w_s^{T_i} - w_s^{T_{i+1}} = (\lambda_{i+1} - \lambda_i) \cdot \frac{\|A_j\|_2 + \|B_j\|_2}{\|A_j\|_2} \cdot w_s^X = (\lambda_{i+1} - \lambda_i) \cdot \frac{w_s^X}{\lambda_j}. \quad (2.5)$$

Next, consider $s \in B_j, j \leq i$. Then

$$\begin{aligned} w_s^{T_i} &= \frac{\lambda_i \|B_j\|_2 - (1 - \lambda_i) \|A_j\|_2}{\|B_j\|_2} \cdot w_s^T \\ w_s^{T_{i+1}} &= \frac{\lambda_{i+1} \|B_j\|_2 - (1 - \lambda_{i+1}) \|A_j\|_2}{\|A_j\|_2} \cdot w_s^T \end{aligned}$$

and subtracting yields

$$w_s^{T_i} - w_s^{T_{i+1}} = -(\lambda_{i+1} - \lambda_i) \cdot \frac{\|A_j\|_2 + \|B_j\|_2}{\|B_j\|_2} \cdot w_s^T. \quad (2.6)$$

We want to reformulate $\frac{\|A_j\|_2 + \|B_j\|_2}{\|B_j\|_2}$ in terms of λ_j . With

$$\begin{aligned} \frac{\|A_j\|_2}{\|A_j\|_2 + \|B_j\|_2} + \frac{\|B_j\|_2}{\|A_j\|_2 + \|B_j\|_2} &= 1 \\ \Leftrightarrow \frac{\|B_j\|_2}{\|A_j\|_2 + \|B_j\|_2} &= 1 - \lambda_j \\ \Leftrightarrow \frac{\|A_j\|_2 + \|B_j\|_2}{\|B_j\|_2} &= \frac{1}{1 - \lambda_j}, \end{aligned}$$

we get

$$w_s^{T_i} - w_s^{T_{i+1}} = -(\lambda_{i+1} - \lambda_i) \cdot \frac{w_s^T}{1 - \lambda_j}.$$

As the third case, consider $s \in C$. Then

$$\begin{aligned} w_s^{T_i} &= (1 - \lambda_i) \cdot w_s^X + \lambda_i \cdot w_s^T \\ w_s^{T_{i+1}} &= (1 - \lambda_{i+1}) \cdot w_s^X + \lambda_{i+1} \cdot w_s^T \end{aligned}$$

and subtracting yields

$$w_s^{T_i} - w_s^{T_{i+1}} = (\lambda_{i+1} - \lambda_i)(w_s^X - w_s^T). \quad (2.7)$$

Now, we plug in the results from (2.5),(2.6),(2.7) into (2.4) and factor out $(\lambda_{i+1} - \lambda_i)^2$:

$$\begin{aligned} d(T_i, T_{i+1}) &= \\ &= \sqrt{\sum_{j=1}^i \sum_{s \in B_j} (w_s^{T_i} - w_s^{T_{i+1}})^2 + \sum_{j=i+1}^k \sum_{s \in A_j} (w_s^{T_i} - w_s^{T_{i+1}})^2 + \sum_{s \in C} (w_s^{T_i} - w_s^{T_{i+1}})^2} \\ &= \sqrt{(\lambda_{i+1} - \lambda_i)^2 \cdot \sum_{j=1}^i \sum_{s \in B_j} \left(\frac{w_s^T}{1 - \lambda_j}\right)^2 + \sum_{j=i+1}^k \sum_{s \in A_j} \left(\frac{w_s^X}{\lambda_j}\right)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2} \\ &= (\lambda_{i+1} - \lambda_i) \cdot \sqrt{\sum_{j=1}^i \sum_{s \in B_j} \left(\frac{w_s^T}{1 - \lambda_j}\right)^2 + \sum_{j=i+1}^k \sum_{s \in A_j} \left(\frac{w_s^X}{\lambda_j}\right)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2} \end{aligned}$$

It turns out that the terms for B_j and A_j cancel out as well, after factoring out $1 - \lambda_j$ or λ_j respectively:

$$\begin{aligned}\sum_{s \in A_j} \left(\frac{w_s^T}{\lambda_j} \right)^2 &= \frac{1}{\lambda_j^2} \sum_{s \in A_j} (w_s^T)^2 = \frac{\|A_j\|_2^2}{\lambda_j^2} = (\|A_j\|_2 + \|B_j\|_2)^2 \\ \sum_{s \in B_j} \left(\frac{w_s^X}{1 - \lambda_j} \right)^2 &= \frac{1}{(1 - \lambda_j)^2} \sum_{s \in B_j} (w_s^X)^2 = \frac{\|B_j\|_2^2}{(1 - \lambda_j)^2} = (\|A_j\|_2 + \|B_j\|_2)^2,\end{aligned}$$

where we used that $\lambda_j = \frac{\|A_j\|_2}{\|A_j\|_2 + \|B_j\|_2}$ and $\sum_{s \in A_j} (w_s^X)^2 = \|A_j\|_2$ (for T and B_j analogously) which hold by definition. Plugging this in yields the desired equality

$$\begin{aligned}d(T_i, T_{i+1}) &= \\ &= (\lambda_{i+1} - \lambda_i) \cdot \sqrt{\sum_{j=1}^i \sum_{s \in B_j} \left(\frac{w_s^T}{1 - \lambda_j} \right)^2 + \sum_{j=i+1}^k \sum_{s \in A_j} \left(\frac{w_s^X}{\lambda_j} \right)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2} \\ &= (\lambda_{i+1} - \lambda_i) \cdot \sqrt{\sum_{j=1}^k (\|A_j\|_2 + \|B_j\|_2)^2 + \sum_{s \in C} (w_s^X - w_s^T)^2} = \\ &= (\lambda_{i+1} - \lambda_i) \cdot d(X, T) = |\lambda_{i+1} - \lambda_i| \cdot d(\gamma(0), \gamma(1)).\end{aligned}$$

Next we show that (2.2) holds for $\lambda = \lambda_0 = 0$ and $\lambda' = \lambda_i$ with $i \in \{1, \dots, k, k+1\}$. For $i = k+1$ we have $\lambda_{k+1} = 1$ and the statement holds trivially. For $i \neq k+1$, we first divide the geodesic $\Gamma(\gamma(0), \gamma(\lambda_i))$ into $\Gamma(\gamma(0), \gamma(\lambda_1))$, $\Gamma(\gamma(1), \gamma(\lambda_2))$, \dots , $\Gamma(\gamma(\lambda_{i-1}), \gamma(\lambda_i))$, using (F3) and (F4) of Theorem 2.22. After that we apply the result (2.3) successively and receive

$$\begin{aligned}d(\gamma(0), \gamma(\lambda_i)) &= \sum_{j=1}^i d(\gamma(\lambda_j), \gamma(\lambda_{j+1})) = \sum_{j=1}^i |\lambda_{j+1} - \lambda_j| \cdot d(\gamma(0), \gamma(1)) \\ &= d(\gamma(0), \gamma(1)) \cdot \sum_{j=1}^i (\lambda_{j+1} - \lambda_j) = (\lambda_i - \lambda_0) \cdot d(\gamma(0), \gamma(1)) \\ &= |\lambda_i| \cdot d(\gamma(0), \gamma(1)).\end{aligned}$$

We generalize once more and prove (2.2) for $\lambda = 0$ and $\lambda' \in (0, 1)$ arbitrary. Here it is important to first recall how the parametrization behaves for two trees Y_1, Y_2 in the same orthant. In this case, the splits of the trees are all contained in the set of double compatible splits $C(Y_1, Y_2)$ and we immediately see that the weight vectors of the parametrization yield a Euclidean straight line segment, i.e., the constant speed property holds, as the Euclidean norm and the straight line satisfy it.

Recall that we consider $\lambda' \in (0, 1)$ arbitrary. When $\lambda' \in \{\lambda_0, \dots, \lambda_{k+1}\}$, then the statement holds as we have shown above. When $\lambda' \notin \{\lambda_0, \dots, \lambda_{k+1}\}$, then there exists an $i \in \{0, \dots, k\}$ such that $\lambda' \in (\lambda_i, \lambda_{i+1})$ and there exists a $\mu \in (0, 1)$ such that $(1 - \mu)\lambda_i + \mu\lambda_{i+1} = \lambda'$. Since $\gamma(\lambda_i) = T_i$ and $\gamma(\lambda_{i+1}) = T_{i+1}$ are contained in the same orthant, the geodesic between them is a Euclidean line segment and satisfies the constant speed property, as discussed above. Moreover, as T_i and T_{i+1} are contained in the geodesic from T to X , $\Gamma(T, X)$, the geodesic between T_i and T_{i+1} , $\Gamma(T_i, T_{i+1})$, satisfies $\Gamma(T_i, T_{i+1}) \subset \Gamma(T, X)$, see (F4) from Theorem 2.22. In particular, this implies that $\gamma(\lambda') \in \Gamma(T_i, T_{i+1})$.

Let $\gamma' : [0, 1]$ denote the parametrization of the geodesic from T_i to T_{i+1} . Then we can calculate

$$\begin{aligned} d(T_i, \gamma(\lambda')) &= d(\gamma'(0), \gamma'(\mu)) = \mu d(\gamma'(0), \gamma'(1)) \\ &= \mu \cdot d(T_i, T_{i+1}) = (\mu \cdot (\lambda_{i+1} - \lambda_i)) \cdot d(X, T). \end{aligned}$$

Using (F3) from Theorem 2.22, we get the desired equality,

$$\begin{aligned} d(\gamma(0), \gamma(\lambda')) &= d(\gamma(0), \gamma(\lambda_i)) + d(\gamma(\lambda_i), \gamma(\lambda')) \\ &= \lambda_i \cdot d(X, T) + \mu(\lambda_{i+1} - \lambda_i) \cdot d(X, T) \\ &= (1 - \mu)\lambda_i + \mu \cdot \lambda_{i+1} d(X, T) = \lambda' \cdot d(X, T). \end{aligned}$$

With this we can finally prove that γ satisfies (2.2) for general $\lambda, \lambda' \in [0, 1]$. W.l.o.g. assume that $\lambda < \lambda'$. Then

$$\begin{aligned} d(\gamma(\lambda), \gamma(\lambda')) &= d(\gamma(0), \gamma(\lambda)) + d(\gamma(\lambda), \gamma(\lambda')) - d(\gamma(0), \gamma(\lambda)) \\ &= d(\gamma(0), \gamma(\lambda')) - d(\gamma(0), \gamma(\lambda)) = \lambda' \cdot d(X, T) - \lambda \cdot d(X, T) \\ &= (\lambda' - \lambda) \cdot d(X, T) = |\lambda' - \lambda| \cdot d(\gamma(0), \gamma(1)). \end{aligned}$$

□

Example 2.2.3 has shown that the support of a geodesic is not always unique. In order to get a unique representation of the geodesic one needs to require more than **(P1)**, **(P2)**, **(P3)**. This motivates the notion of the *minimal support*:

Definition 2.24. [MOP15] A support $(\mathcal{A}, \mathcal{B})$ satisfying **(P2)** with strict inequalities is called minimal support.

Moreover, [MOP12] show (Lemma 1.5) that there always exists a unique minimal support for each geodesic. One receives a minimal support by combining two support pairs who have equal ratios to one pair, i.e., when $\frac{\|A_i\|_2}{\|B_i\|_2} = \frac{\|A_{i+1}\|_2}{\|B_{i+1}\|_2}$, then set $A'_i := A_i \cup A_{i+1}$ and $B'_i := B_i \cup B_{i+1}$ and replace the two sets in the support by their union. Repeating this process eventually yields a minimal support. Note, that combining

two sets of equal ratios in this matter does not change the parametrization, and splitting a set up to two different sets with equal ratio does not as well, as hinted at in Example 2.2.3. This shows that the non-uniqueness arises from only requiring **(P2)** to hold with equality. Nonetheless, it will at times be convenient to also work with a support that is not minimal and since the path and the parametrization coincide, the non-uniqueness is not a big issue.

Even when the support of the geodesic $(\mathcal{A}, \mathcal{B})$ is not explicitly known, it is often helpful to work with it.

If a support *sequence* (not a support!) $(\mathcal{A}, \mathcal{B})$ satisfies the property **(P1)** then we can find a path from X to T through these orthants. But we do not know the points on the boundaries of the orthants yet, which give us the complete path, since the remaining points are defined by straight lines through the orthants. Nonetheless, when **(P2)** additionally holds, then there exists an exact parametrization for such paths, that we call *proper*. The parametrization for this is given in Theorem 2.4 in [OP11]. We slightly changed this to fit our notation.

Theorem 2.25. *Let $X, T \in \mathcal{T}_n$ and let a support sequence $(\mathcal{A}, \mathcal{B})$ be given that satisfies **(P1)** and **(P2)**. Then the path $P(\mathcal{A}, \mathcal{B})$ parametrized as in Theorem 2.21 yields a path from X to T with length*

$$L(P_{\mathcal{A}\mathcal{B}}) = \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2}.$$

Example 2.2.4. *We want to illustrate the concept of supports and the paths they define. Therefore, we return to the paths shown in Example 2.2.2. Again, consider the trees $X = (\{s_2, s_9\}, (2, 2))$ and $T = (\{s_1, s_6\}, (1, 4))$. Let us verify that the path P_1 in Example 2.2.2 is the geodesic by using the concept of supports.*

*Starting at X , the first removed split is s_2 , so $A_1 = \{s_2\}$ and the first split of T to be added is s_6 , so $B_1 = \{s_6\}$. The second removed split is s_9 , the second added split s_6 , so $A_2 = \{s_9\}$ and $B_2 = \{s_1\}$. We verify that **(P1)**, **(P2)** and **(P3)** hold:*

(P1) A_2 only contains s_9 and s_9 is compatible with s_6 , the only split in B_1 .

(P2) The ratio sequence is

$$\frac{\|A_1\|_2}{\|B_1\|_2} = \frac{2}{4} < \frac{2}{1} = \frac{\|A_2\|_2}{\|B_2\|_2}$$

(P3) Since A_i and B_i all contain at most one element, there is no way to split them up.

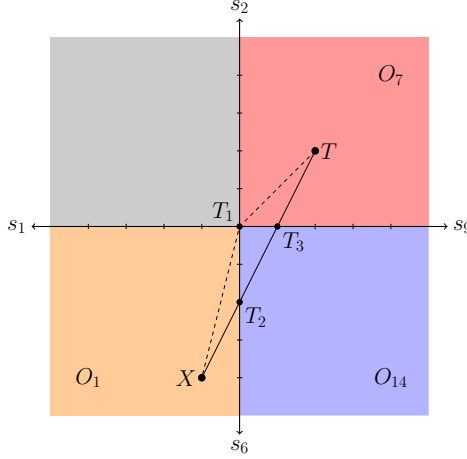


Figure 2.12: Two paths from X to T that are defined by different partitions $(\mathcal{A}, \mathcal{B})$. The geodesic is the solid path.

In contrast, let us now look at the partition belonging to the dashed path P_2 , the cone path. First, all splits of X are removed, so $A_1 = \text{Split}(X) = \{s_2, s_9\}$ and then all splits of T are added, so $B_1 = \text{Split}(T) = \{s_1, s_6\}$. The cone path trivially satisfies **(P1)** and **(P2)**: There exists no $i > j \geq 1$ such that A_i needed to be compatible with B_j and there is only one ratio $\frac{\|A_1\|_2}{\|B_1\|_2}$.

Since we have already verified that the solid line, P_1 is the geodesic, it is clear that **(P3)** can not be satisfied for $(\mathcal{A}, \mathcal{B}) = (\{s_2, s_9\}, \{s_1, s_6\})$. It fails since the partition $C_1 = \{s_2\}$, $D_1 = \{s_6\}$, $C_2 = \{s_9\}$, $D_2 = \{s_1\}$ satisfies $\frac{C_1}{D_1} = 0.5 < 2 = \frac{C_2}{D_2}$: We have already calculated this above because that gives the partition for the geodesic.

The example in \mathcal{T}_4 above gives a good first intuition about how supports and geodesics relate and work. Nonetheless, intuition of only the space \mathcal{T}_4 may be deceitful: For \mathcal{T}_n with $n \geq 5$, there exist cases where the geodesic between X and T traverses a non-maximal orthant, whereas this is not possible in \mathcal{T}_4 . In order to enter a non-maximal orthant, at least $j \geq 2$ splits of X have to be removed in order to be able to add at most $j - 1$ splits of T (at least one must be added). In \mathcal{T}_4 , when two splits are removed, one already has the star tree and the geodesic hence has to be the cone path. The following example shows a geodesic in \mathcal{T}_5 that traverses a non-maximal orthant. This behavior of geodesics has already been mentioned in [Owe11] where a similar example has been given.

Example 2.2.5. *Let*

$$\begin{aligned}
e_1 &= (\{4, 5\}|\{0, 1, 2, 3\}), & e_2 &= (\{1, 2\}|\{0, 3, 4, 5\}), & e_3 &= (\{1, 2, 4, 5\}|\{0, 3\}) \\
X &= \left((e_1, e_2, e_3), \left(\frac{8}{10}, \frac{\sqrt{18}}{10}, \frac{\sqrt{18}}{10} \right) \right), \\
f_1 &= (\{\{2, 5\}|\{0, 1, 3, 4\}\}), & f_2 &= (\{3, 4\}|\{0, 1, 2, 5\}), & f_3 &= (\{2, 3, 4, 5\}|\{0, 1\}) \\
T &= \left((f_1, f_2, f_3), \left(\frac{\sqrt{18}}{10}, \frac{\sqrt{18}}{10}, \frac{8}{10} \right) \right).
\end{aligned}$$

Then the support of the geodesic $\Gamma(X, T)$ is given by

$$(\mathcal{A}, \mathcal{B}) = ((A_1, A_2), (B_1, B_2)) = ((\{e_2, e_3\}, \{e_1\}), (\{f_3\}, \{f_1, f_2\})).$$

Again, we check **(P1)**, **(P2)** and **(P3)**:

(P1) It holds that A_2 and B_1 are compatible as $e_1 = (\{4, 5\}|\{0, 1, 2, 3\}) \in A_2$ and $f_3 = (\{2, 3, 4, 5\}|\{0, 1\}) \in B_1$ are the only splits the sets contain and they are compatible because $\{0, 1\} \subset \{0, 1, 2, 3\}$.

(P2)

$$\frac{\|A_1\|_2}{\|B_1\|_2} = \frac{\sqrt{\frac{18}{100} + \frac{18}{100}}}{\frac{8}{10}} = \frac{3}{4} < \frac{4}{3} = \frac{\frac{8}{10}}{\sqrt{\frac{18}{100} + \frac{18}{100}}} = \frac{\|A_2\|_2}{\|B_2\|_2}$$

(P3) There is no way split up A_1, B_1 into two subsets. First of all, a partition of B_1 would be trivial and additionally no split in A_1 is compatible with f_3 , the only split in B_1 . The same argument holds for A_2, B_2 . Here, e_1 is incompatible with f_1 and f_2 .

The topology of the non-maximal orthant is depicted in the middle of Figure 2.13 and the splitset defining that orthant is $\{e_1, f_3\}$. Finally we want to give the path of the geodesic. The sequence of trees defining the path is (X, I_1, I_2, T) , where $I_1 = ((e_1), (\frac{1}{5}))$ and $I_2 = ((f_3), (\frac{1}{5}))$, where the lengths may be calculated with the parametrization of Theorem 2.21. Recalling Definition 2.11, we see that $\text{Split}(I_1) = \{e_1\} \subset \mathcal{O}(\{e_1, f_3\})$ and $\text{Split}(I_2) = \{f_3\} \subset \mathcal{O}(\{e_1, f_3\})$ and any tree of the leg with I_1 and I_2 as endpoints only contains these two splits. So the geodesic passes through the two-dimensional, non-maximal orthant $\mathcal{O}(\{e_1, f_3\})$.

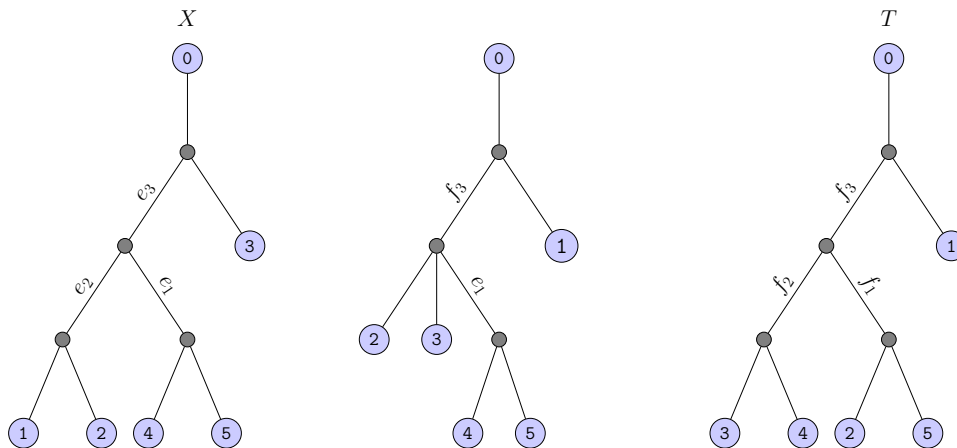


Figure 2.13: The topologies of the trees X and T and the degenerate tree topology on the geodesic $\Gamma(X, T)$.

2.3 Modeling the Species Tree Problem in Tree Space

Now that the tree space and the geodesic distance are thoroughly defined, we give a more formal description of our approach to the species tree problem in tree space. First, recall the species tree problem: Given a set of sample trees (these samples are usually gene trees), find the phylogenetic tree that depicts the correct evolutionary history of the given set of species, i.e., the species tree. Of course there is no proof that the correct tree has been found, so methods that yield candidates for the species tree can generally only yield hypotheses that the resulting tree is the species tree. In comparison to many 'heuristic' rules to infer candidates for the species tree from the set of given gene trees, the tree space \mathcal{T}_n has one big advantage: It allows for an underlying mathematical model to search for the species tree. Thus, when conjecturing that the species tree is the minimizer of a certain function there is a justification for the choice of the tree. Nevertheless, these justifications are limited to the assumptions of the model space being the correct one.

When interpreting the given set of gene trees as samples in tree space, it is a natural thing to statistically investigate the sample (of trees) and try to determine characteristics of this sample in this metric space as, e.g., the sample mean. As we mentioned in the introduction, [OP11] have developed an algorithm to calculate geodesics in polynomial time. With this algorithm at hand it is possible to calculate a sequence converging to the *barycenter* of the given sample trees, which we denote by $\mathbf{T} := \{T_1, \dots, T_M\}$. The barycenter is the minimizer of the *Fréchet variance* F

with respect to a given data set,

$$F(X) = \sum_{i=1}^M d(X, T_i)^2 \quad \text{for } X \in \mathcal{T}_n.$$

This minimizer is also called *Fréchet mean*. [Stu02] has developed an iterative approach in metric spaces of global non-positive curvature that generates a sequence $(X_l)_{l \in \mathbb{N}}$ that converges to the barycenter, i.e.,

$$X_l \rightarrow X^*, \text{ where } f(X^*) = \min_{X \in \mathcal{T}_n} f(X).$$

Note, that the barycenter always exists and is unique in metric spaces of global non-positive curvature as noted in [Stu02]. [Bac14a] and [MOP15] have independently considered the notion of the Fréchet mean in the tree space setting. [Bac14a] has proven a convergence theorem of his algorithm to calculate the mean which is very similar to Sturm’s algorithm [Stu03] but has different ‘step lengths’ in each iteration. A lot of researchers attempt to find better algorithms to calculate the barycenter, or investigate statistical properties of the space. The goal is to eventually be able to define distributions on this space, prove central limit theorems, to do proper statistics and hypothesis testing within the tree space, related papers are [Nye11, BLO⁺13, Nye15, NTWY17].

Nevertheless, all of these lines of research heavily focus on statistics on the space and on finding the barycenter, even though it is not clear, whether the barycenter actually yields good hypotheses for the species tree.

In practice it has turned out that $0 \in \mathcal{T}_n$ often is the unique barycenter and that it still remains the unique minimizer even when perturbing the sample trees significantly. This is a non-desirable property that lead to another research topic discussing *stickiness* of barycenters, see [HHL⁺13].

One possible approach to try to get different results is to change the space and the distance in which the problem is considered. This approach is taken for example in [YZZ17], where the Fréchet mean has been calculated after applying principal component analysis to the set of trees. The foundation to work out statistics on this ultrametric space is attempted in [LMY18].

Unfortunately, the calculations of barycenters in this alternative modeling space seem to be as challenging as for the tree space \mathcal{T}_n , so that it is not clear until now if this approach will turn out to be advantageous. Additionally it is not clear yet if the phenomenon of stickiness is less present in this space. Nevertheless, the approach is rather new and may yield good hypotheses for the species tree or better ways to calculate barycenters with additional research.

Our approach to get different candidates for the species tree is to investigate the problem in \mathcal{T}_n but to take a step back and consider various minimization problems,

not only minimizing the Fréchet variance. From the point of view of a researcher in Location Theory a given set of points is a set of *facilities*; one aims to find the location in the space for which some objective function g , that is non-decreasing in the distances to the given facilities, is minimized. Minimizing the Fréchet variance is just a special kind of a location problem and our goal is to model two other classic location problems in the tree space that yield different trees than the barycenter. We want to find trees X_1^* , X_2^* that minimize

$$g_1(X) = \sum_{i=1}^M d(X, T_i) \quad \text{for } X \in \mathcal{T}_n, \quad \text{or} \quad g_2(X) = \max_{i=1, \dots, M} d(X, T_i) \quad \text{for } X \in \mathcal{T}_n,$$

respectively. Even though it is not clear whether the trees X_1^* , X_2^* perform better than the barycenter X^* in the sense of better inferring the species tree from the set of given trees $\{T_1, \dots, T_M\}$, there are two reasons for investigating them: Firstly, it will turn out that the location problems in tree space have interesting relations to specific Euclidean location problems and are hence interesting to study on their own, disregarding the species tree problem as background. Secondly, since X^* is not guaranteed to be the species tree, it is good to have other methods or guesses that somehow describe 'centroids' of trees based on mathematical models, where optimality of some sense may be proven.

Now that we have motivated the use of Location Theory in tree space we finally formally introduce Location Theory and our location problems in tree space in the next chapter.

3 Location Theory in Tree Space

In the previous section we have thoroughly defined tree space \mathcal{T}_n . In the introduction we motivated the problem that we want to model and solve within the tree space. We want to find a tree that, in some sense, incorporates the structure and similarities of a given set of phylogenetic trees. The hope is that the trees that we get as solution to our models describe the *true* species tree, which is hard to verify. Nevertheless, we develop new approaches, so we also receive other trees as hypotheses for the species tree, which may be presented to biologists as possible solutions.

The chapter is structured as follows: First we give a short introduction to the broad field of *Location Theory*. Then we present our new approach to the problem: interpret the given trees in tree space as *facilities* and model the species tree problem as location problems. At the end of the section we give some general results concerning the split sets of optimal trees.

3.1 Introduction to Location Theory

Location Theory, often times also referred to as *Facility Location* or *Locational Analysis* is a field of research in Optimization, Computational Geometry and Operations research and has a broad range of applications.

Nowadays, problems in *modern* Location Theory are more often than not motivated by an economic background asking for an optimal location of a facility to maximize profit. These basic ideas go back to the 19th century and early 20th century where, for example Alfred Weber wrote his book [Web22] about the location of industry, thereby describing classic problems of Location Theory.

The very first problem of *classic* Location Theory dates back many centuries and may be described as follows: *Given three points in the plane, locate a point in the plane such that the sum of Euclidean distances to the three given points is minimized.* This is the birth of the so-called *median problem* in the case of only three given points. Even though there is a lot of discussion whether this median problem was first introduced by Pierre de Fermat, Evangelista Torricelli or Battista Cavalieri it is certain that the problem was stated as early as in the beginning of the 17th century. The real awakening of Location Theory beyond such small problems took place in the 20th century, when analytical approaches for classic location problems have been developed with which these problems can be solved for more than three points.

After that introduction of analytical methods to Location Theory many researchers started to investigate several kinds of location problems. The basic description of these problems is the following.

One is given a set of facilities $\mathcal{F} = \{F_1, \dots, F_M\}$ that are usually contained in the Euclidean plane, oftentimes these facilities are also assigned weights $w_1, \dots, w_M > 0$ describing their importance.

The task is to locate a point, or, as a typical variant, a set of points, in \mathbb{R}^2 as to minimize some measure of proximity. This measure of proximity is modeled by the objective function, which is usually given by a function which is non-decreasing in the weighted distances to the facilities. When research continued many of these problems have been generalized and solved in \mathbb{R}^n instead of only \mathbb{R}^2 . Classical examples of location problems are the *Weber problem* (L_W), where the sum of weighted distances to the facilities is minimized and the *center problem* (L_{\max}), where the maximum of the weighted distances to the facilities is minimized:

$$\begin{array}{ll}
 (L_W) & \min \sum_{i=1}^M w_i \cdot \|x - F_i\|_2 \\
 & \text{s.t. } x \in \mathbb{R}^n
 \end{array}
 \qquad
 \begin{array}{ll}
 (L_{\max}) & \min \max_{i=1, \dots, M} w_i \cdot \|x - F_i\|_2 \\
 & \text{s.t. } x \in \mathbb{R}^n
 \end{array}$$

Note, that the Weber problem is also referred to as the median problem in the unweighted case. These problems have been generalized in numerous ways: one may consider different spaces \mathcal{X} instead of \mathbb{R}^n and / or change the distance measure $\|\cdot\|_2$ to specific distances d , including gauges, metrics and norms. Moreover, a common variant is to find a minimizer within the set $\mathcal{X} \setminus R$ where R is some forbidden region, where one is neither allowed to place a facility nor may paths connecting a facility and the located point cross this region.

In analogy to the standard location problems in Euclidean space we define and investigate these classic location problems in the tree space in the next section. This is an interesting new ansatz, because the vast majority of location problems is studied in the Euclidean space or even only in the plane. Here we present and start to solve location problems coming from an application that is very different from these standard problems and stresses that Location Theory should not limit itself to Euclidean settings.

Even though we tackle location problems in a much more complicated space, we still benefit from the knowledge about Euclidean location problems, as some of the subproblems that arise may be solved with algorithms for Euclidean location problems.

There are two statements about the Euclidean Weber-problem (L_W) that we will frequently use in the remainder of the thesis:

Theorem 3.1 ([Kuh73]). (i) (L_W) does not have a unique optimal solution if and only if there exists a line $L \subset \mathbb{R}^n$ such that $\{F_1, \dots, F_M\} \subset L$.

(ii) A facility F_k is optimal to (L_W) if and only if

$$\text{Test}_k := \left\| \sum_{\substack{i=1 \\ i \neq k}}^M \frac{w_i(F_i - F_k)}{\|F_i - F_k\|_2} \right\|_2 \leq w_k.$$

3.2 Location Problem Models in Tree Space

Suppose we are given M trees in the phylogenetic tree space \mathcal{T}_n for which we want to determine the species tree. We interpret this set of trees $\mathbf{T} = \{T_1, \dots, T_M\}$ as facilities. Therewith we define and investigate the following three tree space location problems:

$$(P_{\text{Med}}) \min_{X \in \mathcal{T}_n} \sum_{i=1}^M d(X, T_i) \quad \left| \quad (P_{\text{max}}) \min_{X \in \mathcal{T}_n} \max_{i=1, \dots, M} d(X, T_i) \quad \left| \quad (P_{\text{F}}) \min_{X \in \mathcal{T}_n} \sum_{i=1}^M d(X, T_i)^2$$

Note that we only consider the unweighted versions of location problems here, as it is unclear how to purposefully assign weights for a given set of gene trees.

The first two problems, the *median problem* (P_{Med}) and the *center problem* (P_{max}) are the tree space versions of the classical location problems (L_W) and (L_{max}) , that we introduced in Section 3.1.

The third problem (P_{F}) is what we call the *Fréchet problem*. The Fréchet variance is frequently used in statistics. The map that assigns x to the sum of squared distances $\sum_{i=1}^M d(x, y_i)^2$ to a given set of points $\{y_1, \dots, y_M\}$ in some metric space is called *Fréchet variance*. Its minimizer yields the *barycenter*, or *Fréchet mean* of the sample $\{y_1, \dots, y_M\}$, which is a generalization of the arithmetic mean to metric spaces. The Fréchet problem is not the most popular location problem but is definitely of interest here: The barycenter, i.e., the sample mean of our given set of trees is a natural thing to investigate when searching for the species tree and it has been the most common approach for the species tree problem in tree space until now.

Notation 3.2.1. We denote the objective functions of the three location problems by

$$f_{\text{Med}} = \sum_{i=1}^M d(X, T_i), \quad f_{\text{max}} = \max_{i=1, \dots, M} d(X, T_i), \quad f_{\text{F}} = \sum_{i=1}^M d(X, T_i)^2.$$

Most of the times we focus on one of the above problems (P_{\max}) , (P_{Med}) or (P_{F}) at a time and simply write f for the objective function. Whenever there might be ambiguities we specify which objective function we talk about by using the above notation.

In general, solving the problems (P_{\max}) , (P_{Med}) , (P_{F}) is extremely hard. We will see that specific, easy cases of these location problems reduce to their Euclidean counterparts, which implies that they are at least as hard to solve. But we do not have a general approach for (P_{\max}) , (P_{Med}) , (P_{F}) as no one has studied Location Theory in tree space yet.

In order to solve location problems one usually tries to use features of the space or properties of the objective function like differentiability or convexity to design a solution algorithm. For the tree space this is a lot harder than for Euclidean space as it has a lot less structure to work with. For example, the tree space is not a vector space, let alone an inner product space. Nonetheless, there exists recent literature about Hadamard space optimization, [Bac14a], that tries to develop a general theory and framework for optimization within these spaces. Hadamard spaces are complete metric spaces of globally non-positive curvature, and as we mentioned in Section 2.2, the tree space is a Hadamard space. This is why the approximation algorithms from [Bac14b] that calculate the median and the mean in Hadamard spaces may be applied in tree space and yield approximate solutions to the problems (P_{Med}) and (P_{F}) . Nonetheless, there are a lot of reasons to still study these problems as the approximation algorithms converge rather slowly, as we will discuss later, and do not incorporate any of the tree space specific structure. In comparison to a general Hadamard space we have a local Euclidean structure in the orthants and additionally our trees have splits sets that allow to investigate common features of trees. We exploit these structures to state general results about optimal solutions and use it to design solution methods. Later on we will of course compare our tree space specific algorithms with the algorithm Bacak provides for general Hadamard spaces.

3.2.1 Split Sets of Optimal Solutions

The first thing we establish for the tree space location problems are criteria for splits of optimal solutions. In order to prove the correctness of the criteria we need the following lemma that shows that the distance between two trees reduces when a non-common split is removed from one of the trees.

Lemma 3.2. *Suppose that for*

$$X = (\{e_1, \dots, e_{n_1}\}, (w_1, \dots, w_{n_1})), T = (\{f_1, \dots, f_{n_2}\}, (w_1, \dots, w_{n_2})) \in \mathcal{T}_n$$

there exists a split f_l such that $f_l \in \text{Split}(T) \setminus \text{Split}(X)$.

Let $T' = (\{f_1, \dots, f_{l-1}, f_{l+1}, \dots, f_{n_2}\}, (w_1, \dots, w_{l-1}, w_{l+1}, \dots, w_{n_2})) \in \mathcal{T}_n$, i.e., the tree which is derived by taking X and removing f_l . Then $d(X, T) < d(X, T')$.

Proof. Suppose that the geodesic $\Gamma(X, T)$ is given by the path

$$P = (X = T_0, T_1, \dots, T_k = T),$$

i.e., $d(X, T) = L(\Gamma(X, T)) = \sum_{j=1}^k \|t_j - t_{j-1}\|_2$. Since $f_l \in \text{Split}(T) = \text{Split}(T_k)$ but $f_l \notin \text{Split}(X) = \text{Split}(T_0)$, there exists some $i \in 1, \dots, k$ such that $f_l \in \text{Split}(T_j)$ for $j \geq i$ and $f_l \notin \text{Split}(T_j)$ for $j < i$.

With this we define a path P' from X to T' by modifying the sequence of trees that yield the geodesic path. Let $P' = (X = T_0, T_1, \dots, T_{i-1}, T'_i, T'_{i+1}, T'_k = T')$ where we receive T'_j by removing the split f_l , as it was done for T' .

Now let $l' \in \{1, \dots, N\}$ be the component of the embedded vectors $t_j, t'_j \in \mathbb{R}_+^N$ that corresponds to the length of the split f_l . Then we have $t'_{j l'} = 0$ for all $j = 1, \dots, k$, as T'_j was derived from T_j by removing f_l . Moreover, the length of the split f_l grows monotonically from T_i, T_{i+1}, \dots to $T_k = T$, which follows from the parametrization of the geodesic from X to T , and thus $t_{j l'} > t_{j-1 l'}$ for $j > i$. Then it holds that

$$\begin{aligned} 0 &= |t'_{i l'} - t'_{i-1 l'}| \leq |t_{i l'} - t_{i-1 l'}| && \text{since } t_{i l'} \geq 0 = t'_{i l'} = t_{i-1 l'}, \\ 0 &= |t'_{j l'} - t'_{j-1 l'}| < |t_{j l'} - t_{j-1 l'}| && \text{since } t_{j l'} > t_{j-1 l'} > 0 = t'_{j l'} = t'_{j-1 l'} \quad \forall j > i. \end{aligned}$$

By definition of the T'_j , we have $t_{j l} = t_{j l'}$ for all other components $l \neq l' \in \{1, \dots, N\}$. Together this yields

$$\begin{aligned} L(P') &= \sum_{j=1}^k \|t'_j - t'_{j-1}\|_2 = \sum_{j=1}^{i-1} \|t_j - t_{j-1}\|_2 + \sum_{j=i}^k \|t'_j - t'_{j-1}\|_2 \\ &< \sum_{j=1}^{i-1} \|t_j - t_{j-1}\|_2 + \sum_{j=i}^k \|t_j - t_{j-1}\|_2 = \sum_{j=1}^k \|t_j - t_{j-1}\|_2 \\ &= L(P) = d(X, T). \end{aligned}$$

This implies that we have found a path P' from X to T' that satisfies $L(P') < d(X, T)$. Since $d(X, T')$ is the infimum over the lengths of all paths from X to T' it follows that $d(X, T') \leq L(P') < d(X, T)$. \square

Having this auxiliary lemma we can show that there are certain splits that have to be contained in $\text{Split}(T)$ for any optimal tree T .

Theorem 3.3. *Let $e \in \bigcap_{i=1}^M \text{Split}(T_i)$. Then $e \in \text{Split}(X^*)$ for all optimal solutions $X^* \in \mathcal{T}_n$ to the median problem (P_{Med}) and $e \in \text{Split}(X^*)$ for all optimal solutions $X^* \in \mathcal{T}_n$ to the center problem (P_{max}).*

Proof. In order to prove that any optimal solution X^* to (P_{Med}) or (P_{max}) has to satisfy $e \in \text{Split}(X^*)$, we show that any tree $X \in \mathcal{T}_n$ with $e \notin \text{Split}(X)$ is not optimal.

First, assume that for $X \in \mathcal{T}_n$ there exists some split $s \in \text{Split}(X)$ such that s is incompatible with e . Apply Lemma 3.2 to X and X' where X' is derived from X by removing s . This yields

$$d(X', T_i) < d(X, T_i) \quad \forall i = 1, \dots, M.$$

Thus,

$$\sum_{i=1}^M d(X', T_i) < \sum_{i=1}^M d(X, T_i) \quad \text{and} \quad \max_{i=1}^M d(X', T_i) < \max_{i=1}^M d(X, T_i)$$

so X cannot be an optimal solution for (P_{Med}) or (P_{max}) .

Now, let $X = ((s_1, \dots, s_{n_1}), (w_1, \dots, w_{n_1}))$ be a tree for which no $s \in \text{Split}(X)$ is incompatible with e . If $e \in \text{Split}(X)$ then the statement holds, so we assume that $e \notin \text{Split}(X)$. As no split in $\text{Split}(X)$ is incompatible with e this implies that $\text{Split}(X) \cup \{e\}$ is a compatible set of splits and recalling Definition 2.18, e is a double compatible split of X and T_i for all $i = 1, \dots, M$.

Now define $c := \min_{i=1, \dots, M} w_e^{T_i} > 0$, i.e., c is the minimal length of the split e among all T_i . We now construct a tree X' for which $d(X', T_i) < d(X, T_i)$ for all $i = 1, \dots, M$. Let $X' = ((s_1, \dots, s_{n_1}, e), (w_1, \dots, w_{n_1}, c))$ and fix a tree T_i . As e is a double compatible split of T_i and X as well as for T_i and X' it holds that $\text{Split}(X) \setminus C = \text{Split}(X') \setminus C$. So if $(\mathcal{A}, \mathcal{B}) = ((A_1, \dots, A_k), (B_1, \dots, B_k))$ is a support of the geodesic from X to T , i.e., satisfies $(P1) - (P3)$, then it also is a support of the geodesic from X' to T , as the weights of all splits $s \in \text{Split}(X) \setminus C = \text{Split}(X') \setminus C$ coincide. Now, apply the length formula for the geodesic from Theorem 2.21 for both trees:

$$d(X, T_i) = \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^{T_i} - w_s^X)^2}$$

$$d(X', T_i) = \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^{T_i} - w_s^{X'})^2}$$

As the support and all split lengths of X and X' are the same except for the split $e \in C$, we get

$$d(X, T_i)^2 - d(X', T_i)^2 = (w_e^{T_i} - w_e^X)^2 - (w_e^{T_i} - w_e^{X'})^2 > 0$$

due to

$$(w_e^{T_i} - w_e^X)^2 = (w_e^{T_i})^2 > (w_e^{T_i} - c)^2 = (w_e^{T_i} - w_e^{X'})^2,$$

where $(w_e^{T_i})^2 > (w_e^{T_i} - c)^2$ follows from the choice of c . Hence, we get $d(X, T_i) > d(X', T_i)$ for all $i = 1, \dots, M$ which implies

$$\sum_{i=1}^M d(X', T_i) < \sum_{i=1}^M d(X, T_i) \quad \text{and} \quad \max_{i=1}^M d(X', T_i) < \max_{i=1}^M d(X, T_i),$$

so X cannot be an optimal solution for (P_{Med}) or (P_{max}) . \square

We have seen that there are certain splits that are always in the split sets of optimal trees. The counterpart of this idea is to ask which splits are never contained in the split set of optimal trees.

Theorem 3.4. *Let $e \notin \cup_{i=1}^M \text{Split}(T_i)$. Then $e \notin \text{Split}(X^*)$ for any optimal solution X^* to (P_{Med}) and $e \notin \text{Split}(X^*)$ for any optimal solution $X^* \in \mathcal{T}_n$ to (P_{max}) .*

Proof. Let $X = ((s_1, \dots, s_{n_X}), (w_1, \dots, w_{n_X})) \in \mathcal{T}_n$ be such that $e \in \text{Split}(X)$, i.e., $e = s_i$ for some $i \in \{1, \dots, n_X\}$. Then define X' by removing s_i ,

$$X' = ((s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{n_X}), (w_1, \dots, w_{i-1}, w_{i+1}, w_{n_X})),$$

as in 3.2. By the prerequisite it holds that $e \notin \text{Split}(T_i)$ for $i = 1, \dots, M$. Applying Lemma 3.2 then yields

$$d(X', T_i) < d(X, T_i) \quad \forall i = 1, \dots, M$$

so X cannot be optimal for (P_{Med}) or (P_{max}) as X' yields a lower objective. \square

The analogous statements are known for the Fréchet variance and have been shown in [MOP12], Lemma 5.1. Additionally, the Fréchet problem has a unique optimal solution, as shown in [Stu03]. We summarize this:

Theorem 3.5. *[MOP12],[Stu03] (P_{F}) has a unique optimal solution X^* . Moreover, if $e \notin \cup_{i=1}^M \text{Split}(T_i)$ then $e \notin \text{Split}(X^*)$ and if $e \in \bigcap_{i=1}^M \text{Split}(T_i)$ it follows that $e \in \text{Split}(X^*)$.*

4 Solving Tree Space Location Problems by Transformations to Euclidean Location Problems

In the previous chapter we have introduced Location Theory and modeled and defined the location problems in tree space that we want to solve. We already mentioned that it is in general a very difficult task to solve these problems.

Therefore our approach is to first try to understand the structure of simple versions of the problems (P_{\max}) , (P_{Med}) , (P_{F}) , e.g., problems in low dimensions or problems where we have strong assumptions on the given set of trees $\mathbf{T}\{T_1, \dots, T_M\}$. This helps to understand the key difficulties in solving these problems and, if possible, we try to generalize the methods for these easy cases to a more general setting. We start with the lowest-dimensional, non-trivial tree space, that is, \mathcal{T}_3 .

4.1 Location Problems in \mathcal{T}_3

In this section we completely describe the solutions of the location problems (P_{\max}) , (P_{Med}) , (P_{F}) in \mathcal{T}_3 . The solutions to the problems are easy to obtain once we see that the \mathcal{T}_3 is just a star graph with three infinitely long edges emanating from the center vertex. Hence, the location problems in \mathcal{T}_3 boil down to well-known location problems on networks.

First we investigate the structure of \mathcal{T}_3 : There are only three splits on the set of leaves $\{0, 1, 2, 3\}$, these being $s_1 = (\{2, 3\}|\{0, 1\})$, $s_2 = (\{1, 3\}|\{0, 2\})$, $s_3 = (\{1, 2\}|\{0, 3\})$. Moreover, maximal orthants of \mathcal{T}_n consist of $n - 2$ splits; this means that a maximal set of compatible splits has cardinality 1 in \mathcal{T}_3 . These simple observations have already been stated in Example 2.2.1, where the embedding of \mathcal{T}_3 into \mathbb{R}^3 was given. Furthermore, geodesics in \mathcal{T}_3 are easy to obtain: If two trees X_1, X_2 are contained in the same orthant, the geodesic is the straight line connecting the two. If X_1 and X_2 are in different orthants then the geodesic is the cone path, since all three splits in \mathcal{T}_3 are pairwise incompatible.

Denote $\mathcal{O}_i := \mathcal{O}(\{s_i\})$ for $i = 1, 2, 3$ and $\mathbf{T}_i := \{T \in \mathbf{T} : T \in \mathcal{O}_i\}$ for $i = 1, 2, 3$, which are the facilities in the respective orthants. Moreover, let $l_i = \max_{T \in \mathbf{T} \cap \mathcal{O}_i} \|t\|_2$

for $i = 1, 2, 3$, where $t \in \mathbb{R}_+^3$ is the embedded vector of edge lengths, compare Definition 2.9.

Theorem 4.1. *Let $\mathbf{T} = \{T_1, \dots, T_m\} \subset \mathcal{T}_3$ and deploy the notation introduced above. Then the solutions to the problem are characterized as follows:*

(P_{\max}) : *Let $i \in \{1, 2, 3\}$ be the index with $l_i \geq l_j$ for $j \neq i$, then the optimal solution is contained in \mathcal{O}_i with the length of s_i being $\frac{l_i - l_j}{2}$ where l_j is the second biggest value among l_1, l_2, l_3 .*

(P_{Med}) : *If there exists no $i \in \{1, 2, 3\}$ such that $|\mathbf{T}_i| \geq |\mathbf{T} \setminus \mathbf{T}_i|$ then 0 is the unique optimal solution. Otherwise let i be that index and suppose that $\mathbf{T}_i = \{T_{j_1}, \dots, T_{j_k}\}$ is ordered such that $\|t_{j_i}\|_2 \geq \|t_{j_{i+1}}\|_2$ and let*

$$d = \begin{cases} \frac{|\mathbf{T}_i| - |\mathbf{T} \setminus \mathbf{T}_i|}{2} & \text{for } \mathbf{T} \text{ even} \\ \frac{|\mathbf{T}_i| - |\mathbf{T} \setminus \mathbf{T}_i| + 1}{2} & \text{for } \mathbf{T} \text{ odd} \end{cases}$$

Then for $|\mathbf{T}|$ odd, the unique optimal solution is T_d and for $|\mathbf{T}|$ even the optimal solution is every tree on the edge between T_{d+1} and T_d , where $T_0 := 0$ is the star tree.

(P_{F}) : *If there exists no $i \in \{1, 2, 3\}$ such that $\sum_{T \in \mathbf{T}_i} \|t\|_2 \geq \sum_{T \in \mathbf{T} \setminus \mathbf{T}_i} \|t\|_2$ then 0 is the unique optimal solution. Otherwise let i be that index, then the unique optimal solution X is contained in \mathcal{O}_i and the length of $s_i \in \text{Split}(X)$ is*

$$\frac{1}{M} \left(\sum_{T \in \mathbf{T}_i} \|t\|_2 - \sum_{T \in \mathbf{T} \setminus \mathbf{T}_i} \|t\|_2 \right).$$

These results may be obtained by calculating the objective of the above solutions and showing that the objective value for all non-optimal trees is higher by estimations. Instead of doing this we use this possibility to point out the equivalence of the above location problems in \mathcal{T}_3 to specific location problems on networks. We refer to [LPT95] for a nice overview over facility location on networks.

In [LPT95] one is given a set of facilities $V = \{v_1, \dots, v_M\}$ that are contained in the network N . Conveniently, [LPT95] defines networks as subsets of \mathbb{R}^d with a finite amount of vertices and the edges are the straight line segments connecting the vertices. Our goal is to model our problems in this setup.

As all objective functions are non-decreasing in $d(X, T_i)$, it is clear that there exists no optimal solution in \mathcal{O}_i with a weight greater than the maximal weight of trees in this orthant l_i . This way, we may bound the star graph with infinitely long edges: Let $O_i := \{T \in \mathcal{O}_i : w_{s_i}^T \leq l_i\}$ for $i = 1, 2, 3$ and let $\mathbb{T}_3 := \cup_{i=1,2,3} O_i$.

Now simply setting $N = \chi(\mathbb{T}_3)$ and $V = \chi(\mathbf{T})$, we may apply Theorem 3.5 [LPT95] to receive the solutions to the median problem and Theorem 4.4 [LPT95] to receive the unique solution to the center problem as described in Theorem 4.1.

In order to get the solution to the Fréchet problem, we note another connection of \mathcal{T}_3 to a geometrical object called “open book”, [HHL⁺13]. The authors of [HHL⁺13] mention \mathcal{T}_3 as one of the first examples for open books. Open books consist of “pages” of dimension d , here the orthants, that are joined along a common “spine” of dimension $d - 1$. Hence, the \mathcal{T}_3 is a very simple example of an open book. The solution to the Fréchet problem as stated in Theorem 4.1 is then gained by applying Theorem 2.9 in combination with Lemma 3.3 from [HHL⁺13].

4.2 Adjacent Orthants

In this section we do not restrict ourselves to the very simple low-dimensional case $n = 3$. The approach here is to try to find an isometry to \mathbb{R}^{n-2} equipped with an appropriate distance that yield us familiar and solvable location problems.

We show that there exists such an isometry under the following assumption: The existing trees $\mathbf{T} = \{T_1, \dots, T_M\}$ satisfy $\mathbf{T} \subset \mathcal{O}(S_1) \cup \mathcal{O}(S_2)$ for two *adjacent* maximal orthants $\mathcal{O}(S_1), \mathcal{O}(S_2)$.

Definition 4.2. *Two maximal orthants $\mathcal{O}(S_1), \mathcal{O}(S_2)$ in \mathcal{T}_n are called adjacent if $|S_1 \cap S_2| = n - 3$.*

So, if two maximal orthants are adjacent, there exists one split $e \in S_1 \setminus S_2$ and one split $f \in S_2 \setminus S_1$, respectively. Note that e and f are incompatible, because otherwise $S_2 \cup \{e\}$ would be a set of compatible splits with $n - 1$ elements which is not possible.

Let us abbreviate $\mathcal{O}_i = \mathcal{O}(S_i), i = 1, 2$ in what follows. To further simplify notation let us assume (w.l.o.g. after renumbering the splits in \mathcal{S}) that $S_1 \cap S_2 = \{s_1, s_2, \dots, s_{n-3}\}$, $S_1 \setminus S_2 = \{s_{n-2}\}$ and $S_2 \setminus S_1 = \{s_{n-1}\}$. For two trees

$$\begin{aligned} T &= (\{s_1, \dots, s_{n-2}\}, (w_1, \dots, w_{n-2})) \in \mathcal{O}(S_1) \text{ and} \\ X &= (\{s_1, \dots, s_{n-3}, s_{n-1}\}, (w'_1, \dots, w'_{n-2})) \in \mathcal{O}(S_2) \end{aligned}$$

we then have that the embedded vectors (compare Definition 2.9) $\chi(T) = t, \chi(X) = x$ are given by

$$t = \begin{pmatrix} w_1 \\ \vdots \\ w_{n-2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, x = \begin{pmatrix} w'_1 \\ \vdots \\ w'_{n-3} \\ 0 \\ w'_{n-2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (4.1)$$

This is a case where the two trees T and X have all but one split in common, so $\{s_1, \dots, s_{n-3}\} \subset C$ for the set of double compatible splits, compare Definition 2.18. So the support $(\mathcal{A}, \mathcal{B})$ only needs to take care of the splits s_{n-2}, s_{n-1} which are incompatible. Hence, $\mathcal{A} = (A_1) = (\{s_{n-2}\})$ and $\mathcal{B} = (B_1) = (\{s_{n-1}\})$. Using the distance formula from the parametrization in Theorem 2.21 we receive

$$\begin{aligned} d(X, T) &= \sqrt{(\|A_1\|_2 + \|B_1\|_2)^2 + \sum_{s \in C} (w_s^T - w_s^X)^2} \\ &= \sqrt{(w_{n-2} + w'_{n-2})^2 + \sum_{s \in C} (w_s^T - w_s^X)^2}. \end{aligned}$$

Using the embedded vectors $t, x \in \mathbb{R}_+^N$ this yields the following formula:

Lemma 4.3. *For T, X being two trees in adjacent orthants $\mathcal{O}_1, \mathcal{O}_2$ with their embedded vectors t, x as in (4.1), the geodesic distance between them is*

$$d(T, X) = \sqrt{\sum_{i=1}^{n-3} (t_i - x_i)^2 + (t_{n-2} + x_{n-1})^2}$$

We now define the map $\varphi : \mathcal{O}_1 \cup \mathcal{O}_2 \rightarrow \mathbb{R}^{n-2}$ by

$$\varphi(Y) = \begin{pmatrix} \chi(Y)_1 \\ \vdots \\ \chi(Y)_{n-3} \\ \chi(Y)_{n-2} \end{pmatrix} \text{ for } Y \in \mathcal{O}_1, \quad \varphi(Y) = \begin{pmatrix} \chi(Y)_1 \\ \vdots \\ \chi(Y)_{n-3} \\ -\chi(Y)_{n-1} \end{pmatrix} \text{ for } Y \in \mathcal{O}_2,$$

where we again used that the embedded vectors have the form of (4.1). We want to verify that $\varphi : (\mathcal{O}_1 \cup \mathcal{O}_2, d) \rightarrow (\mathbb{R}^{n-2}, \|\cdot\|_2)$ is an isometry: For $A \in \mathcal{O}_1, B \in \mathcal{O}_2$ and $a = \chi(A), b = \chi(B)$ we have

$$\begin{aligned} \|\varphi(A) - \varphi(B)\|_2 &= \left\| \begin{pmatrix} a_1 \\ \vdots \\ a_{n-3} \\ a_{n-2} \end{pmatrix} - \begin{pmatrix} b_1 \\ \vdots \\ b_{n-3} \\ -b_{n-1} \end{pmatrix} \right\|_2 \\ &= \sqrt{\sum_{i=1}^{n-3} (a_i - b_i)^2 + (a_{n-2} + b_{n-1})^2} = d(A, B), \end{aligned}$$

where the last equality follows from Lemma 4.3. For the case that we have two trees A, B that are in the same orthant \mathcal{O}_i for $i \in \{1, 2\}$, we have that the distance

between them is the Euclidean distance by definition of the geodesic distance, so $\|\varphi(A) - \varphi(B)\|_2 = d(A, B)$ holds as well.

This isometry allows to reduce (P_{Med}) , (P_{max}) and (P_F) to equivalent location problems in \mathbb{R}^{n-2} , where we call two problems equivalent if their objective values coincide and there is a one-to-one map between their set of minimizers.

Theorem 4.4. *Let $\mathbf{T} \subset \mathcal{O}_1 \cup \mathcal{O}_2$ and f be one of the three objective functions $\{f_{\text{Med}}, f_{\text{max}}, f_F\}$. Then*

$$(P) \quad \min_{X \in \mathcal{T}_n} f(X) \quad \Leftrightarrow \quad (\tilde{P}) \quad \min_{x \in \mathbb{R}^{n-2}} \tilde{f}(x)$$

s.t. $X \in \mathcal{T}_n$ s.t. $x \in \mathbb{R}^{n-2}$,

where \tilde{f} depends on f and $\varphi(\mathbf{T})$ as follows:

$$\begin{aligned} f = f_{\text{max}} &\rightarrow \tilde{f}(x) = \max_{i=1, \dots, M} \|x - \varphi(T_i)\|_2 \\ f = f_{\text{Med}} &\rightarrow \tilde{f}(x) = \sum_{i=1, \dots, M} \|x - \varphi(T_i)\|_2 \\ f = f_F &\rightarrow \tilde{f}(x) = \sum_{i=1, \dots, M} (\|x - \varphi(T_i)\|_2)^2 \end{aligned}$$

Proof. First, for $X \in \mathcal{O}(S_1) \cup \mathcal{O}(S_2)$ we define $u = \varphi(X) \in \mathbb{R}^{n-2}$. Since φ is an isometry it holds that $d(X, T_i) = \|\varphi(X) - \varphi(T_i)\|_2 = \|u - \varphi(T_i)\|_2$, and thus for any choice of $f \in \{f_{\text{Med}}, f_{\text{max}}, f_F\}$ we have $\tilde{f}(u) = f(X)$.

We now show that φ and φ^{-1} (restricted on $\text{im}(\varphi)$) map a minimizer of one problem to a minimizer of the other problem.

First, let u^* be optimal for (\tilde{P}) and let $T^* = \varphi^{-1}(u^*)$. We want to show that T^* is optimal for (P) . To this end, assume T^* is not an optimal solution for (P) and let T' be the minimizer.

Theorem 3.3 and the analogous statements for the Fréchet variance in [MOP12], Lemma 5.1 (depending on which objective function $\{f_{\text{Med}}, f_{\text{max}}, f_F\}$ is deployed) imply that all optimal trees to (P) have to contain all $n - 3$ splits in $S_1 \cap S_2$ in their set of splits. Additionally, Lemma 3.2 implies that all optimal trees X' satisfy a split $\text{Split}(X') \setminus S_1 \cup S_2$. So any split $s \in \text{Split}(X')$ other than s_1, \dots, s_{n-3} is either $\{e\} = S_1 \setminus S_2$ or $\{f\} \in S_2 \setminus S_1$.

Together this shows that the optimal solution T' has to be in $\mathcal{O}(S_1) \cup \mathcal{O}(S_2)$. Setting $u' := \varphi(T') \in \mathbb{R}^{n-2}$, we derive

$$\tilde{f}(u') = f(T') < f(T^*) = \tilde{f}(u^*),$$

which is a contradiction to the optimality of u^* . Thus, T^* is an optimal solution for (P) .

For the other direction let T^* be optimal for (P) . We have seen above that this implies $T^* \in \mathcal{O}(S_1) \cup \mathcal{O}(S_2)$, so we may define $u^* := \varphi(T^*)$. Assume $\tilde{f}(u') <$

$\tilde{f}(u^*)$ for some $u' \in \mathbb{R}^{n-2}$, i.e., that u^* is not optimal. Since all objective functions $\{f_{Med}, f_{max}, f_F\}$ are convex it follows that the optimal solution to (\tilde{P}) is contained in $\text{conv}(\varphi(\mathbf{T}))$. Then u' is a convex combination of the $\varphi(T_i)$, hence $u'_1, \dots, u'_{n-3} \geq 0$, as this holds for all T_i . This implies that $u' \in \text{im}(\varphi)$, so we may set $T' := \varphi^{-1}(u')$. Then $f(T') = \tilde{f}(u') < \tilde{f}(u^*) = f(T^*)$, again a contradiction. \square

Theorem 4.4 allows for an easy solution procedure for our location problems in tree space, when $\mathbf{T} \subset \mathcal{O}_1 \cup \mathcal{O}_2$:

- 1 Map the existing facilities \mathbf{T} to $\varphi(\mathbf{T}) \subset \mathbb{R}^{n-2}$.
- 2 Use a solution algorithm for the Euclidean problem to get an optimal solution u^* to (\tilde{P}) .
- 3 $\varphi^{-1}(u^*) \in \mathcal{O}(S_1) \cup \mathcal{O}(S_2)$ is an optimal solution to (P) .

Hence, the location problems in tree space for the adjacent orthant case are as easy to solve as their respective Euclidean versions. Naturally, step 2 depends on the specific objective function f that we have given. For our investigated objective functions, f_{Med}, f_{max}, f_F we can apply the Weiszfeld Algorithm, the Elzinga-Hearn Algorithm or compute the arithmetic mean respectively to determine the optimal solution to the Euclidean problem. For the center and the Fréchet problem exact solutions are obtained with short computation times and for the median problem there exist several fast algorithms that yield approximate optimal solutions, e.g., the Weiszfeld algorithm.

4.3 Completely Incompatible Orthants

After we have solved the adjacent orthant case, where the two maximal sets of splits S_1, S_2 are the same but for one split we now consider the other extreme case. Here we assume that we have $\kappa \geq 2$ maximal orthants containing the data, i.e., $\mathbf{T} \subset \mathcal{O}_1 \cup \dots \cup \mathcal{O}_\kappa \subset \mathcal{T}_n$, where the \mathcal{O}_j are pairwise *completely incompatible*.

Definition 4.5. *Two sets of splits S_1, S_2 are called completely incompatible if e and f are incompatible splits for any $e \in S_1, f \in S_2$. Moreover, two orthants $\mathcal{O}(S_1), \mathcal{O}(S_2)$ are called completely incompatible if S_1 and S_2 are completely incompatible.*

Even though this case is somehow the opposite of the adjacent orthant case, the basic approach is still the same. We use specific knowledge about the path of the geodesic between two trees in completely incompatible orthants and formulate equivalent location problems in \mathbb{R}^{n-2} . Again we abbreviate $\mathcal{O}_i = \mathcal{O}(S_i), i = 1, 2, \dots, \kappa$ in what follows.

Lemma 4.6. [BHV01] *If $\mathcal{O}_1, \mathcal{O}_2$ are completely incompatible, then for all $X \in \mathcal{O}_1$ and $T \in \mathcal{O}_2$ it holds that the geodesic is the cone path (compare p. 22) and*

$$d(X, T) = \|x\|_2 + \|t\|_2.$$

First we show that for the completely incompatible orthants we have easily verifiable conditions that enable us to find an orthant that contains an optimal solution with little effort. We start by providing conditions for (P_{Med}) and (P_{max}) and later for (P_{F}) , for which we need some auxiliary lemmas.

Theorem 4.7. *i) If there exists $k \in \{1, \dots, \kappa\}$ such that $|\mathbf{T} \cap \mathcal{O}_k| > \frac{M}{2}$, then all optimal solutions to (P_{Med}) are contained in \mathcal{O}_k . Otherwise 0 is an optimal solution to (P_{Med}) .*

ii) If there exists $k \in \{1, \dots, \kappa\}$ such that all facilities in $\arg \max_{T_i \in \mathbf{T}} \|t_i\|_2$ are in one orthant, say \mathcal{O}_k , then the optimal solution to (P_{max}) is contained in \mathcal{O}_k . Otherwise 0 is the unique optimal solution to (P_{max}) .

Proof. We first prove that all optimal solutions to both of the problems have to be contained in $\mathcal{O}_1 \cup \dots \cup \mathcal{O}_\kappa$. Theorem 3.4 implies that for each of the three objective functions it holds that every optimal solution $X^* \in \mathcal{T}_n$ has to satisfy $\text{Split}(X^*) \subset \text{Split}(T_1) \cup \dots \cup \text{Split}(T_M) \subseteq S_1 \cup \dots \cup S_\kappa$, where S_i is the set of splits of \mathcal{O}_i . Since S_i and S_j are incompatible for each $i \neq j$, $\text{Split}(X^*) \subset S_1 \cup \dots \cup S_\kappa$ implies that $\text{Split}(X^*) \subset S_i$ for some $i \in \{1, \dots, \kappa\}$, i.e., $X^* \in \mathcal{O}_i$ for some $i \in \{1, \dots, \kappa\}$. We start with statement *i*): Let \mathcal{O}_k satisfy $|\mathbf{T} \cap \mathcal{O}_k| > \frac{M}{2}$. Then for $0 \neq X \in \mathcal{O}_i \neq \mathcal{O}_k$ the estimation

$$\begin{aligned} f_{\text{Med}}(X) &= \sum_{T \in \mathbf{T}} d(X, T) = \sum_{T \in \mathbf{T} \setminus \mathcal{O}_i} (\|x\|_2 + \|t\|_2) + \sum_{T \in \mathbf{T} \cap \mathcal{O}_i} \|t - x\|_2 \\ &\geq \sum_{T \in \mathbf{T} \setminus \mathcal{O}_i} (\|t\|_2 + \|x\|_2) + \sum_{T \in \mathbf{T} \cap \mathcal{O}_i} \|t\|_2 - |\mathbf{T} \cap \mathcal{O}_i| \cdot \|x\|_2 \\ &= \sum_{T \in \mathbf{T}} \|t\|_2 + (|\mathbf{T} \setminus \mathcal{O}_i| - |\mathbf{T} \cap \mathcal{O}_i|) \cdot \|x\|_2 \\ &> \sum_{T \in \mathbf{T}} \|t\|_2 = \sum_{T \in \mathbf{T}} d(0, T) = f_{\text{Med}}(0) \end{aligned}$$

shows that the optimal solution has to be contained in \mathcal{O}_k .

Next, we prove statement *ii*): If there exist two or more different orthants containing a facility $T_k \in \arg \max_{T \in \mathbf{T}} \|t\|_2$, then the unique optimal solution is $X^* = 0 \in \mathcal{T}_n$: If $0 \neq X \in \mathcal{O}_i$, then we can choose a facility with maximal norm $T_k \in \mathcal{O}_j$ with $j \neq i \in \{1, \dots, \kappa\}$ and

$$f_{\text{max}}(X) \geq d(X, T_k) = \|x\|_2 + \|t_k\|_2 > \|t_k\|_2 = f_{\text{max}}(0), \quad (4.2)$$

showing that 0 yields a better objective value than any $0 \neq X \in \mathcal{O}$ with \mathcal{O} arbitrary, so 0 is the unique optimal solution in this case.

The other case is that there exists an orthant \mathcal{O}_k such that all facilities with maximal norm are contained in \mathcal{O}_k and let T_l be one of them. Then for any orthant $\mathcal{O}_i \neq \mathcal{O}_k$ and an arbitrary point $0 \neq X \in \mathcal{O}_i$ it holds that

$$f_{\max}(0) = \max_{i=1, \dots, M} \|t_i\|_2 = \|t_l\|_2 < \|t_l\|_2 + \|x\|_2 = f_{\max}(X),$$

so X cannot be optimal. Hence, the optimal solution has to be in \mathcal{O}_k . \square

As we mentioned we want to show a similar result for (P_F) . To simplify notation, we write f for f_F in the following. In order to prove the condition if an optimal solution is contained in an orthant we need the directional derivatives of f at 0 in the directions of a fixed orthant \mathcal{O}_j . To this end we use that the objective in each orthant \mathcal{O}_j , $j = 1, \dots, \kappa$ may be written as

$$f(X) = \sum_{T \in \mathbf{T}} d(T, X)^2 = \sum_{T \in \mathbf{T} \cap \mathcal{O}_j} \|t - x\|_2^2 + \sum_{T \in \mathbf{T} \setminus \mathcal{O}_j} (\|t\|_2 + \|x\|_2)^2 \quad \forall X \in \mathcal{O}_j.$$

Hence, restricted to a fixed orthant, the objective only depends on the embedding x , so we may calculate directional derivatives as in Euclidean space.

Lemma 4.8. *Consider an orthant \mathcal{O}_j . Let $V \in \mathcal{O}_j$ be s.t. $\|v\|_2 = 1$, where $v = \psi_{\mathcal{O}_j}(V) \in \mathbb{R}_+^{n-2}$. Then*

$$\begin{aligned} \nabla_V d(\cdot, T_i)(0) &= 1 && \text{if } T_i \notin \mathcal{O}_j, \\ \nabla_V d(\cdot, T_i)(0) &\geq -1 && \text{otherwise.} \end{aligned}$$

Proof. For any $V \in \mathcal{O}_j$ and $\lambda > 0$ we define $\lambda \cdot V := \psi_{\mathcal{O}_j}^{-1}(\lambda \cdot \psi_{\mathcal{O}_j}(V))$, basically scaling the weight vector v of V in \mathbb{R}_+^{n-2} by λ .

We start with the case where $T_i \notin \mathcal{O}_j$. Then, applying the distance formula from Lemma 4.6 and the definition of directional derivatives we have

$$\begin{aligned} \nabla_V d(\cdot, T_i)(0) &= \lim_{h \rightarrow 0} \left(\frac{d(h \cdot V, T_i) - d(0, T_i)}{h} \right) \\ &= \lim_{h \rightarrow 0} \left(\frac{\|h \cdot v\|_2 + \|t_i\|_2 - \|t_i\|_2}{h} \right) \\ &= \lim_{h \rightarrow 0} \left(\frac{|h| \|v\|_2}{h} \right) \\ &= \lim_{h \rightarrow 0} \left(\frac{h}{h} \right) \\ &= 1. \end{aligned}$$

Now, let $T_i \in \mathcal{O}_j$. Then

$$\begin{aligned}
\nabla_V d(\cdot, T_i)(0) &= \lim_{h \rightarrow 0} \left(\frac{d(h \cdot V, T_i) - d(0, T_i)}{h} \right) \\
&= \lim_{h \rightarrow 0} \left(\frac{\|h \cdot v - t_i\|_2 - \|t_i\|_2}{h} \right) \\
&\geq \lim_{h \rightarrow 0} \left(\frac{\|h \cdot \frac{t_i}{\|t_i\|_2} - t_i\|_2 - \|t_i\|_2}{h} \right) \\
&= \lim_{h \rightarrow 0} \left(\frac{\|(1 - \frac{h}{\|t_i\|_2})t_i\|_2 - \|t_i\|_2}{h} \right) \\
&= \lim_{h \rightarrow 0} \left(\frac{|(1 - \frac{h}{\|t_i\|_2})\|t_i\|_2 - \|t_i\|_2}{h} \right) \\
&= \lim_{h \rightarrow 0} \left(\frac{\|t_i\|_2 - h - \|t_i\|_2}{h} \right) \\
&= \lim_{h \rightarrow 0} \left(\frac{-h}{h} \right) \\
&= -1.
\end{aligned}$$

where the “ \geq ” in the second line holds because $v^* = \frac{t_i}{\|t_i\|_2}$ minimizes $\|h \cdot v - t_i\|_2$ amongst all v with $\|v\|_2 = 1$, when h is sufficiently small such that $\frac{h}{\|t_i\|_2} \leq 1$. \square

With this we can calculate the directional derivative of $f = f_{\mathbf{F}}$. To simplify notation, define $\mathbf{T}_j = \mathbf{T} \cap \mathcal{O}_j$ for $j = 1, \dots, \kappa$.

Lemma 4.9. *Consider an orthant \mathcal{O}_j for some $j = 1, \dots, \kappa$ and let $V \in \mathcal{O}_j$ be s.t. $\|v\|_2 = 1$. Then*

$$\begin{aligned}
\nabla_V f(0) &= 2 \cdot \sum_{i=1}^M \nabla_V d(\cdot, T_i)(0) \cdot d(0, T_i) \\
&\geq 2 \cdot \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} \|t_i\|_2 - 2 \cdot \sum_{T_i \in \mathbf{T}_j} \|t_i\|_2
\end{aligned}$$

Proof. This result is an immediate consequence of Lemma 4.8 and the chain rule. \square

With these two statements we are able to prove a reduction result that reduces the problem to one orthant for the Fréchet problem $(P_{\mathbf{F}})$. Recall that $(P_{\mathbf{F}})$ always has a unique solution, as established in Theorem 3.5.

Theorem 4.10. *If there exists $k \in \{1, \dots, \kappa\}$ such that*

$$\sum_{T_i \in \mathbf{T}_k} \|t_i\|_2 \geq \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_k} \|t_i\|_2,$$

then the optimal solution to (P_F) is contained in \mathcal{O}_k . If such an orthant does not exist, 0 is the unique optimal solution to (P_F) .

Proof. We first show

$$\sum_{T \in \mathbf{T}_j} \|t\|_2 \leq \sum_{T \in \mathbf{T} \setminus \mathbf{T}_j} \|t\|_2 \quad \Rightarrow \quad \bar{A} \text{ optimal solution in } \mathcal{O}_j \setminus \{0\} \quad (4.3)$$

Using the explicit form of f on \mathcal{O}_j we receive

$$f(X) = \sum_{T_i \in \mathbf{T}_j} \|t_i - x\|_2^2 + \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} (\|t_i\|_2 + \|x\|_2)^2.$$

Note, that on \mathcal{O}_j , f only depends on the embedded vector $\psi(X) = x \in \mathbb{R}_+^{n-2}$, and that the right hand side is convex in x , as it is a sum of convex functions. Now suppose there exists some $X^* \in \mathcal{O}_j$, $X^* \neq 0$, such that X^* is a minimizer of f on \mathcal{O}_j . Let $V = \psi_{\mathcal{O}_j}^{-1}\left(\frac{x^*}{\|x^*\|_2}\right)$. By construction $\|\psi_{\mathcal{O}_j}(V)\|_2 = 1$. Since

$$\sum_{T \in \mathbf{T}_j} \|t\|_2 \leq \sum_{T \in \mathbf{T} \setminus \mathbf{T}_j} \|t\|_2,$$

Lemma 4.9 implies that $\nabla_V f(0) \geq 0$, i.e., V is a direction of ascent at 0. Thus, there exists some $\varepsilon > 0$ such that $f(\delta V) > f(0)$ for all $0 < \delta \leq \varepsilon$, where $\delta V := \psi_{\mathcal{O}_j}^{-1}(\delta \psi(V))$. In particular we have that $f(\varepsilon V) > f(0)$ and it holds that $f(\varepsilon V) \geq f(X^*)$, since X^* is the supposed minimizer. Hence

$$f(\varepsilon \cdot V) > \mu f(X^*) + (1 - \mu)f(0) \quad \text{for all } \mu \in (0, 1). \quad (4.4)$$

Now, define $\lambda = \frac{\varepsilon}{\|x^*\|_2} > 0$, such that $\varepsilon v = \frac{\varepsilon}{\|x^*\|_2} x^* = \lambda x^*$. If $\lambda \geq 1$, we have $\varepsilon \geq \|x^*\|_2$ which yields

$$f(X^*) = f\left(\frac{\|x^*\|_2}{\|x^*\|_2} X^*\right) = f(\|x^*\|_2 \cdot V) > f(0)$$

as we have just established that δV is a direction of ascent at 0 for $\delta \leq \varepsilon$. This contradicts optimality of X^* when $\lambda \geq 1$.

Let now $\lambda \in (0, 1)$. Then we apply convexity of f for $\varepsilon v = \lambda x^* + (1 - \lambda) \cdot 0$:

$$\begin{aligned}
f(\varepsilon \cdot V) &= \sum_{T_i \in \mathbf{T}_j} \|t_i - \varepsilon \cdot v\|_2^2 + \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} (\|t_i\|_2 + \|\varepsilon v\|_2)^2 \\
&= \sum_{T_i \in \mathbf{T}_j} \|t_i - (\lambda x^* + (1 - \lambda) \cdot 0)\|_2^2 + \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} (\|t_i\|_2 + \|(\lambda x^* + (1 - \lambda) \cdot 0)\|_2)^2 \\
&\leq \lambda \left(\sum_{T_i \in \mathbf{T}_j} \|t_i - x^*\|_2^2 + \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} (\|t_i\|_2 + \|x^*\|_2)^2 \right) \\
&\quad + (1 - \lambda) \left(\sum_{T_i \in \mathbf{T}_j} \|t_i - 0\|_2^2 + \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} (\|t_i\|_2 + \|0\|_2)^2 \right) \\
&= \lambda f(X^*) + (1 - \lambda) f(0).
\end{aligned}$$

This contradicts (4.4). Hence, there cannot exist a minimizer $0 \neq X^*$ in \mathcal{O}_j .

With this we have proven (4.3) and we now prove the claim of the theorem that either the unique optimal solution is 0 or that some orthant satisfies the optimality condition. First note that, in analogy to the proof of Theorem 4.7 it holds that any optimal solution X^* to (P_F) has to satisfy $\text{Split}(X^*) \subset \text{Split}(\mathcal{O}_1) \cup \dots \cup \text{Split}(\mathcal{O}_\kappa)$ which again implies $\text{Split}(X^*) \subset \text{Split}(\mathcal{O}_l)$ for some $l \in \{1, \dots, \kappa\}$ due to the complete incompatibility of the orthants.

Now, suppose that for $k \in \{1, \dots, \kappa\}$ it holds that $\sum_{T_i \in \mathbf{T}_k} \|t_i\|_2 \geq \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} \|t_i\|_2$. Then we get

$$\sum_{T_i \in \mathbf{T}_j} \|t_i\|_2 \leq \sum_{T_i \in \mathbf{T}_k} \|t_i\|_2 \leq \sum_{T_i \in \mathbf{T} \setminus \mathbf{T}_j} \|t_i\|_2$$

for all $j \neq k$. (4.3) then implies that no optimal solution can be contained in $\mathcal{O}_j \setminus \{0\}$ with $j \neq k$, so all optimal solutions are contained in \mathcal{O}_k .

The case that such an orthant \mathcal{O}_k does not exist, directly implies that for all j it holds that there cannot be an optimal solution in $\mathcal{O}_j \setminus \{0\}$ for all j , so 0 is the unique optimal solution to the problem. \square

For all three problems we have now established conditions when 0 is the optimal solution and in the case that 0 is not the optimal solution we know in which orthant to search for the optimal solution. Let us assume w.l.o.g. that the orthant containing the optimal solution is \mathcal{O}_1 . With this knowledge we can treat all other trees in the orthants $\mathcal{O}_2, \dots, \mathcal{O}_\kappa$ in the same way, since the geodesic distance does not depend on their split sets, but only the norm of their weight vectors, as

$$d(X, T) = \|x\|_2 + \|t\|_2 \quad \text{for } X \in \mathcal{O}_1, T \in \mathcal{O}_j, j = 2, \dots, \kappa.$$

Therefore it makes no difference when we replace the given trees $\mathbf{T}_i \subset \mathcal{O}_i$, $i \geq 3$, by trees in \mathcal{O}_2 which have a weight vector with the same norm. To this end, for $T_k = ((s_1^k, \dots, s_{n_k}^k), (w_1^k, \dots, w_{n_k}^k)) \in \mathbf{T} \setminus \mathbf{T}_1$, define

$$\tau_2 : \mathbf{T} \setminus \mathbf{T}_1 \rightarrow \mathcal{O}_2 : \tau_2(T_k) = ((s'_1, \dots, s'_{n_k}), (w_1^k, \dots, w_{n_k}^k)),$$

where the s'_i are the splits of the maximal orthant \mathcal{O}_2 , i.e., $\text{Split}(\mathcal{O}_2) = \{s'_1, \dots, s'_{n-2}\}$. Since the weight vectors are identical we have $\|\chi(T_k)\|_2 = \|\chi(\tau_2(T_k))\|_2$, and since $s'_i \in \text{Split}(\mathcal{O}_2)$ it follows that $\tau_2(T_k) \in \mathcal{O}_2$.

With this map at hand we can state equivalent problem formulations for all problems (P_{Med}) , (P_{max}) , (P_{F}) where we only have facilities in two orthants: We keep the facilities \mathbf{T}_1 and replace the facilities T_i in all other orthants by $\tau_2(T_i) \in \mathcal{O}_2$, so that we end up with data in only two orthants.

Theorem 4.11. *For all three problems assume that \mathcal{O}_1 is the orthant containing all optimal solutions to the problem. Define $\mathbf{T}' = \mathbf{T}_1 \cup \tau_2(\mathbf{T} \setminus \mathbf{T}_1)$. Then*

i)

$$(P_{\text{Med}})(\mathbf{T}) = \min_{\substack{\sum_{T \in \mathbf{T}} d(X, T) \\ \text{s.t. } X \in \mathcal{O}_1}} \Leftrightarrow (P_{\text{Med}})(\mathbf{T}') = \min_{\substack{\sum_{T \in \mathbf{T}'} d(X, T) \\ \text{s.t. } X \in \mathcal{O}_1}}$$

ii)

$$(P_{\text{max}})(\mathbf{T}) = \min_{\substack{\max_{T \in \mathbf{T}} d(X, T) \\ \text{s.t. } X \in \mathcal{O}_1}} \Leftrightarrow (P_{\text{max}})(\mathbf{T}') = \min_{\substack{\max_{T \in \mathbf{T}'} d(X, T) \\ \text{s.t. } X \in \mathcal{O}_1}}$$

iii)

$$(P_{\text{F}})(\mathbf{T}) = \min_{\substack{\sum_{T \in \mathbf{T}} d(X, T)^2 \\ \text{s.t. } X \in \mathcal{O}_1}} \Leftrightarrow (P_{\text{F}})(\mathbf{T}') = \min_{\substack{\sum_{T \in \mathbf{T}'} d(X, T)^2 \\ \text{s.t. } X \in \mathcal{O}_1}}$$

Proof. We only need to show that the objectives of the problems coincide in all three cases. This, however, is an immediate consequence of the distances being the same for all facilities or the transformed facilities, respectively: \mathbf{T}_1 is contained in \mathbf{T} as well as in \mathbf{T}' . For $T \in \mathbf{T} \setminus \mathbf{T}_1$ we have

$$d(X, T) = \|x\|_2 + \|t\|_2 = \|x\|_2 + \|\chi(T)\|_2 = \|x\|_2 + \|\chi(\tau_2(T))\|_2 = d(X, \tau_2(T))$$

which is the corresponding transformed tree in \mathbf{T}' . Thus, the objectives of all problems are identical for each $X \in \mathcal{O}_1$. \square

Now having the formulation of the equivalent problems where data is only contained in two orthants it is possible to formulate equivalent problems in \mathbb{R}^{n-2} . This is not necessarily possible for data in $\kappa \geq 3$ orthants: As an example, in \mathcal{T}_4 there exist exactly five triplets of incompatible orthants. Assume that the data is contained in one of these triplets. Without the reduction to two orthants, there is no way to isometrically embed the three orthants into \mathbb{R}^2 . The image of every orthant of \mathcal{T}_4 is a closed orthant in \mathbb{R}^2 , thus, one pair of the orthants has to intersect in a line segment of an axis. The problem is, that the trees of the different orthants that would be mapped to this axis would not be the same, so this map can not be isometric.

This problem is solved after applying Theorem 4.11. When we only have data of two orthants in \mathcal{T}_n , we can map one orthant to \mathbb{R}_+^{n-2} and the other one to \mathbb{R}_-^{n-2} , so that the intersection of the orthants in tree space, as well as in Euclidean space after the transformation is only the origin.

Just as for the adjacent orthant case in Section 4.2, we use the specific form of the geodesic distance to define an isometry to \mathbb{R}^{n-2} , where \mathbb{R}^{n-2} is equipped with a special distance measure. We define the map $\varphi : \mathcal{O}_1 \cup \mathcal{O}_2 \rightarrow \mathbb{R}^{n-2}$ by

$$\varphi(Y) = \psi_{\mathcal{O}_1}(Y) \in \mathbb{R}_+^{n-2} \text{ for } Y \in \mathcal{O}_1, \quad \varphi(Y) = -\psi_{\mathcal{O}_2}(Y) \in \mathbb{R}_-^{n-2} \text{ for } Y \in \mathcal{O}_2. \quad (4.5)$$

Our goal is again to find an isometry to (\mathbb{R}^{n-2}, d^0) so that $d(X, T) = d^0(\varphi(X), \varphi(T))$ holds for all $X, T \in \mathcal{O}_1 \cup \mathcal{O}_2$. To this end, define d^0 on $\text{im}(\varphi) = I_1 \cup I_2 \subset \mathbb{R}^{n-2}$, where $I_1 := \mathbb{R}_+^{n-2}$ and $I_2 := \mathbb{R}_-^{n-2}$ as $d^0 : \text{im}(\varphi) \times \text{im}(\varphi) \rightarrow [0, \infty)$ with

$$d^0(y_1, y_2) = \begin{cases} \|y_1 - y_2\|_2 & \text{if } \exists j \in \{1, 2\} : \{y_1, y_2\} \subset I_j, \\ \|y_1\|_2 + \|y_2\|_2 & \text{otherwise.} \end{cases}$$

With Lemma 4.6 it follows by definition of d^0 that $\varphi : (\mathcal{O}_1 \cup \mathcal{O}_2, d) \rightarrow (\text{im}(\varphi), d^0)$ is an isometry. With this transformation at hand, we define the Euclidean facilities A_i via

$$A_i := \varphi(T_i) \quad i = 1, \dots, M \quad \text{and set } \mathcal{A} = \varphi(\mathbf{T}) \text{ and } \mathcal{A}_i = \varphi(\mathbf{T}_j) \text{ for } j = 1, 2. \quad (4.6)$$

Now we can state the equivalent problems in the Euclidean space, just as in Theorem 4.4.

Theorem 4.12. *Let $\mathbf{T} \subset \mathcal{O}_1 \cup \mathcal{O}_2$ and f be one of the three objective functions $f_{\text{Med}}, f_{\text{max}}, f_{\text{F}}$. Then*

$$(P) \quad \min_{X \in \mathcal{T}_n} f(X) \quad \Leftrightarrow \quad (\tilde{P}) \quad \min_{x \in \text{im}(\varphi)} \tilde{f}(x)$$

where \tilde{f} depends on f and $\mathcal{A} = \varphi(\mathbf{T})$ as follows:

$$\begin{aligned} f = f_{\text{Med}} &\rightarrow \tilde{f}(x) = \sum_{i=1, \dots, M} d^0(x, A_i) \\ f = f_{\text{max}} &\rightarrow \tilde{f}(x) = \max_{i=1, \dots, M} d^0(x, A_i) \\ f = f_F &\rightarrow \tilde{f}(x) = \sum_{i=1, \dots, M} d^0(x, A_i)^2 \end{aligned}$$

Proof. The only thing we need to show is that we may restrict \mathcal{T}_n to \mathcal{O}_1 and \mathcal{O}_2 , because then the problems are exactly the same, just that we transformed them by a bijective isometry.

Using Theorem 3.4 and Theorem 3.5, we get that for each of the three objective functions it holds that every optimal solution $X^* \in \mathcal{T}_n$ has to satisfy $\text{Split}(X^*) \subset \text{Split}(T_1) \cup \dots \cup \text{Split}(T_M) \subseteq S_1 \cup S_2$. Since every split in S_1 is incompatible with every split in S_2 this implies that either $X^* \in \mathcal{O}_1$ or $X^* \in \mathcal{O}_2$, which was to show. \square

Note that, in contrast to Theorem 4.4 we have to restrict the variables to $\text{im}(\varphi)$, because it is otherwise possible that the optimal solution is contained in $\mathbb{R}^{n-2} \setminus \text{im}(\varphi)$, since $\text{im}(\varphi)$ is not convex. Moreover, we do not have the Euclidean distance in the resulting problems in \mathbb{R}^{n-2} , it actually is the intrinsic distance of $\|\cdot\|_2$ restricted to the subset $\text{im}(\varphi)$.

On the one hand, this means that we did not really gain many additional insights by this reformulation, because the metric d^0 is not the standard Euclidean norm and we cannot apply standard algorithms from Facility Location as in the adjacent orthant case in Section 4.2.

On the other hand, the reformulated versions belong to a special class of Location Theory problems which are problems with *barrier distances* or *forbidden regions*, see [KC81] or [AP94] for early works on this subject. Since the problems may be interpreted as special cases of these problems, and some familiar problems are studied in the field of Location Theory, we also formulated and solved the problems in the Euclidean setting.

The Euclidean problems with the distance d^0 that arise from the isometry φ (4.5) are what we call *fixed gate point* location problems. The origin 0, may be interpreted as the single point in the orthants where one is allowed to traverse from one orthant to another, hence the name ‘gate’. In the following we first describe the origin of gate point problems in Location Theory and related research in Location Theory before we develop solution methods for each objective function. The results and the solution methods for the fixed gate point problems are submitted for publication and a preprint [Bot18] is available.

4.4 Location Problems with a Fixed Gate Point at the Origin

Fixed gate point problems are a specific kind of location problems that we introduce in this paper. [BKW05] coined the term *gate point*. *Gate points* introduce a deviation to classical location problems by introducing new distance measures. These distance measures arise when a space is partitioned into several regions that are equipped with different distance measures or when *forbidden regions* are introduced that connections between facilities and the placed location may not cross.

The gate point variant we present here was introduced in [BKW03] and [FVGA12], as it is the closest variant to our specific *fixed gate point* version.

Given a line R in \mathbb{R}^2 , let the two half-spaces H^+, H^- induced by R carry different ℓ_p norms, ℓ_{p_1}, ℓ_{p_2} , respectively. Then the distance between two points in the same half-space is the respective ℓ_p distance, and the distance between point $x \in H^+, y \in H^-$ is given by

$$\min_{z \in R} \{d_{p_1}(x, z) + d_{p_2}(z, y)\}, \quad (4.7)$$

and the minimizer $z^* \in R$ of (4.7), where one traverses from one half space into the other, is called the *gate point* for x and y .

An early work of the problem of calculating different distances in subdivisions of the plane is [MP91], where they calculate distances in a bounded part of the plane that is subdivided into bounded polygons, each equipped with the Euclidean distance with a different positive scaling factor. For the specific distance (4.7) induced by the two hyperplanes, H^+, H^- , [BKW03] investigate general properties and several cases of the problem and develop solution procedures for the investigated cases. [FVGA12] solved the optimization problem analytically for $p_2 = \infty$ and $p_1 > 1$ under slight assumptions and have given several theoretical statements in which segments of R the gate point has to be contained.

More general variants of the distance (4.7) where the regions are not restricted to half-spaces have been studied before, see e.g. [BKW05], [Par94], where the Weber problem is investigated for this distance.

A possible application for gate point problems is presented in [BKW03] and [Par94]: Traveling within a city is only possible using the ℓ_1 distance due to the street layout, whereas distances outside the city are considered to be Euclidean, for example when traveling with a helicopter, or via straight highways.

Nonetheless, instead of considering regions where one is allowed to traverse anywhere on their shared boundary, one could also think of applications, where there is only a prespecified set of what we call *fixed gate points* $G \subset R$, where one is allowed to traverse from one region into the other. For example, these gate points may model bridges over a river or bridges and tunnels for pedestrians that want to cross highways.

Of course, our motivation originates from the location problems in tree space, more specifically from the case where data is contained in completely incompatible orthants. In this case we only have a single fixed gate point at the origin. Recall that the given facilities are denoted by $\mathcal{A} = \{A_1, \dots, A_M\}$, compare (4.6), in contrast to the data for location problems in tree space for which we used $\mathbf{T} = \{T_1, \dots, T_M\}$. In Theorem 4.11 we reformulate general problems to equivalent problems where the data is only contained in two orthants, so we focus on solving these problems in Euclidean space and use $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$, with $\mathcal{A}_1, \mathcal{A}_2$ defined as in (4.6). We denote the fixed gate point problems, as stated in Theorem 4.12 by

$$\begin{aligned} (G_{\text{Med}}) \quad & \min \sum_{i=1}^M d^0(x, A_i) \quad \text{s.t.} \quad x \in \mathbb{R}_+^{n-2} \cup \mathbb{R}_-^{n-2}, \\ (G_{\text{max}}) \quad & \min \max_{i=1, \dots, M} d^0(x, A_i) \quad \text{s.t.} \quad x \in \mathbb{R}_+^{n-2} \cup \mathbb{R}_-^{n-2}, \\ (G_{\text{F}}) \quad & \min \sum_{i=1}^M d^0(x, A_i)^2 \quad \text{s.t.} \quad x \in \mathbb{R}_+^{n-2} \cup \mathbb{R}_-^{n-2}. \end{aligned}$$

and study them in the following sections. As we only focus on solving these problems in the Euclidean setting, we simplify notation and write n instead of $n-2$, d instead of d^0 for the remainder of Chapter 4. Then,

$$\begin{aligned} (G_{\text{Med}}) \quad & \min \sum_{i=1}^M d(x, A_i) \quad \text{s.t.} \quad x \in \mathbb{R}_+^n \cup \mathbb{R}_-^n, \\ (G_{\text{max}}) \quad & \min \max_{i=1, \dots, M} d(x, A_i) \quad \text{s.t.} \quad x \in \mathbb{R}_+^n \cup \mathbb{R}_-^n, \\ (G_{\text{F}}) \quad & \min \sum_{i=1}^M d(x, A_i)^2 \quad \text{s.t.} \quad x \in \mathbb{R}_+^n \cup \mathbb{R}_-^n. \end{aligned}$$

are the formulations with which we work in the following sections.

4.4.1 Median Problem

At first, we consider the median problem (G_{Med}). Denote $k_i = |\mathcal{A}_i|$ for $i = 1, 2$. We first characterize the set of optimal solutions in the case $k_1 = k_2$, and then reformulate the problem for the case $k_1 \neq k_2$. Recall that $\mathcal{A}_1 \subset \mathbb{R}_+^n$ and $\mathcal{A}_2 \subset \mathbb{R}_-^n$.

Lemma 4.13. *Suppose $k_1 = k_2$. Then 0 is an optimal solution. Moreover, if $\mathcal{A}_i \subset \{\lambda v_i | \lambda \geq 0\}$ for some $v_i \in \mathbb{R}^n$ with $\|v_i\|_2 = 1$, then all elements of $L_i =$*

$\{\lambda v \mid \lambda \in [0, c]\}$ are optimal solutions, where $c = \min_{A_j \in \mathcal{A}_i} \|A_j\|_2$ and no other optimal solution exists.

Proof. Define $k := k_1 = k_2$. Then for any $x \in \mathbb{R}_+^n$

$$\begin{aligned} f(0) &= \sum_{A \in \mathcal{A}} \|A\|_2 = \sum_{A \in \mathcal{A}_1} (\|A\|_2 - \|x\|_2) + k\|x\|_2 + \sum_{A \in \mathcal{A}_2} \|A\|_2 \\ &\stackrel{(*)}{\leq} \sum_{A \in \mathcal{A}_1} \|x - A\|_2 + \sum_{A \in \mathcal{A}_2} (\|x\|_2 + \|A\|_2) = f(x). \end{aligned}$$

The same estimate may be done for $x \in \mathbb{R}_-^n$ by simply exchanging roles of \mathcal{A}_1 and \mathcal{A}_2 . It follows that $f(0) \leq f(x)$ for all feasible x , so 0 is an optimal solution. Even more, we see from (*) that 0 is the unique optimal solution if and only if (*) is strict for $i = 1, 2$. In the case of equality we have that x satisfies

$$\begin{aligned} \text{"=" holds in (*)} &\Leftrightarrow x = \lambda_A \cdot A \text{ with } 0 \leq \lambda_A \leq 1 \quad \forall A \in \mathcal{A}_i \\ &\Leftrightarrow x \in L_i \text{ and } \mathcal{A}_i \subset \{\lambda v_i \mid \lambda \geq 0\} \end{aligned}$$

with L_i and v_i as in the statement of the lemma. □

The remaining case is $k_1 \neq k_2$ and w.l.o.g. assume $k_1 > k_2$.

Theorem 4.14. *If $k_1 > k_2$, (G_{Med}) is equivalent to*

$$\begin{aligned} \min \quad & \sum_{A \in \mathcal{A}_1} \|x - A\|_2 + k_2 \cdot \|x\|_2 \\ \text{s.t.} \quad & x \in \mathbb{R}_+^n \end{aligned}$$

Proof. This follows from Theorem 4.7: The optimal solution has to be in \mathcal{O}_1 , which translates to \mathbb{R}_+^n after the reformulation. Then plugging in the explicit distances for $x \in \mathbb{R}_+^n$ into the objective we get

$$\sum_{A \in \mathcal{A}_1} \|x - A\|_2 + \sum_{A \in \mathcal{A}_2} (\|A\|_2 + \|x\|_2) = \sum_{A \in \mathcal{A}_1} \|x - A\|_2 + k_2 \cdot \|x\|_2 + \sum_{A \in \mathcal{A}_2} \|A\|_2$$

and we simply omit the constant part to get the objective

$$\sum_{A \in \mathcal{A}_1} \|x - A\|_2 + k_2 \cdot \|x\|_2.$$

□

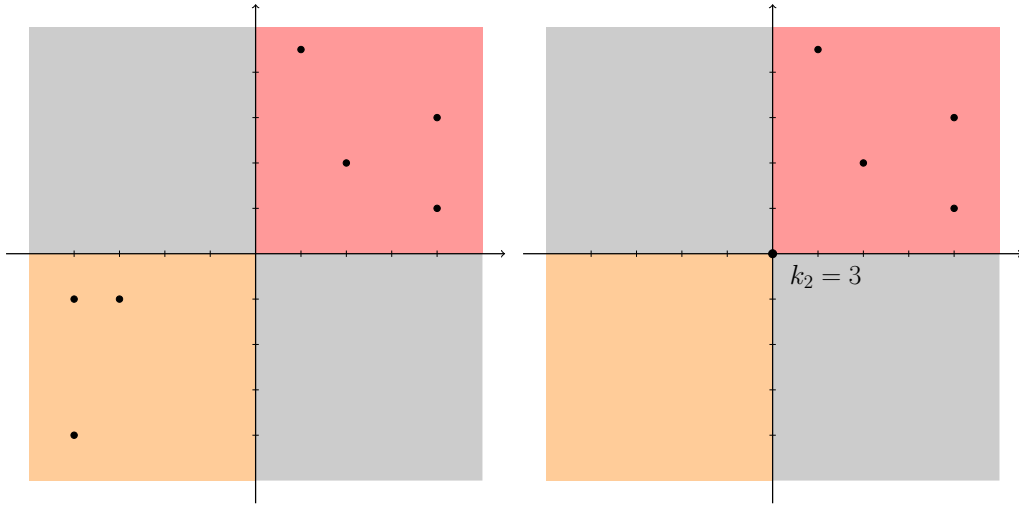


Figure 4.1: Example for the transformation of the fixed gate point median problem to a classical median problem.

This is now a standard Weber problem with a facility at 0 with weight of k_2 and it can hence be solved with, e.g., the Weiszfeld algorithm. Note that in cases where k_2 is rather big (but still, k_2, k_1) it might happen that 0 is the unique optimal solution to the reformulated standard median problem and the Weiszfeld algorithm might not detect it. This is no problem, since Theorem 3.1 (ii) gives an easily verifiable optimality criterion. 0 is the optimal solution to the problem if and only if

$$\left\| \sum_{i=1}^{k_1} \frac{-A_i}{\|A_i\|_2} \right\|_2 \leq k_2.$$

With the reformulation to a standard median problem for the case $k_2 < k_1$ and the explicit description of the set of optimal solutions in the case $k_1 = k_2$ we are able to solve the problem (G_{Med}) in all cases and proceed with the next problem.

4.4.2 Center Problem

In this subsection we turn our attention towards the center problem (G_{max}):

$$\begin{aligned} \min \quad & \max_{i=1, \dots, M} d(x, A_i) \\ \text{s.t.} \quad & x \in \mathbb{R}_+^n \cup \mathbb{R}_-^n \end{aligned}$$

Theorem 4.7 states that 0 is the unique optimal solution in the case that two orthants contain a facility with maximal norm. As this case is solved we suppose that there

exists only one such orthant and assume w.l.o.g. that the facility with maximal norm is contained in $\mathcal{A}_1 \subset \mathbb{R}_+^n$. Then we can reformulate the problem to

$$(G_{\max})' \quad \min \quad \max \{ \max_{A \in \mathcal{A}_1} (\|x - A\|_2), \max_{A \in \mathcal{A}_2} (\|x\|_2 + \|A\|_2) \}$$

$$\text{s.t. } x \in \mathbb{R}_+^n,$$

as Theorem 4.7 states that the optimal solution is contained in \mathbb{R}_+^n in this case. To simplify notation, define $r_2 := \max_{A \in \mathcal{A} \setminus \mathcal{O}} \|A\|_2$ and assume w.l.o.g, that $\mathcal{A}_1 = \{A_1, \dots, A_k\}$, $\mathcal{A}_2 = \{A_{k+1}, \dots, A_M\}$.

A commonly used notion in the context of center problems is the *smallest enclosing ball* for a set of points $B_1, \dots, B_m \subset \mathbb{R}^d$:

Definition 4.15. *The smallest enclosing ball of $B_1, \dots, B_m \subset \mathbb{R}^d$ is the ball $B_r(x) = \{y \in \mathbb{R}^d : \|x - y\|_2 \leq r\}$ s.t. $B_i \in B_r(x)$ for $i = 1, \dots, m$ and there exists no other ball $B_{r'}(x')$ such that $r' \leq r$.*

In the definition it is implicitly stated that the ball is unique. That can easily be seen by assuming there were two such balls with center points $x_1 \neq x_2$ and then the ball centered at $\frac{x_1 + x_2}{2}$ has a smaller radius. We will use the term smallest enclosing ball for fixed gate point center problems as well, where the only difference is that the balls are with respect to the fixed gate point distance d .

Theorem 4.16. *Consider the problem*

$$(P_{r_2}) \quad \min \quad r$$

$$\text{s.t. } \begin{aligned} \|A_i - x\|_2 &\leq r & i = 1, \dots, k \\ \|x\|_2 + r_2 &\leq r \\ x &\in \mathbb{R}_+^n \\ r &\in \mathbb{R}. \end{aligned}$$

x^* is an optimal solution to (G_{\max}) if and only if (x^*, r^*) is an optimal solution to (P_{r_2}) with

$$r^* = \max \left\{ \max_{i=1, \dots, k} (\|A_i - x\|_2), \|x\|_2 + r_2 \right\}.$$

Proof. The problem (P_{r_2}) is obtained from (G_{\max}) by introducing the *bottleneck* variable r . Note that (x, r) is feasible if and only if

$$r \geq \max \left\{ \max_{i=1, \dots, k} \|A_i - x\|_2, r_2 + \|x\|_2 \right\}.$$

Hence, for any $x \in \mathbb{R}_+^n$ the lowest attainable objective value is

$$\begin{aligned} & \max \left\{ \max_{i=1, \dots, k} \|A_i - x\|_2, r_2 + \|x\|_2 \right\} \\ &= \max \left\{ \max_{A \in \mathcal{A}_1} (\|x - A\|_2), \max_{A \in \mathcal{A}_2} (\|x\|_2 + \|A\|_2) \right\} \\ &= \max_{A \in \mathcal{A}} d(x, A) \end{aligned}$$

which is the objective value of x for (G_{\max}) . This concludes the proof as the feasible sets for x for the two problems coincide. \square

For the remainder of the section, we choose a fixed $A_0 \in \arg \max_{A \in \mathcal{A}_2} \|A\|_2$. A_0 models an *active* facility in \mathcal{A}_2 .

The goal is to find structural parallels of (P_{r_2}) to a standard center problem. We start with the helpful and intuitive concept of *active* facilities.

Definition 4.17. *Let x^* be the optimal solution to (P_{r_2}) with optimal objective value r^* . Then a facility A_i , $i = 1, \dots, k$ that satisfies $\|A_i - x^*\|_2 = r^*$ is called active. A_0 is called active if $\|A_0\|_2 + \|x^*\|_2 = r^*$.*

The next theorem proves that the optimal solution to our problem is unique, thereby justifying that we only spoke of active facilities in general, not with respect to a specific optimal solution.

Theorem 4.18. *The problem (P_{r_2}) has a unique optimal solution (x^*, r^*) .*

Proof. We start by showing existence. For any $x \in \mathbb{R}_+^n$ we may choose $r > 0$ such that (x, r) is feasible, as \mathcal{A}_1 is bounded and r_2 is finite. The objective is continuous and bounded from below so it follows that an optimal solution exists.

Now, we show uniqueness of the optimal solution. To this end, suppose there were two optimal solutions $x_1^* \neq x_2^*$ with objective value r^* . We show that $x^* := \frac{1}{2}(x_1^* + x_2^*) \in \mathbb{R}_+^n$ yields a lower objective value. For $i = 1, \dots, n$ it holds that

$$\begin{aligned} \|x^* - A_i\|_2 &= \left\| \frac{1}{2}(x_1^* - A_i) + \frac{1}{2}(x_2^* - A_i) \right\|_2 \\ &\leq \frac{1}{2} (\|x_1^* - A_i\|_2 + \|x_2^* - A_i\|_2) \leq r^* \end{aligned} \quad (4.8)$$

and

$$\|x^*\|_2 + \|A_0\|_2 \leq \frac{1}{2} (\|x_1^*\|_2 + \|A_0\|_2) + \frac{1}{2} (\|x_2^*\|_2 + \|A_0\|_2) \leq r^* \quad (4.9)$$

where feasibility of (x_j^*, r^*) , $j = 1, 2$ is used in the last inequalities. This shows that (x^*, r^*) is a feasible solution. Now investigate the above terms in more detail:

- Suppose A_i is active for x_1^* and x_2^* , i.e., $\|x_1^* - A_i\|_2 = r^* = \|x_2^* - A_i\|_2$.

$$\|x^* - A_i\|_2 = \left\| \frac{1}{2}(x_1^* - A_i) + \frac{1}{2}(x_2^* - A_i) \right\|_2 \leq \frac{1}{2} (\|x_1^* - A_i\|_2 + \|x_2^* - A_i\|_2) \quad (4.10)$$

holds with equality $\Leftrightarrow (x_1^* - A_i) = \lambda(x_2^* - A_i)$ for $\lambda \geq 0$. If equality in (4.10) would hold, then $\lambda = 1$ follows, as their distance to A_i is equal. But then $x_1^* = x_2^*$ yields a contradiction, so (4.10) is a strict inequality.

- If A_i is not active for one of the x_j^* , then using $\|x_j^* - A_i\|_2 < r^*$ in (4.8) implies $\|x^* - A_i\|_2 < r^*$.
- If A_0 is active for x_1^* and x_2^* , then it follows that $\|x_1^*\|_2 = \|x_2^*\|_2$, which implies $\|x^*\|_2 < \|x_1^*\|_2 = \|x_2^*\|_2$ by strict convexity of the norm ball. Then $\|x^*\|_2 + \|A_0\|_2 < \|x_1^*\|_2 + \|A_0\|_2 = r^*$.
- If A_0 is not active for at least one of the x_j^* , then using $\|x_j^*\|_2 + \|A_0\|_2 < r^*$ in (4.9) implies $\|x^*\|_2 + \|A_0\|_2 < r^*$.

Altogether these cases prove $\|x^* - A_i\|_2 < r^*$ for $i = 1, \dots, k$ and $\|x^*\|_2 + \|A_0\|_2 < r^*$. Thus, $r' := \max \{ \max_{i=1, \dots, k} \{ \|x^* - A_i\|_2 \}, \|x^*\|_2 + \|A_0\|_2 \} < r^*$. Then (x^*, r') is feasible and $r' < r^*$, which contradicts optimality of the x_j^* . □

The activity of the A_0 plays an important role for the problem (P_{r_2}) . Suppose A_0 was not active for an instance $\{A_0, A_1, \dots, A_k\}$. Then (P_{r_2}) reduces to

$$(S_{r_2}) \quad \begin{array}{ll} \min & r \\ \text{s.t.} & \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k \\ & x \in \mathbb{R}_+^n \\ & r \in \mathbb{R}. \end{array}$$

This is the easier case, as we receive a standard center problem which can be solved via different methods, compare [Meg83],[FGK03], [KMY03]. It is to be mentioned here, that to our knowledge there exists no genuinely polynomial algorithm that yields an exact solution to the center problem in dimension $n \geq 3$. [Meg83] is not polynomial in the dimension, [KMY03] only provides an $(1 + \varepsilon)$ approximation and [FGK03] is not guaranteed to terminate after a polynomial number of steps. Nevertheless, [FGK03] show that their exact solver performs well in practice even for high dimensions. In case that their algorithm does not solve an instance in a reasonable running time one might still refuge to the approximation algorithm of [KMY03] or to [Meg83], when the dimension is not too big.

In the other case, where A_0 is active and we do not end up with a standard center problem, we make use of this a priori knowledge. So the general approach for a solution algorithm is as follows:

- Solve the relaxation of (S_{r_2}) and obtain (\tilde{x}, \tilde{r}) as unique optimal solution.
- Check if $\|A_0\|_2 + \|\tilde{x}\|_2 \leq \tilde{r}$:
 - If so, then (\tilde{x}, \tilde{r}) is optimal to (P_{r_2}) .
 - Otherwise, solve the problem (P_{r_2}) , but with the constraint $\|x\|_2 + r_2 = r$ instead of $\|x\|_2 + r_2 \leq r$, as A_0 has to be active.

Since there exist several solution methods to the standard Euclidean center-problem, the main problem is the version of (P_{r_2}) where A_0 is active. Thus, we concentrate on solving this problem in what follows.

It turns out that it is not necessary to use the relaxation (S_{r_2}) to check if A_0 is active when $n = 2$:

Theorem 4.19. *In \mathbb{R}^2 , A_0 is always active.*

Proof. Our goal is to show that the smallest enclosing ball $B_{r^*}(x^*)$ of the points $\{A_1, \dots, A_k\}$ satisfies $B_{r^*}(x^*) \cap \mathbb{R}_-^2 \subset \{0\}$. Then $\|x^*\|_2 \geq r^*$ and $\|A_0\|_2 + \|x^*\|_2 > \|x^*\|_2 \geq r^*$, so that A_0 has to be active in that case.

In \mathbb{R}^2 , a sphere is uniquely generated by at most three points on its boundary, so any smallest enclosing ball is generated by two or three active facilities. We consider these two cases for $B_{r^*}(x^*)$. First, suppose that exactly two facilities are active and call them A, B . Then,

$$x^* = \frac{A + B}{2}, \quad r^* = \left\| \frac{A - B}{2} \right\|_2.$$

Since $A, B \in \{A_1, \dots, A_k\} \subset \mathbb{R}_+^2$, it follows directly that

$$\|x^*\|_2 = \left\| \frac{A + B}{2} \right\|_2 \geq \left\| \frac{A - B}{2} \right\|_2 = r^*.$$

since we have $A_i + B_i \geq A_i - B_i$ in both coordinates $i = 1, 2$.

Now suppose that three facilities are active and call them A, B, C . If A, B, C are all translated by some vector $x \in \mathbb{R}^2$, then their smallest enclosing ball is just the translated ball $B_{r^*}(x^* + x)$. We use this to consider the “worst-case scenario”, where $B_{r^*}(x^*)$ has the biggest intersection with \mathbb{R}_-^2 . Therefore, translate A, B, C by

$$s := - \begin{pmatrix} \min_{(x_1, x_2) \in \{A, B, C\}} x_1 \\ \min_{(x_1, x_2) \in \{A, B, C\}} x_2 \end{pmatrix} \in \mathbb{R}_-^2.$$

Note that by the definition of s it holds that $A + s, B + s, C + s \in \mathbb{R}_+^2$. For any $y \in \mathbb{R}_-^2$ it holds that if $B_r(x + y) \cap \mathbb{R}_-^2 \subset \{0\}$ it follows that $B_r(x) \cap \mathbb{R}_-^2 \subset \{0\}$, so

it suffices to consider this case where we shift the ball as far as possible by using s . Now there are two different cases. First assume there exists one point, say A , for which $A + s = 0$. Since $A + s, B + s, C + s$ are in \mathbb{R}_+^2 the center of the optimal solution $x^* + s$ is also in \mathbb{R}_+^2 , since it has to lie in $\text{conv}\{A + s, B + s, C + s\} \subset \mathbb{R}_+^2$. Now consider the unique supporting hyperplane of $B_{r^*}(x^* + s)$ at $A + s = 0$. It has the normal vector $x^* + s \in \mathbb{R}_+^2$, so the components of the orthogonal vector have different signs. Thus, this supporting hyperplane is a line through 0 with negative slope, so it is a separating hyperplane of $B_{r^*}(x^* + s)$ and $\mathbb{R}_-^2 \Rightarrow B_{r^*}(x^*) \cap \mathbb{R}_-^2 = \{0\}$.

In the second case the situation is that one of the three points $A + s, B + s, C + s$ is contained in $\{0\} \times \mathbb{R}_+$ and another one is contained in $\mathbb{R}_+ \times \{0\}$. W.l.o.g. let this be $A + s \in \{0\} \times \mathbb{R}_+, B + s \in \mathbb{R}_+ \times \{0\}$. To simplify notation, set $A' = A + s, B' = B + s, C' = C + s$, compare Figure 4.2 for a depiction after the translation. We additionally know that all three points are active, so C' may not be contained in the smallest enclosing ball of A' and B' which is given by $B_{\tilde{r}}(\tilde{x})$. In formulae, we have $C' \in \mathbb{R}_+^2 \setminus B_{\tilde{r}}(\tilde{x})$, with $\tilde{r} = \left\| \frac{A' - B'}{2} \right\|_2$ and $\tilde{x} = \frac{A' + B'}{2}$.

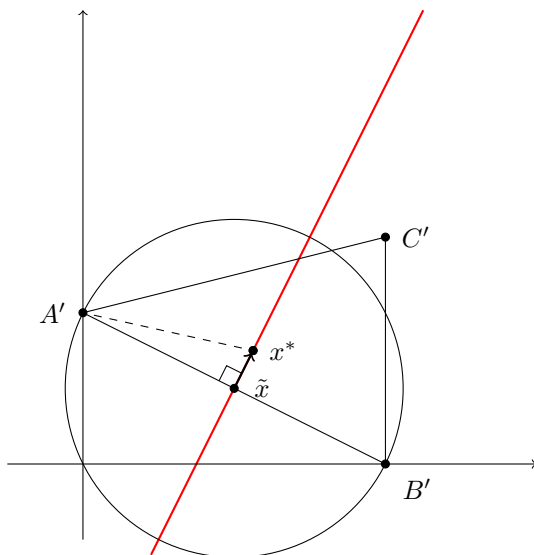


Figure 4.2: Depiction of the smallest enclosing ball for $\{A', B'\}$ and the perpendicular bisector which contains the center of the smallest enclosing ball for $\{A', B', C'\}$.

x^* and \tilde{x} both have to be on the equidistant line of A' and B' , which is given by the orthogonal vector $v \perp (A' - B')$ which can be chosen such that $\|v\|_2 = 1, v \in \mathbb{R}_+^2$.

Then it follows that there exists a $\delta \in \mathbb{R}$ such that $x^* = \tilde{x} + \delta \cdot v$. If $\delta > 0$, then

$$\begin{aligned} \|x^*\|_2^2 &= \|\tilde{x} + \delta \cdot v\|_2^2 = \left\| \frac{A' + B'}{2} \right\|_2^2 + \delta^2 \cdot \|v\|_2^2 + 2\langle \tilde{x}, \delta \cdot v \rangle \\ &> \left\| \frac{A' + B'}{2} \right\|_2^2 + \delta^2 \cdot \|v\|_2^2 = \|A - x^*\|_2^2 = r^{*2}, \end{aligned}$$

using the Pythagorean theorem in the second but last equation. Thus, $\|x^*\|_2 > r^*$, so $B_{r^*}(x^*) \cap \mathbb{R}_-^2 = \emptyset$.

So it remains to be shown that $\delta > 0$ for the representation $x^* = \tilde{x} + \delta \cdot v$. As A', B' and 0 are all in $B_{r^*}(x^*)$ it follows that the triangle $\Delta(A', B', 0)$ is a subset of $B_{r^*}(x^*)$, i.e., $\Delta(A', B', 0) \subset B_{r^*}(x^*)$. As $C' \in \mathbb{R}_+^2 \setminus B_{r^*}(x^*)$ this implies $C' \notin \Delta(A', B', 0)$. Thus the minimal distance from $C' \in \mathbb{R}_+^2$ to the perpendicular bisector of A, B is attained for $\tilde{x} + t \cdot v$ for some $t > 0$. Since $\|C' - x^*\|_2 = r^* < \|C' - \tilde{x}\|_2$ this implies that $x^* = \tilde{x} + \delta \cdot v$ with $\delta > 0$. \square

Theorem 4.19 does not generalize to higher dimensions, i.e., in \mathbb{R}^n , for any $n \geq 3$, it is not guaranteed that A_0 is active. We show this with the following example:

Example 4.4.1. *First, consider the problem in \mathbb{R}^3 with facilities*

$$A_0 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 9 \\ 0 \\ 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 \\ 9 \\ 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 \\ 0 \\ 9 \end{pmatrix}.$$

We want to find the smallest enclosing ball for A_1, A_2, A_3 . Since for no pair of $A_i, A_j, i, j \in \{1, 2, 3\}$ it holds that $A_l, l \neq \{i, j\}$, is in the smallest enclosing ball of A_i, A_j , it follows that all of the three facilities have to be active. Thus, the optimal solution lies on the line L of equal distance of those three points, which is given by

$$L = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3 : \lambda \in \mathbb{R} \right\}.$$

Then the point $x^ \in L$ with minimal distance to A_1, A_2, A_3 is the center of the smallest enclosing ball of A_1, A_2, A_3 . Computing the minimal squared distance of A_1 to L , we receive*

$$\min_{\lambda \in \mathbb{R}} \left\{ \left\| \begin{pmatrix} 9 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} \lambda \\ \lambda \\ \lambda \end{pmatrix} \right\|_2^2 \right\} = \min_{\lambda \in \mathbb{R}} \{3\lambda^2 - 18\lambda + 81\} = 54 =: r^{*2},$$

which is attained for $\lambda = 3$ with $x^ = (3 \ 3 \ 3)^t$ as optimal solution and the unique smallest enclosing ball for A_1, A_2, A_3 is $B_{r^*}(x^*)$. As $d(x^*, A_0) = \|x^*\|_2 + \|A_0\|_2 =$*

$\sqrt{3^2 + 3^2 + 3^2} + 1 = \sqrt{27} + 1 < 7 < \sqrt{54} = r^*$ it follows that $A_0 \in B_{r^*}(x^*)$. So (x^*, r^*) is the optimal solution to (P_{r^*}) and A_0 is not active.

This counterexample can be generalized to \mathbb{R}^n , $n > 3$, by adding additional zeros to the vectors A_1, A_2, A_3 . It is clear that the optimal solution to these three points has to have zeros in all components that are added (otherwise, removing them yields a better solution), so the optimal solution coincides with the one in \mathbb{R}^3 with additional zeros. Then the optimal solution still satisfies $\|x^*\|_2 + \|A_0\|_2 < r^*$, i.e., A_0 is not active.

An Exact Solution Algorithm in the Plane

Our goal is now to develop a solution algorithm in the plane, where Theorem 4.19 gives us the a priori knowledge that A_0 is active. There are several solution approaches to solve the center problem in the plane, but here we will focus on adapting the Elzinga-Hearn algorithm [EH72] which makes use of active facilities for which our a priori knowledge is useful. First, recall that a ball in \mathbb{R}^2 is uniquely determined by three points on its boundary. Then any optimal solution to the center problem either has two or three active facilities whose smallest enclosing ball is the ball about the optimal solution with minimal radius. The Elzinga-Hearn algorithm uses this idea and searches for the subset of cardinality two or three that determines the optimal solution.

In order to apply this idea to our setting we need to show that for the fixed gate point center version, i.e., for the balls with respect to the fixed gate point distance d , it also holds that three points on the boundary suffice to uniquely determine a ball. Before we prove the result for three points we investigate the case that only two points are active. Then the optimal solution is given as follows:

Lemma 4.20. *Let $x_1 \in \mathbb{R}_+^2$, $x_2 \in \mathbb{R}_+^2$ and suppose that $\|x_1\|_2 \geq \|x_2\|_2$. Then the smallest enclosing ball of x_1 and x_2 is $B_{\bar{r}}(\bar{x})$ with*

$$\bar{x} = \frac{\|x_1\|_2 - \|x_2\|_2}{2\|x_1\|_2} \cdot x_1 \quad \bar{r} = \frac{\|x_1\|_2 + \|x_2\|_2}{2}$$

Proof. First we check if $B_{\bar{r}}(\bar{x})$ contains both points:

$$\begin{aligned} d(x_1, \bar{x}) &= \|\bar{x} - x_1\|_2 = \left| \left(\frac{\|x_1\|_2 - \|x_2\|_2}{2\|x_1\|_2} - 1 \right) \right| \cdot \|x_1\|_2 \\ &= \left| \left(\frac{\|x_1\|_2 - \|x_2\|_2 - 2\|x_1\|_2}{2\|x_1\|_2} \right) \right| \cdot \|x_1\|_2 = \frac{\|x_1\|_2 + \|x_2\|_2}{2} \end{aligned}$$

and

$$d(x_2, \bar{x}) = \|x_2\|_2 + \|\bar{x}\|_2 = \|x_2\|_2 + \frac{\|x_1\|_2 - \|x_2\|_2}{2} = \frac{\|x_1\|_2 + \|x_2\|_2}{2}.$$

Data: Facilities $\mathcal{A} = \{A_1, \dots, A_M\} \subset \mathbb{R}^2$
Initialization: Active set $S = \{A_1, A^2\}$, $x_{\max} = 0 \in \mathbb{R}^2$, $r_{\max} = 0$
 $x = \frac{A_1 + A^2}{2}$, $r = \frac{\|A_1 - A^2\|_2}{2}$;
if (x, r) *feasible for* \mathcal{A} **then**
| **return** (x, r)
end
else
| $S = S \cup \{A_k\}$, for A_k that is farthest away from x .
end
while 1 **do**
| $x_{\max} = (0, 0)$, $r_{\max} = 0$
| $x = (0, 0)$, $r = 0$
| **for** $\mathcal{A}_c = \{A_i, A_j, A_k\} \subset S$ **do**
| | Construct the optimal solution x with objective value r for \mathcal{A}_c .
| | **if** $r > r_{\max}$ **then**
| | | $x_{\max} = x$, $r_{\max} = r$, $S' = \{A_i, A_j, A_k\}$
| | **end**
| **end**
| **if** (x_{\max}, r_{\max}) *feasible* **then**
| | **return** (x_{\max}, r_{\max})
| **end**
| **else**
| | $S = S' \cup \{A_k\}$, for A_k that is farthest away from x_{\max} .
| **end**
end
Output: $x_{\max} \in \mathbb{R}^2$, with optimal objective value r_{\max} .

Algorithm 1: Algorithm of Elzinga-Hearn.

As $d(x_1, x_2) = \|x_1\|_2 + \|x_2\|_2$ it follows that the radius of a ball containing both points has to be greater or equal $\frac{\|x_1\|_2 + \|x_2\|_2}{2}$. Thus, the radius \bar{r} is minimal. Moreover, this is the unique ball with this radius, as there does not exist a point $z \in \mathbb{R}_+^2$, s.t. $d(x_1, z) = \bar{r}$ as well as $\|z\|_2 = \bar{r} - \|x_2\|_2$. \square

Knowing the solution for this case we can prove that three points determine a unique smallest enclosing ball.

Lemma 4.21. *Let $x, y \in \mathbb{R}_+^2, z \in \mathbb{R}_-^2$ be such that $\|x\|_2 \geq \max\{\|y\|_2, \|z\|_2\}$. Moreover, assume that the smallest enclosing ball with respect to the distance d for no two of those points contains the third one.*

Then there exists a unique circle $C = \partial B_{\bar{r}}(\bar{x})$ such that $x, y, z \in C$.

Proof. First, assume $\|x\|_2 = \|z\|_2$. Lemma 4.20 implies that the smallest enclosing ball of x and z is $\bar{x} = 0, \bar{r} = \|x\|_2$. Then $y \in B_{\bar{r}}(\bar{x})$, so the smallest enclosing ball assumption is not met. Thus $\|z\|_2 < \|x\|_2$. Moreover, this implies that the center \bar{x} has to be contained in \mathbb{R}_+^2 . If it was in \mathbb{R}_-^2 , $\|z\|_2 < \|x\|_2$ would imply that $\|\bar{x} - z\|_2 < \|\bar{x}\|_2 + \|x\|_2$, so the two points cannot be equidistant to \bar{x} .

Consider the problem (P_{r_2}) with $r_2 = s$ and facilities $\mathcal{A} = \{x, y\}$. According to Theorem 4.18 this problem has a unique optimal solution (x^*, r^*) , where x, y, z all have distance r^* to the point x^* , i.e., $x, y, z \in \partial B_{r^*}(x^*)$. This already proves existence of such a circle $C = \partial B_{r^*}(x^*)$ and we now show its uniqueness. Assume there exists another ball $B_{r'}(x')$ such that $x, y, z \in \partial B_{r'}(x')$. Then $r' > r^*$, due to the uniqueness of the optimal solution to (P_{r_2}) . Since $\|x - x'\|_2 = \|y - x'\|_2$, it holds that $x' \in F = \{u \in \mathbb{R}^2 : \|u - x\|_2 = \|u - y\|_2\}$, the perpendicular bisector of x and y .

For F we can give the following representation: $F = \{M + \lambda v : \lambda \in \mathbb{R}\}$ where $M := \frac{x+y}{2}$ and $v \perp (x - y)$ such that $x^* = M + \mu v$ with $\mu < 0$. We show that $v \in \mathbb{R}_+^2$.

v only depends on x and y , so we first carefully investigate the relationship between x and y . Consider the following cases:

- 1) $y \geq x$, where \geq means \geq in both components. This is impossible, since $x \neq y$ and $\|x\|_2 > \|y\|_2$.
- 2) $y \leq x$. The smallest enclosing ball for z and x is given by $B_{\tilde{r}}(\tilde{x})$ as in Lemma 4.20. In particular, $\tilde{x} \in \mathbb{R}_+^2$ and $x' \leq x$, so $\|y - \tilde{x}\|_2 < \|x - \tilde{x}\|_2$ which implies that y is contained $B_{\tilde{r}}(\tilde{x})$, violating the prerequisites of the lemma.
- 3) The only remaining case is $x_1 \geq y_1$ and $y_2 \geq x_2$ or $x_1 \leq y_1$ and $y_2 \leq x_2$. In both cases, one of the two orthogonal vectors to $x - y$

$$\begin{pmatrix} y_2 - x_2 \\ x_1 - y_1 \end{pmatrix}, \begin{pmatrix} x_2 - y_2 \\ y_1 - x_1 \end{pmatrix}.$$

is contained in \mathbb{R}_+^2 and we choose v (for the definition of F) as the version above such that $v \in \mathbb{R}_+^2$ holds. Because z is not contained in the smallest enclosing ball of x and y , which is $B_{\tilde{r}}(\tilde{x})$, with $\tilde{r} = \frac{\|x\|_2 + \|y\|_2}{2}$ due to Lemma 4.20, it automatically follows that $x^* = M + \mu v$ with $\mu < 0$, since the distance to z gets smaller in this direction.

We now get back to $B_{r'}(x')$, the ball which is not the unique smallest enclosing ball for x, y, z but for which we assumed $x, y, z \in \partial B_{r'}(x')$.

We already know that $x' \in F \cap \mathbb{R}_+^2$. Furthermore,

$$\|z\|_2 + \|x'\|_2 = r' > r^* = \|z\|_2 + \|x^*\|_2,$$

so $\|x'\|_2 > \|x^*\|_2$. Our goal is to show that $d(x, x') \neq d(z, x')$ and we do this for two different cases: First, assume $x' = M + tv$ with $t > 0$. Then

$$\begin{aligned}
& \|x'\|_2^2 > \|x - x'\|_2^2 \Leftrightarrow \|M + tv\|_2^2 > \|x - M\|_2^2 + \|x' - M\|_2^2 \\
& \Leftrightarrow \|M\|_2^2 + \lambda^2 \|v\|_2^2 + 2\lambda \langle M, v \rangle < \left(\frac{\|x - y\|_2}{2} \right)^2 + \lambda^2 \|v\|_2^2 \\
& \Leftrightarrow \|M\|_2^2 + 2\lambda \langle M, v \rangle > \frac{\|x - y\|_2^2}{4} \\
& \Leftrightarrow 4 \left\| \frac{x + y}{2} \right\|_2^2 + 8\lambda \langle M, v \rangle > \|x\|_2^2 + \|y\|_2^2 - 2\langle x, y \rangle \\
& \Leftrightarrow 2\langle x, y \rangle + 8\lambda \langle M, v \rangle > 0
\end{aligned}$$

and the latter inequality holds since $\lambda > 0$ and $x, y, M, v \in \mathbb{R}_+^2$ with non-zero components. Hence, $d(z, x') = \|z\|_2 + \|x'\|_2 > \|x'\|_2 > \|x - x'\|_2 = d(x, x')$, what was to show.

The second case is that $x' = M + tv$ with $t \leq 0$. As $\|x'\|_2 > \|x^*\|_2$, $x' = M + tv$ may be expressed as $x' = \lambda x^* + (1 - \lambda)M \in F$ for some $\lambda \in [0, 1)$. Note that $\lambda \neq 1$ as $x' \neq x^*$. Using $v \in \mathbb{R}_+^2$ and the fact that the distance of any point p on F to x is $\sqrt{\|x - M\|_2^2 + \|M - p\|_2^2}$ we get

$$\begin{aligned}
r'^2 &= \|x - x'\|_2^2 = \|x - M\|_2^2 + \|M - x'\|_2^2 \\
&= \|x - M\|_2^2 + \|M - (\lambda x^* + (1 - \lambda)M)\|_2^2 = \|x - M\|_2^2 + \|\lambda(M - x^*)\|_2^2 \\
&< \|x - M\|_2^2 + \|M - x^*\|_2^2 = \|x - x^*\|_2^2 = r^{*2},
\end{aligned}$$

so $d(x, x') = \|x - x'\|_2 < r^*$. Moreover,

$$d(z, x') = \|z\|_2 + \|x'\|_2 > \|z\|_2 + \|x^*\|_2 = r^*,$$

so $d(z, x') > r^* > d(x, x')$.

So for both cases we get that $d(z, x') \neq d(x, x')$ contradicting $x, z \in \partial B_{r'}(x')$ and thus implying uniqueness of $B_{r^*}(x^*)$. \square

Lemma 4.21 implies that at most three active facilities suffice to find an optimal solution to (P_{r_2}) within \mathbb{R}_+^2 . Additionally, we know by Theorem 4.19 that there always exists exactly one active facility, A_0 , in \mathbb{R}_-^2 . Such an a priori knowledge of an active facility does not exist in the standard center problem and we make use of that in the adapted Elzinga-Hearn algorithm.

The next lemma shows that the smallest enclosing ball condition of Lemma 4.21 can easily be checked by only looking at one pair of facilities; in the case where the condition is not satisfied, this pair already yields the optimal solution. Recall that $\mathcal{A}_1 = \{A_1, \dots, A_k\}$ is sorted descendingly with respect to the norms of the facilities A_i .

Lemma 4.22. *Let $r_1 = \|A_1\|_2$. Then either $x_A := \frac{r_1 - r_2}{2r_1}A_1$ is the optimal solution or three facilities are active.*

Proof. Take any facility A_j other than A_1 . Then $r_j := \|A_j\|_2 \leq r_1$ and the smallest enclosing ball for A_j and A_0 is centered at $\frac{r_j - r_2}{2r_j}A_j$ with radius $\frac{r_j + r_2}{2}$ due to Lemma 4.20. Moreover, $\|\frac{r_j - r_2}{2r_j}A_j\|_2 = \frac{r_j - r_2}{2}$. Then we can calculate the distance of A_1 to the center of this ball,

$$\begin{aligned} \left\|A_1 - \frac{r_j - r_2}{2r_j}A_j\right\|_2 &\geq \left|\|A_1\|_2 - \left\|\frac{r_j - r_2}{2r_j}A_j\right\|_2\right| = \left|r_1 - \frac{r_j - r_2}{2}\right| \\ &= r_1 - \frac{r_j - r_2}{2} \geq r_1 - \frac{r_1 - r_2}{2} = \frac{r_1 + r_2}{2} \geq \frac{r_j + r_2}{2}, \end{aligned}$$

which is at least $\frac{r_j + r_2}{2}$. Furthermore we know that $A_1 \neq A_j$ so we either have $r_j < r_1$ or $A_1 \neq \lambda A_j$ for all $\lambda \in \mathbb{R}$. If $r_j < r_1$, then the last inequality holds strictly and if $A_1 \neq \lambda A_j$, the first inequality holds strictly. $\Rightarrow \left\|A_1 - \frac{r_j - r_2}{2r_j}A_j\right\|_2 > \frac{r_j + r_2}{2}$, so the optimal ball corresponding to A_0 and A_j does not contain A_1 . Thus, x_A is the only possible candidate for x^* defined by two active facilities on $\partial B_{r^*}(x^*)$. In all other cases it must be defined by three points on its boundary. \square

Lemma 4.22 implies that we can simply check, whether $B_{\frac{r_1 + r_2}{2}}x_A$ contains all A_j and is optimal. If it does, x_A is optimal and in the case where it does not, we know that three facilities are active. Thus, the essential problem that remains to be solved is a fixed gate point center problem with one given facility $A_0 \in \mathbb{R}_-^2$ and two facilities $A, B \in \mathbb{R}_+^2$, where no two facilities yield an optimal solution, i.e., no smallest enclosing ball of two facilities contains the third one. We denote this problem as $(P_{\text{ess}}) = (P_{\text{ess}})(A_0, A, B)$, where w.l.o.g. let $r_1 = \|A\|_2 \geq \|B\|_2$.

$$\begin{aligned} (P_{\text{ess}})(A_0, A, B) \quad &\min \max_{A \in \{A_0, A, B\}} d(x, A) \\ \text{s.t.} \quad &x \in \mathbb{R}_+^2 \end{aligned}$$

When no smallest enclosing ball of two facilities yields an optimal solution, then all three facilities are active. Then the center of the smallest enclosing ball of A_0, A, B , x^* satisfies $d(x, A_0) = d(x, A) = d(x, B)$. We use this fact to characterize the optimal solution. As $x^* \in \mathbb{R}_+^2$, we can rewrite the distances d to get $\|x\|_2 + \|A_0\|_2 = \|x - A\|_2 = \|x - B\|_2$. and define the sets

$$E = \{x \in \mathbb{R}^2 : \|x\|_2 + r_2 = \|A - x\|_2\}, \text{ and } F = \{x \in \mathbb{R}^2 : \|A - x\|_2 = \|B - x\|_2\}.$$

Lemma 4.23. *Assume $\|A\|_2 = r_1 > r_2$. Then $E = \{x \in \mathbb{R}^2 : \|A - x\|_2 = \|x\|_2 + r_2\}$ is a branch of a hyperbola.*

Proof. A hyperbola in \mathbb{R}^2 is given by

$$\{x \in \mathbb{R}^2 : \left| \|F_1 - x\|_2 - \|F_2 - x\|_2 \right| = 2 \cdot a \},$$

where $F_1, F_2 \in \mathbb{R}^2$ are the two *foci* of the hyperbola. Setting $a = \frac{r_2}{2}$, $F_1 = A$, $F_2 = 0$ this reformulates to

$$\{x \in \mathbb{R}^2 : \left| \|A - x\|_2 - \|x\|_2 \right| = r_2 \}.$$

Taking the branch of the hyperbola where $\|A - x\|_2 \geq \|x\|_2$ holds, we get exactly

$$\{x \in \mathbb{R}^2 : \|A - x\|_2 = r_2 + \|x\|_2\} = E.$$

□

Notation 4.4.2. *We use the following notation:*

$$\begin{aligned} H &= \{x \in \mathbb{R}^2 : \left| \|A - x\|_2 - \|x\|_2 \right| = r_2\}, \text{ the hyperbola containing } E \text{ as a branch} \\ F &= \{x \in \mathbb{R}^2 : \|A - x\|_2 = \|B - x\|_2\} \text{ the perpendicular bisector of } A, B. \end{aligned}$$

Corollary 4.24. *The optimal solution x^* is contained in $D := \{x \in \mathbb{R}^2 : \|A - x\|_2 = \|x\|_2 + r_2 = \|B - x\|_2\} = E \cap F$ and $|D| \leq 2$.*

Proof. We already know that there exists a unique optimal solution $x^* \in \mathbb{R}_+^2$ to our fixed gate point center problem, that satisfies $d(x^*, A_0) = d(x^*, A) = d(x^*, B)$, so $x^* \in D$. The intersection of a line and a hyperbola is given by a quadratic equation, thus $|H \cap F| \leq 2$. For our branch $E \subset H$ this implies $|E \cap F| \leq |H \cap F| \leq 2$. □

In order to calculate the intersection of E and F , we need a convenient representation of the hyperbola H . First of all, let us clarify notation: Let P be the *midpoint* of the two foci A and 0 . Then the distance of P to the intersection of the hyperbola with the *major axis* $M = \{\lambda A : \lambda \geq 0\}$ is $a = \frac{r_2}{2}$ and the *linear eccentricity* e of the hyperbola is given by $e = \frac{\|A\|_2}{2}$. The values a and e are invariant under translations and rotations of the hyperbola, so when we transform the hyperbola H such that the midpoint of the two foci is 0 and the major axis is the x-axis, then the transformed version \mathcal{H} of H can be represented as

$$\mathcal{H} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\}, \quad (4.11)$$

which is the standard equation form of a hyperbola. We now determine σ , the transformation of the hyperbola to this standard form. σ consists of two maps. First, a rotation is applied, then a translation applied. An example for the rotation and the translation afterwards is depicted in the Figures 4.3 , 4.4 , 4.5. The figures

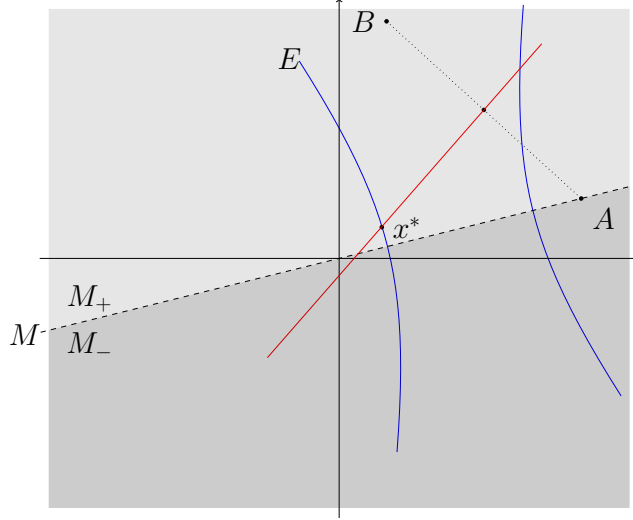


Figure 4.3: Initial Position: The perpendicular bisector of A and B (red) intersects both branches of the hyperbola (blue) yielding the optimal solution x^* in the branch E .

also show the perpendicular bisector F of A and B , to which we also apply the transformation. Then the intersection of the correct branch of the transformed hyperbola $\sigma(H) = \tilde{H}$ and the transformed bisector $\sigma(F)$ contains the transformed optimal solution $\sigma(x^*)$ and when we invert the transformation we get the optimal solution x^* . First of all we rotate the branch, such that its major axis M is the x -axis. Since the major axis is $\{\lambda A : \lambda \in \mathbb{R}\}$, we rotate by $\theta = -\varphi_A$, where φ_A is determined by $A = \|A\|_2 \cdot \begin{pmatrix} \cos(\varphi_A) \\ \sin(\varphi_A) \end{pmatrix}$ and denote the map as

$$R_\theta \left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

After applying R_θ to E , we only need to translate the rotated midpoint of the foci to the origin. The rotated midpoint is the midpoint of the two rotated foci, thus given by $\begin{pmatrix} \frac{\|A\|_2}{2} \\ 0 \end{pmatrix}$ and the translation T of this point to the origin is

$$T \left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right) = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} -\frac{\|A\|_2}{2} \\ 0 \end{pmatrix}.$$

After having applied the rotation and the translation, the branch of the hyperbola fits the standard form as in (4.11), and our transformation σ can be expressed as

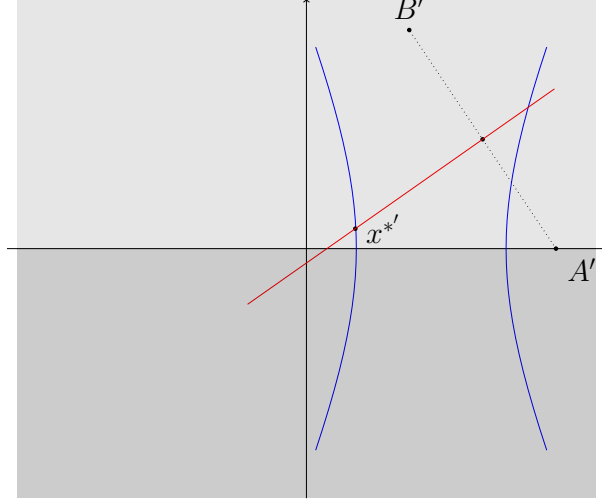


Figure 4.4: The points and sets from Figure 4.3 after applying $\sigma = T \circ R_\theta$.

$\sigma = T \circ R_\theta$. The standard parametrization of \mathcal{H} is

$$\mathcal{H} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\} = \left\{ \begin{pmatrix} -a \cdot \cosh(t) \\ b \cdot \sinh(t) \end{pmatrix} : t \in \mathbb{R} \right\}. \quad (4.12)$$

One branch of the standard parametrized \mathcal{H} corresponds to $a \cdot \cosh(t)$ and the other branch to $-a \cdot \cosh(t)$. Since $\sigma(E)$ is the branch of the hyperbola where

$$\|\sigma(A) - x\|_2 = \left\| \begin{pmatrix} \frac{\|A\|_2}{2} \\ 0 \end{pmatrix} \right\|_2 \geq \left\| \begin{pmatrix} -\frac{\|A\|_2}{2} \\ 0 \end{pmatrix} \right\|_2 = \|\sigma(0) - x\|_2$$

holds, we get the branch with negative values of x , i.e.,

$$\sigma(E) = \left\{ \begin{pmatrix} -a \cdot \cosh(t) \\ b \cdot \sinh(t) \end{pmatrix} : t \in \mathbb{R} \right\}.$$

When we apply σ to the line F as in the figures, we can calculate the intersection of $\sigma(F)$ and $\sigma(E)$ and obtain the optimal solution to our problem by applying σ^{-1} . An important observation is, that $|E \cap F| = 2$ might hold. In this case, one point of $E \cap F$ is contained in M_+ and one is contained in M_- , where M_+, M_- are the half spaces induced by the major axis M . The next lemma states a condition, which of the two points in $E \cap F$ yields the optimal solution in case that $|E \cap F| = 2$.

Recall, that $A \neq B \in \mathbb{R}_+^2$, $\|A\|_2 \geq \|B\|_2$ and that if $B = \lambda A$ with $0 \leq \lambda < 1$ it follows that $B \in B_{r_A}(x_A)$, the smallest enclosing ball of A and A_0 as in Lemma 4.20. For (P_{ess}) we assumed that no facility is contained in the smallest enclosing ball of the other two. Thus, $B \neq \lambda A$ for $\lambda \in \mathbb{R}$.

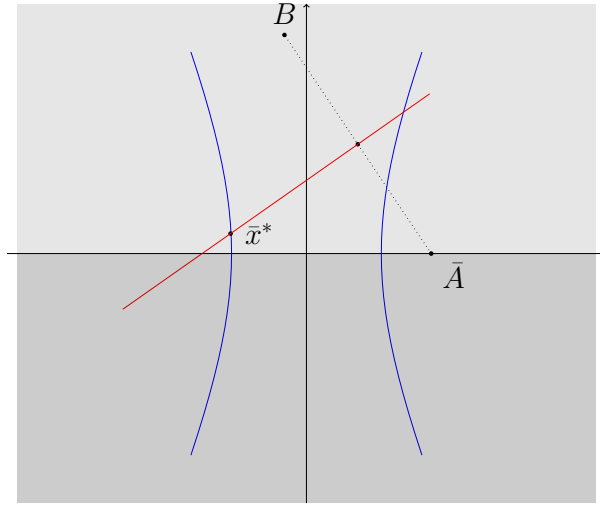


Figure 4.5: The points and sets from Figure 4.3, after applying R_θ and then σ .

Lemma 4.25. *Let $\varphi_B \in [0, \frac{\pi}{2}]$ be such that $B = \|B\|_2 \cdot \begin{pmatrix} \cos(\varphi_B) \\ \sin(\varphi_B) \end{pmatrix}$ and let $\varphi_A \in [0, \frac{\pi}{2}]$ be determined analogously. Then*

$$\{x^*\} = D \cap M_+ \Leftrightarrow \varphi_B > \varphi_A \quad \text{and} \quad \{x^*\} = D \cap M_- \Leftrightarrow \varphi_B < \varphi_A,$$

where $D = E \cap F$.

Proof. Since the hyperbola H is symmetric to its major axis M (the line through 0 and A) we consider w.l.o.g. the case where $\varphi_B > \varphi_A$. If $\varphi_B < \varphi_A$ we may reflect B on the major axis to receive this situation and reflect back the optimal solution resulting from A and the reflected B .

We have already shown that the optimal solution x^* is contained in $D = E \cap F$. So we need to show that $x^* \in M_+$.

As D is the intersection of a line with a hyperbola it follows that $|D| \leq 2$. We explicitly construct the point $y \neq x^*$ in D . Let \mathcal{R} denote the reflection upon the major axis M . As reflections are bijective, there exists a unique point y such that $\mathcal{R}(x^*) = y$. It follows from the definition of the reflection and the half spaces that $y \in M_- \Leftrightarrow x^* \in M_+$ and vice versa. Moreover, $x^* \in E \Leftrightarrow y \in E$ holds due to the symmetry of the branches of the hyperbola with respect to the major axis M . Thus, $D = \{x^*, y\}$. (On a side note, $|D| = 1 \Leftrightarrow x^* = \mathcal{R}(x^*)$.)

Now in order to show $x^* \in M_+$ we simply prove that the objective of the point $w \in D \cap M_-$ is bigger than the objective of the point $v \in M_+$. This suffices, as we already know that the optimal solution is contained in D .

Consider the transformed versions $\sigma(v), \sigma(w)$. The major axis $\sigma(M)$ is the x-axis, so for the coordinates of the vectors we have $\sigma(v)_1 = \sigma(w)_1$, $\sigma(v)_2 = -\sigma(w)_2$, as \mathcal{R} is simply the reflection upon the x -axis. Since $w \in D \cap M_-$ we have $\sigma(w)_2 < 0$, as $\sigma(M_-) = \mathbb{R} \times \mathbb{R}_-$. Moreover, $\varphi(B) > \varphi(A)$ implies $\sigma(B) \in \mathbb{R} \times \mathbb{R}_+$. Then

$$(\sigma(B)_2 + \sigma(w)_2)^2 < (\sigma(B)_2 - \sigma(w)_2)^2$$

and we can calculate

$$\begin{aligned} \|\sigma(B) - \sigma(v)\|_2^2 &= (\sigma(B)_1 - \sigma(v)_1)^2 + (\sigma(B)_2 - \sigma(v)_2)^2 \\ &= (\sigma(B)_1 - \sigma(w)_1)^2 + (\sigma(B)_2 + \sigma(w)_2)^2 \\ &< (\sigma(B)_1 - \sigma(w)_1)^2 + (\sigma(B)_2 - \sigma(w)_2)^2 \\ &= \|\sigma(B) - \sigma(w)\|_2^2. \end{aligned}$$

$\|\sigma(B) - \sigma(v)\|_2^2 < \|\sigma(B) - \sigma(w)\|_2^2$ is equivalent to

$$(1) \quad d(w, B) = \|B - w\|_2 > \|B - v\|_2 = d(v, B)$$

as $\sigma = T \circ R_\theta$ and rotations and translations are distance preserving. Furthermore, $\sigma(w)_1 = \sigma(v)_1$, $\sigma(w)_2 = -\sigma(v)_2$ imply that $\|w\|_2 = \|v\|_2$, so

$$(2) \quad d(w, A_0) = \|w\|_2 + r_2 = \|v\|_2 + r_2 = d(v, A_0).$$

As $w, v \in E$ we have $\|w\|_2 + r_2 = d(w, A_0) = d(w, A)$ and $\|v\|_2 + r_2 = d(v, A_0) = d(v, A)$. Together with (2) we receive

$$(3) \quad d(w, A) = d(v, A).$$

Now, (1), (2), (3) imply

$$\max(d(A_0, w), d(A, w), d(B, w)) \geq \max(d(A_0, v), d(A, v), d(B, v)).$$

If equality would hold then we would have two optimal solutions to (P_{ess}) , which contradicts the uniqueness proven in Theorem 4.18. Thus, it follows that

$$\max(d(A_0, w), d(A, w), d(B, w)) > \max(d(A_0, v), d(A, v), d(B, v))$$

and the optimal solution has to be $v = x^*$. As $w \in M_-$, $v \in M_+$ by choice it follows that $x^* \in \{D \cap M_+\}$.

That $|D \cap M_+| = 1$ follows from $|D \cap M| \leq |D| \leq 2$ and $|\{w\}| = |D \cap M_-| = 1$. \square

In the next theorem we combine the gathered knowledge: The optimal solution is found in $E \cap F$ and in case that $|E \cap F| = 2$ we know which point yields the optimal solution. Now, we basically plug in $y = mx + d$ gained from the line $\sigma(F)$ into the standard equation form (4.11) of the hyperbola

$$\mathcal{H} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\}$$

and then only need to re-transform the obtained solution.

Theorem 4.26. *Let $a, b, d, m \in \mathbb{R}$ be such that*

$$\sigma(F) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : y = mx + d \right\}, \sigma(E) = \left\{ \begin{pmatrix} -a \cdot \cosh(t) \\ b \cdot \sinh(t) \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Consider the quadratic equation

$$\frac{x^2}{a^2} - \frac{(mx + d)^2}{b^2} = 1. \quad (4.13)$$

If $\varphi_B > \varphi_A$, there exists a unique solution \bar{x} to (4.13) such that $\left(\begin{pmatrix} \bar{x} \\ m\bar{x} + d \end{pmatrix} \right) \in \sigma(D) \cap \mathbb{R} \times \mathbb{R}_+$. If $\varphi_B < \varphi_A$, there exists a unique solution \bar{x} to (4.13) such that $\left(\begin{pmatrix} \bar{x} \\ m\bar{x} + d \end{pmatrix} \right) \in \sigma(D) \cap \mathbb{R} \times \mathbb{R}_-$.

In both cases the optimal solution to (P_{ess}) is given by $x^ = \sigma^{-1} \left(\begin{pmatrix} \bar{x} \\ m\bar{x} + d \end{pmatrix} \right)$.*

Proof. Consider the case $\varphi_B > \varphi_A$. Lemma 4.25 states that the unique optimal solution to (P_{ess}) is $D \cap M_+ = \{x^*\}$, so $\sigma(x^*) \in \sigma(D) \cap \sigma(M_+) = \sigma(D) \cap \mathbb{R} \times \mathbb{R}_+$ holds. Since $\sigma(x^*) \in \sigma(D) = \sigma(E) \cap \sigma(F)$ it follows that $\sigma(x^*) = \begin{pmatrix} x' \\ mx' + d \end{pmatrix}$, so x' satisfies all requirements. In the proof of Lemma 4.25 we have shown that $D = D \cap M_+ \cup D \cap M_- = \{x^*\} \cup \{\mathcal{R}(x^*)\}$, where \mathcal{R} is the reflection upon the major axis M . Let $y = \mathcal{R}(x^*) \in D \cap M_-$. $y \in D$ implies $\sigma(y) = \begin{pmatrix} y' \\ my' + d \end{pmatrix}$ and $y \in M_-$ implies $\sigma(y) \in \sigma(D) \cap \mathbb{R} \times \mathbb{R}_-$. There are two cases: If $\sigma(y) \in \mathbb{R} \times (-\infty, 0)$, then $y \notin \mathbb{R} \times \mathbb{R}_+$ and x' is the sought unique solution of (4.13). In the other case $\sigma(y) \in \mathbb{R} \times \{0\} \subset \mathbb{R} \times \mathbb{R}_+$. But then $\mathcal{R}(y) = x^*$ and $|D| = 1$ so x' is again the sought unique solution of (4.13).

By definition, x^* is the unique optimal solution of (P_{ess}) , and thus

$$\sigma(x^*) = \begin{pmatrix} x' \\ mx' + d \end{pmatrix} \Leftrightarrow x^* = \sigma^{-1} \left(\begin{pmatrix} x' \\ mx' + d \end{pmatrix} \right)$$

holds as required.

The case $\varphi_B < \varphi_A$ works analogously by interchanging roles of M_- and M_+ . \square

Example 4.4.3. In this example we depict a case, where $|E \cap F| = 2$ (contrary to figures above, where $|E \cap F| = 1$) and demonstrate how to calculate the intersection $\sigma(E) \cap \sigma(F)$ and obtain the optimal solution to (P_{ess}) . Let

$$A = \begin{pmatrix} 8 \\ 0 \end{pmatrix}, B = \begin{pmatrix} 0.2 \\ 7.8 \end{pmatrix}, r_2 = 7.5.$$

Now we calculate the intersection of the transformed versions of F and E . As $\varphi(A) = 0$ already, we only need to shift the hyperbola by $\begin{pmatrix} -4 \\ 0 \end{pmatrix}$ to obtain a hyperbola in standard form.

After shifting we get

$$\sigma(A) = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \sigma(B) = \begin{pmatrix} -3.8 \\ 7.8 \end{pmatrix}.$$

Now we calculate the parameters for the hyperbola in standard form by using our original hyperbola in the other representation

$$\sigma(\mathcal{H}) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\} \quad H = \{x \in \mathbb{R}^2 : |||A - x||_2 - \|x\|_2 = r_2\}$$

Thus, $a = \frac{r_2}{2} = 3.75$ and b is calculated via the linear eccentricity $c = \|\sigma(A)\|_2 = 4$ and $a = 3.75$:

$$b = \sqrt{c^2 - a^2} = \sqrt{4^2 - 3.75^2} \approx 1.39$$

Thus, the parametrization of the branch $\sigma(E)$ is given by

$$\sigma(E) = \left\{ \begin{pmatrix} -a \cdot \cosh(t) \\ b \cdot \sinh(t) \end{pmatrix} : t \in \mathbb{R} \right\} \approx \left\{ \begin{pmatrix} -3.75 \cdot \cosh(t) \\ 1.39 \cdot \sinh(t) \end{pmatrix} : t \in \mathbb{R} \right\},$$

$\sigma(F)$ is given as the perpendicular bisector of $\sigma(A), \sigma(B)$:

$$F = \left\{ \begin{pmatrix} 0.1 \\ 3.9 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \text{ with } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \perp \sigma(A) - \sigma(B).$$

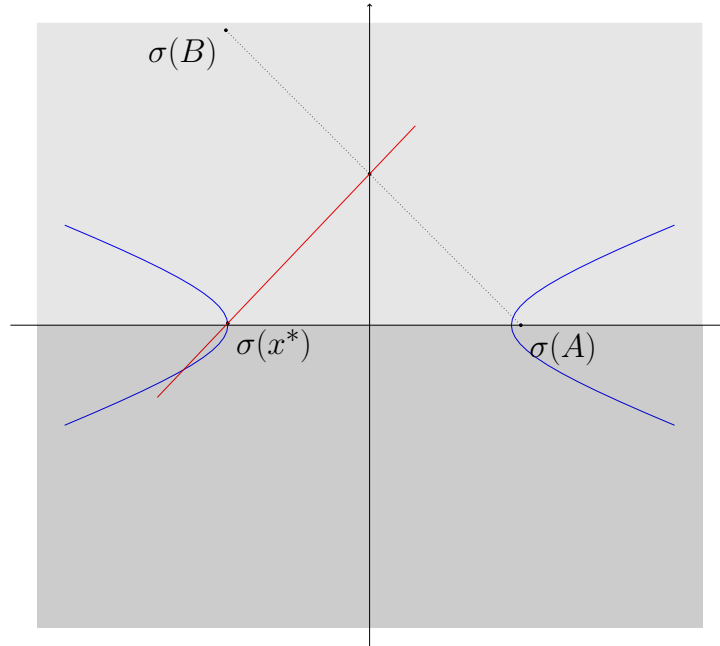


Figure 4.6: An example where the the perpendicular bisector of $\sigma(A)$ and $\sigma(B)$ intersects the branch $\sigma(E)$ of the hyperbola twice.

Then computing the intersection $\sigma(E) \cap \sigma(F)$ we get $\lambda = -0.1 - 3.75 \cdot \cosh(t)$ from the first coordinate and plugging into the second coordinate we receive $3.8 = 3.75 \cdot \cosh(t) + 1.39 \cdot \sinh(t)$. There are two solutions for this equation, $t_1 \approx -0.813$, $t_2 \approx 0.034$, thus $|E \cap F| = 2$. Plugging t_1 and t_2 into the parametrization

$$p(t) = \begin{pmatrix} -3.75 \cdot \cosh(t) \\ 1.39 \cdot \sinh(t) \end{pmatrix}$$

we get

$$p(t_1) \approx \begin{pmatrix} -5.06 \\ -1.26 \end{pmatrix} \quad p(t_2) \approx \begin{pmatrix} -3.75 \\ 0.05 \end{pmatrix}$$

The optimal solution x^* is contained in M_+ , as $\varphi(B) > 0 = \varphi(A)$. So $\sigma(x^*) \in \sigma(M_+)$ which holds for $\sigma(x^*) = p(t_2)$. The retransformation is given by simply shifting with $\begin{pmatrix} 4 \\ 0 \end{pmatrix}$, so

$$x^* = \begin{pmatrix} 0.25 \\ 0.05 \end{pmatrix}.$$

Now that we have found purely algebraic ways to determine the optimal solution to (P_{ess}) , the fixed gate point center problem for only three facilities including A_0 ,

we can come back to solving the general fixed gate point center problem. Recall that the idea is to search for the active facilities $A_i, A_j \in \mathcal{A}_1$ such that the smallest enclosing ball of A_0, A_i, A_j (determined by solving (P_{ess})) is the solution to the whole problem. Since the idea of searching the correct active set of facilities is the same as for the Elzinga-Hearn algorithm we adapt its structure, but based on our gathered knowledge we need to implement three changes:

- i) Theorem 4.19: A_0 is always active. \Rightarrow We only need to find active sets of cardinality at most two out of the remaining facilities in \mathbb{R}_+^2 .
- ii) Lemma 4.22: A_1 (the facility with maximal norm) is the only possible facility to yield an active set of cardinality two including A_0 .
- iii) The subproblem $(P_{\text{ess}})(A, B, A_0)$ is solved with Theorem 4.26. a, b can directly be calculated using $\|A\|_2$ and $\|A_0\|_2 = r_2$. m, d are calculated after the transformation σ has been applied to the line F . The transformation is determined by φ_A and $\|A\|_2$.

Implementing these changes, we state our *adapted Elzinga-Hearn algorithm*, see Algorithm 2.

Data: Facilities $\mathcal{A} = \{A_1, \dots, A_k\} \subset \mathbb{R}_+^2$ ordered descendingly with respect to their norm values, $A_0 \in \mathcal{A}_2$ with maximal norm $\|A_0\|_2 = r_2 > 0$ of all facilities in \mathcal{A}_2

Initialization: Active set $S = \{A_1\}$, $r_1 = \|A_1\|_2$.

$x = \frac{r_1 - r_2}{2 \cdot r_1} \cdot A_1$, $r = \frac{r_1 - r_2}{2}$. **if** (x, r) *feasible* **then**

 | $x_{\max} = x$, $r_{\max} = \frac{r_1 - r_2}{2}$ **return**

end

else

 | $S = S \cup \{A_l\}$, for A_l that is farthest away from x .

end

while 1 **do**

 | $x_{\max} = (0, 0)$, $r_{\max} = 0$

 | $x = (0, 0)$, $r = 0$

for $\{A_i, A_j\} \subset S$ **do**

 | Solve (P_{ess}) for A_i, A_j, A_0 with optimal solution x and objective value r .

if $r > r_{\max}$ **then**

 | $x_{\max} = x$, $r_{\max} = r$, $S' = \{A_i, A_j\}$

end

end

if (x_{\max}, r_{\max}) *feasible* **then**

 | **return**

end

else

 | $S = S' \cup \{A_l\}$, for A_l that is farthest away from x .

end

end

Output: The optimal solution $x_{\max} \in \mathbb{R}_+^2$ to (G_{\max}) with optimal objective value r_{\max} .

Algorithm 2: The *adapted Elzinga-Hearn algorithm* for the fixed gate point center problem.

We repeat some of the earlier arguments to cleanly prove correctness of the algorithm:

Theorem 4.27. *Algorithm 2 returns the optimal solution to (G_{\max}) .*

Recall, that the term *active* is always related to the unique optimal solution.

Proof. First of all, Theorem 4.18 ensures existence of a unique optimal solution x^* to (G_{\max}) . We consider two different cases:

Case 1: There is one active facility in \mathbb{R}_+^2 . Lemma 4.22 states that this can only be A_1 . This is checked in the initialization step, so if they are active, the optimal

solution is found there and the optimal solution is returned.

Case 2: There are at least two active facilities in \mathbb{R}_+^2 . Additionally, Theorem 4.19 states that A_0 is always active, so we have three active facilities. Lemma 4.21 implies that three points on the boundary of a ball with respect to distance d , i.e., our three active facilities, suffice to uniquely determine $B_{r^*}^d(x^*)$, the optimal ball which is the smallest enclosing ball for the three active facilities.

Hence, it suffices to search for two facilities $\{A_i, A_j\}$ in \mathbb{R}_+^2 that determine this ball, which is done in the while loop. The unique optimal solutions of the (P_{ess}) subproblems in the while loop are determined by applying Theorem 4.26.

It remains to be shown that the while loop terminates, i.e., that the correct active set is found. To this end we show that the radius increases in every iteration. Consider a fixed iteration and let $S' = \{A_i, A_j\}$ be the two facilities that yield (x_{\max}, r_{\max}) for this iteration. If the optimality check for x_{\max}, r_{\max} does not fail, it is an optimal solution and we are done. Otherwise, let A_k be the facility which is farthest away from x_{\max} , i.e., not contained in $B_{r_{\max}}(x_{\max})$. Then let (\tilde{x}, \tilde{r}) denote the unique optimal solution to the problem with facilities $\{A_0, A_i, A_j, A_k\}$ which is the problem that is solved in the next iteration, i.e., $S = \{A_i, A_j, A_k\}$. We now need to show that $\tilde{r} > r_{\max}$. Since the first iteration does not yield an optimal solution it follows that $\|A_k - x_{\max}\|_2 > r_{\max}$, so $(x_{\max}, r_{\max}) \neq (\tilde{x}, \tilde{r})$, so the active set for (\tilde{x}, \tilde{r}) can not be $\{A_0, A_i, A_j\}$, so $S' \neq \{A_i, A_j\}$. Thus, A_k has to be active and there exists $A \in \{A_i, A_j\}$ such that $\{A_0, A, A_k\}$ is the active set for (\tilde{x}, \tilde{r}) . Now, if $\tilde{x} = x_{\max}$, then $\tilde{r} > r_{\max}$, because $A_k \notin B_{r_{\max}}(x_{\max})$, but $A_k \in B_{\tilde{r}}(x_{\max})$. If $\tilde{x} \neq x_{\max}$, then $\tilde{r} > r$. Otherwise, if $\tilde{r} \leq r_{\max}$ would hold, then (\tilde{x}, \tilde{r}) was optimal to $\{A_0, A_i, A_j\}$, and contradicts (x_{\max}, r_{\max}) being the unique optimal solution for the active set $\{A_0, A_i, A_j\}$.

Thus, $\tilde{r} > r_{\max}$ and the value of r_{\max} increases in every iteration. Since there only is a finite number of subsets $\{A_i, A_j\} \subset \mathcal{A}_1$, the algorithm terminates after a finite amount of steps. \square

Approximations for the Center Problem in Higher Dimensions

In the previous section we developed an exact solution algorithm based on active facilities for the fixed gate point center problem in \mathbb{R}^2 . This method, however, does not generalize to higher dimensions. The first problem is the increased amount of “degenerate” cases (where $2 \leq k < n + 1$ facilities are active). The second one is to find the solution to the non-degenerate case where $n + 1$ facilities are active, since there is no trivial generalization of hyperbolas and their parametrizations to higher dimensions.

Therefore we propose an approach to approximately solve the fixed gate point center problem for $n \geq 3$. Unfortunately, the counterexample of Theorem 4.19 has shown that with $n \geq 3$ we lose the a priori knowledge that A_0 is active. Recall that in order

to determine whether or not A_0 is active, we may solve the standard Euclidean center problem (S_{r_2}) (p.69) for $\{A_1, \dots, A_k\}$, simply disregarding A_0 . Note that, since we may work in higher dimensions it may not be trivial to solve these standard center problems for very large instances.

Theorem 4.16 has given the reformulation that we heavily used in the previous section and in order to avoid solving the fixed gate point center problem in n -dimensions we state another equivalent reformulation of the problem. The reformulation itself is not useful for solving the problem since it introduces an uncountable amount of constraints, but it is an auxiliary problem that we use to develop an approximation algorithm.

Lemma 4.28. *Let $CS := \partial B_{r_2}(0) \cap \mathbb{R}_-^n$ be the boundary of the sphere with radius r_2 in \mathbb{R}_-^n . Then the following two problems are equivalent.*

$$(P_{r_2}) \quad \begin{array}{ll} \min & r \\ \text{s.t.} & \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k \\ & \|x\|_2 + r_2 \leq r \\ & x \in \mathbb{R}_+^n \end{array}$$

$$(P_{CS}) \quad \begin{array}{ll} \min & r \\ \text{s.t.} & \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k \\ & \|x - A\|_2 \leq r \quad \forall A \in CS \\ & x \in \mathbb{R}_+^n \end{array}$$

Proof. Since the objective functions of both problems are identical, it suffices to show that the feasible sets coincide. Let (x', r') be feasible for (P_{r_2}) . Then for any $A \in CS$ we have that $\|x' - A\|_2 \leq \|x'\|_2 + \|A\|_2 = \|x'\|_2 + r_2 \leq r'$ where the latter inequality holds because x' is feasible. Thus, (x', r') is feasible for (P_{CS}) .

Let now x' be feasible for (P_{CS}) with radius r' . Define $y := \frac{-r_2}{\|x'\|_2} \cdot x'$. Then $\|y\|_2 = r_2$ and $y \in \mathbb{R}_-^n$, so $y \in CS$. Then it follows that $r' \geq \|x - y\|_2 = \|(1 + \frac{r_2}{\|x'\|_2}) \cdot x'\|_2 = \|x'\|_2 + r_2$ and (x', r') is feasible for (P_{r_2}) . \square

An important implication of the proof is the following.

Corollary 4.29. *Let $y^* := \frac{-r_2}{\|x^*\|_2} \cdot x^*$. Then the following problems are equivalent:*

$$(P_{y^*}) \quad \begin{array}{ll} \min & r \\ \text{s.t.} & \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k \\ & \|x - y^*\|_2 \leq r \\ & x \in \mathbb{R}_+^n \end{array}$$

$$(P_{CS}) \quad \begin{array}{ll} \min & r \\ \text{s.t.} & \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k \\ & \|x - A\|_2 \leq r \quad \forall A \in CS \\ & x \in \mathbb{R}_+^n \end{array}$$

Obviously we do not know y^* before we have found the optimal solution (x^*, r^*) . The idea is to generate specific points on CS that approximate y^* , thereby getting an approximate solution to (P_{CS}) . More precisely, we construct a finite subset D of CS such that the distance of some facility in D to y^* is as small as possible and then solve the relaxation

$$(P_D) \quad \min \quad r$$

$$\text{s.t.} \quad \|A_i - x\|_2 \leq r \quad \forall i = 1, \dots, k$$

$$\|x - A\|_2 \leq r \quad \forall A \in D$$

$$x \in \mathbb{R}_+^n$$

A simple way to get such a subset $D \subset CS$ is to construct a grid using Cartesian coordinates in \mathbb{R}^n . All points $A \in CS$ satisfy $\|A\|_2 = r_2$ so we consider

$$\Psi : [0, 2\pi) \times [0, \pi)^{n-2} \rightarrow B_{r_2}(0) \subset \mathbb{R}^n$$

$$\Psi(\varphi_1, \varphi_2, \dots, \varphi_{n-1}) = \begin{pmatrix} r_2 \cos(\varphi_1) \\ r_2 \sin(\varphi_1) \cos(\varphi_2) \\ r_2 \sin(\varphi_1) \sin(\varphi_2) \cos(\varphi_3) \\ \vdots \\ r_2 \sin(\varphi_1) \cdots \sin(\varphi_{n-2}) \cos(\varphi_{n-1}) \\ r_2 \sin(\varphi_1) \cdots \sin(\varphi_{n-1}) \end{pmatrix}.$$

By construction, $\Psi([0, \frac{3}{2}\pi] \times [\frac{\pi}{2}, \pi]^{n-1}) = CS$. Let $D_m \subset CS$ be defined as $D_m = \Psi(I_m)$, with I_m being a uniform segmentation of the intervals for the φ_i , i.e., $I_m = C_1^m \times \dots \times C_{n-1}^m$ with

$$C_1^m = \left\{ \frac{k\pi}{2m} + \pi : k = 0, \dots, m \right\},$$

$$C_i^m = \left\{ \frac{k\pi}{2m} + \frac{\pi}{2} : k = 0, \dots, m \right\}, \text{ for } i = 2, \dots, n-1.$$

If we use our relaxation (P_{D_m}) for a specific m , we actually get a *fully polynomial-time approximation scheme* when the dimension is fixed and an additional slight assumption holds. A fully polynomial-time approximation scheme is an ε approximation algorithm that has a polynomial running time when ε is considered a constant.

Theorem 4.30. *Let an instance $\{A_0, A_1, \dots, A_k\}$ be given, where $r_1 = \|A_1\|_2 = \max_{i=1, \dots, k} \|A_i\|_2$. Let $\varepsilon > 0$ and set $m = \left\lceil \frac{\pi r_2 \sqrt{nr_1}}{\varepsilon} \right\rceil$. Moreover, let (x^*, r^*) be the optimal solution to (P_{r_2}) . If $r_2 \geq \frac{1}{\sqrt{2}}$, then the solution (x_m^*, r_m^*) to (P_{D_m}) satisfies $(1 + \varepsilon)r^* \geq \max_{i=0, \dots, k} d(A_i, x_m^*)$ and is found in polynomial running time when ε is considered a constant.*

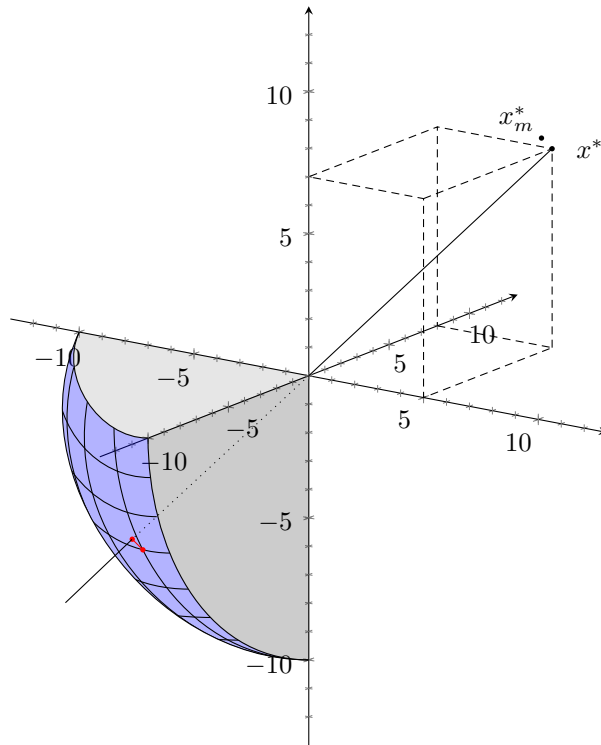


Figure 4.7: Illustration of the grid D_m on the surface of the scaled unit ball in \mathbb{R}^3 and the approximate solution x_m^* . The optimal solution of the center problem is x^* .

Proof. First we show $(1 + \varepsilon)r^* \geq \max_{i=0,\dots,k} d(A_i, x_m^*)$. Since x_m^* is the optimal solution to (P_{D_m}) we know that $\|x - A_i\|_2 \leq r_m^*$ for $i = 1, \dots, m$. Hence, we only need to calculate the distance $d(A_0, x_m^*) = \|x_m^*\|_2 + r_2$. For $y_m^* := \frac{-r_2}{\|x_m^*\|_2} \cdot x_m^*$ it holds that $\|x_m^* - y_m^*\|_2 = \|x_m^*\|_2 + r_2 = d(A_0, x_m^*)$ and then

$$\begin{aligned} \|x_m^* - y_m^*\|_2 &\leq \min_{z \in D_m} (\|x_m^* - z\|_2 + \|z - y_m^*\|_2) \\ &\leq \min_{z \in D_m} (r_m^* + \|z - y_m^*\|_2) = r_m^* + \min_{z \in D_m} \|z - y_m^*\|_2 \end{aligned}$$

Since $y_m^* \in CS$ we can bound the distance of y_m^* to one of the grid points. Let $y_m^* = \Psi(\tilde{\varphi}_1, \tilde{\varphi}_2, \dots, \tilde{\varphi}_{n-1})$. Then by definition of $D_m = \Psi(I_m)$ there exist $\varphi_i \in C_i^m$ such that $|\varphi_i - \tilde{\varphi}_i| \leq \frac{\pi}{4m}$ for $i = 1, \dots, n-1$. Then we can estimate

$$\begin{aligned} \|y_m^* - x\|_2 &= \|\Psi(\tilde{\varphi}_1, \dots, \tilde{\varphi}_{n-1}) - \Psi(\varphi_1, \varphi_2, \dots, \varphi_{n-1})\|_2 \\ &\leq 2r_2 \cdot \left\| \begin{pmatrix} \tilde{\varphi}_1 - \varphi_1 \\ \vdots \\ \tilde{\varphi}_{n-1} - \varphi_{n-1} \end{pmatrix} \right\|_2 \leq 2r_2 \cdot \left\| \begin{pmatrix} \frac{\pi}{4m} \\ \vdots \\ \frac{\pi}{4m} \end{pmatrix} \right\|_2 = \frac{\pi r_2 \sqrt{n}}{2m}, \end{aligned}$$

because Ψ is differentiable and its derivative is bounded by $2r_2$, so Ψ is Lipschitz continuous with constant $2r_2$ therewith yielding the last estimate.

Recall the definition of $m = \left\lceil \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon} \right\rceil$. Thus, $m \geq \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon}$, so

$$\|y_m^* - x\|_2 \leq \frac{\pi r_2 \sqrt{n}}{2m} \leq \frac{\varepsilon \pi r_2 \sqrt{n}}{2\pi r_1 r_2 \sqrt{n}} = \frac{\varepsilon}{2r_1}.$$

Thus, we can bound the distance of A_0 to x_m^* by

$$d(A_0, x_m^*) = \|x_m^* - y_m^*\|_2 \leq r_m^* + \min_{z \in D_m} \|z - y_m^*\|_2 \leq r_m^* + \frac{\varepsilon}{2r_1}$$

and using that $d(A_i, x_m^*) \leq r_m^*$ we get the same bound for the objective value of x_m^* ,

$$\max_{i=0,\dots,k} d(A_i, x_m^*) \leq r_m^* + \frac{\varepsilon}{2r_1}.$$

We now show that $(1 + \varepsilon)r^* \geq \max_{i=0,\dots,k} d(A_i, x_m^*)$. Using $r_1 \geq r_2$ and the assumption $r_2 \geq \frac{1}{\sqrt{2}}$, we have

$$r^* \geq \frac{r_1 + r_2}{2} \geq \frac{\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}}{2} = \frac{1}{\sqrt{2}} = \frac{1}{2 \cdot \frac{1}{\sqrt{2}}} \geq \frac{1}{2r_1},$$

where the first inequality is a lower bound on r^* derived by the radius of the smallest enclosing ball for A_0, A_1 . Moreover, recall that (P_{D_m}) is a relaxation of (P) , hence $r_m^* \leq r^*$. With these inequalities we get

$$\max_{i=0,\dots,k} d(A_i, x_m^*) \leq r_m^* + \frac{\varepsilon}{2r_1} \leq r^* + \frac{\varepsilon}{2r_1} \leq r^* + \varepsilon r^* = (1 + \varepsilon)r^*.$$

Now it only remains to show that this algorithm has a polynomial running time with respect to the input parameters k, r_1, r_2 , where the dimension n and ε are considered to be constants. For the standard center problem in dimension n , the running time of Megiddo's algorithm on an instance with k facilities is $\mathcal{O}((n+1)(n+1)!k)$. Since the dimension does not depend on k , this is linear with respect to k . For our instance $\{A_0, \dots, A_k\}$ we have the set of facilities $D_m \cup \{A_1, \dots, A_k\}$ when solving (P_{D_m}) . It holds that

$$|D_m| = (m+1)^{n-1} = \left\lceil \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon} \right\rceil^{n-1},$$

which is polynomial in r_1, r_2 and even independent of k . Thus, applying Megiddo's algorithm we receive a running time of

$$\mathcal{O} \left((n+1)(n+1)! \left(k + \left\lceil \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon} \right\rceil^{n-1} \right) \right),$$

which is polynomial in k, r_1, r_2 for fixed n and ε . The generation of the grid is also done in linear time with respect to $\left\lceil \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon} \right\rceil^{n-1}$, so we get

$$\mathcal{O} \left((n+1)(n+1)! \left(k + \left\lceil \frac{\pi r_2 \sqrt{n} r_1}{\varepsilon} \right\rceil^{n-1} \right) \right)$$

as the polynomial overall running time for our approximation. \square

Even though it is theoretically possible to approximate the problem up to any fixed precision, there are two drawbacks. The first one is a general flaw of any sort of fully polynomial-time approximation schemes: The running time is polynomial in $\frac{1}{\varepsilon}$, which results in bad practical computation times when a very small tolerance $\varepsilon > 0$ is chosen. The second problem is that the running time depends on r_1 and r_2 . There is no a priori bound on these radii beforehand and when working in higher dimensions n , $(r_1 r_2)^{n-1}$ might also be problematically big. Note that we use Megiddo's algorithm and the argument of a fixed dimension to get a polynomial running time for the algorithm. There are several algorithms (approximations or exact solvers) for the standard center problem that claim to perform well in practice even for high dimensions. However, an exact solver for the problem that has a guaranteed polynomial running time taking the dimension into account is not available to our knowledge. Thus, it makes sense to choose specific algorithms to solve the relaxation P_{D_m} depending on the instance (and dimension of the space) that one has given.

4.4.3 Fréchet Problem

In this section we consider Fréchet problem (G_F) of the fixed gate point problem:

$$(G_F) \quad \min \sum_{A \in \mathcal{A}} d(x, A) \\ \text{s.t. } x \in \mathbb{R}_+^{n-2} \cup \mathbb{R}_-^{n-2}.$$

In the following theorem we provide a necessary and sufficient condition for 0 to be the optimal solution to (G_F) . Note that the condition that we state in the theorem slightly differs from the one in Theorem 4.10. Moreover we give a closed formula for the unique optimal solution, in case that it is not 0.

Theorem 4.31. *0 is the unique optimal solution to (G_F) if and only if*

$$\left\| \sum_{A \in \mathcal{A}_i} A \right\|_2 \leq \sum_{A \in \mathcal{A}_j} \|A\|_2 \quad \text{for } i \neq j \in \{1, 2\}. \quad (4.14)$$

If 0 is not the optimal solution, let $i \in \{1, 2\}$ be the index such that

$$\left\| \sum_{A \in \mathcal{A}_i} A \right\|_2 > \sum_{A \in \mathcal{A}_j} \|A\|_2 \quad (4.15)$$

where $j \neq i \in \{1, 2\}$. Then

$$x^* = \frac{1}{M} \cdot \left(1 - \frac{\sum_{A \in \mathcal{A}_j} \|A\|}{\left\| \sum_{A \in \mathcal{A}_i} A \right\|} \right) \cdot \sum_{A \in \mathcal{A}_i} A.$$

is the optimal solution to (G_F) .

Proof. Before proving the statements we collect several properties of the objective function. Consider a point x in \mathbb{R}_+^n . Then the objective function of (G_F) is given by

$$f(x) = \sum_{A \in \mathcal{A}_1} \|x - A\|_2^2 + \sum_{A \in \mathcal{A}_2} (\|x\|_2 + \|A\|_2)^2.$$

Analogously, for a point x in \mathbb{R}_-^n , we get

$$f(x) = \sum_{A \in \mathcal{A}_2} \|x - A\|_2^2 + \sum_{A \in \mathcal{A}_1} (\|x\|_2 + \|A\|_2)^2.$$

In both cases we have that, as a sum of convex functions, f is convex. Moreover, $\|x - A\|_2^2$ is differentiable on \mathbb{R}^n and $(\|x\|_2 + \|A\|_2)^2 = \|x\|_2^2 + 2\|A\|_2\|x\|_2 + \|A\|_2^2$ is differentiable everywhere except at 0.

To find an optimum of f restricted to an orthant, it suffices to find a local optimum due to its convexity, which we calculate in what follows. We look for x^* such that $\nabla f(x^*) = 0$ holds and we already know that the gradient exists for $x^* \neq 0$. Now, for $x \in \mathbb{R}_+^n \setminus \{0\}$, $\nabla f(x)$ is given by

$$\nabla f(x) = 2 \cdot \left(\sum_{A \in \mathcal{A}_1} (x - A) + \sum_{A \in \mathcal{A}_2} \frac{x}{\|x\|} \cdot (\|A\| + \|x\|) \right).$$

Setting $\nabla f(x) = 0$ and rearranging yields

$$\begin{aligned} \nabla f(x) = 0 &\Leftrightarrow \sum_{A \in \mathcal{A}_1} x + \sum_{A \in \mathcal{A}_2} \left(x + \frac{\|A\|}{\|x\|} \cdot x \right) = \sum_{A \in \mathcal{A}_1} A \\ &\Leftrightarrow \left(M + \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\|x\|} \right) \cdot x = \sum_{A \in \mathcal{A}_1} A, \end{aligned} \quad (4.16)$$

showing that, if $\nabla f(x) = 0$, there exists a $\lambda > 0$ such that $\lambda x = \sum_{A \in \mathcal{A}_1} A$ holds. Considering this equality and scaling the lengths of both vectors to 1 we receive

$$\frac{x}{\|x\|} = \frac{\lambda \cdot x}{\lambda \|x\|} = \frac{\sum_{A \in \mathcal{A}_1} A}{\|\sum_{A \in \mathcal{A}_1} A\|}.$$

Rearranging terms of (4.16) and then plugging in, we get that

$$\begin{aligned} \nabla f(x) = 0 &\Leftrightarrow \left(M + \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\|x\|} \right) \cdot x = \sum_{A \in \mathcal{A}_1} A \\ &\Leftrightarrow M \cdot x + \sum_{A \in \mathcal{A}_2} \|A\| \cdot \left(\frac{\sum_{A \in \mathcal{A}_1} A}{\|\sum_{A \in \mathcal{A}_1} A\|} \right) = \sum_{A \in \mathcal{A}_1} A \\ &\Leftrightarrow M \cdot x = \sum_{A \in \mathcal{A}_1} A - \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\|\sum_{A \in \mathcal{A}_1} A\|} \cdot \sum_{A \in \mathcal{A}_1} A \\ &\Leftrightarrow x = \frac{1}{M} \cdot \left(1 - \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\|\sum_{A \in \mathcal{A}_1} A\|} \right) \cdot \sum_{A \in \mathcal{A}_1} A. \end{aligned}$$

This formula can analogously be established for $x \in \mathbb{R}_-^n$ by simply exchanging \mathcal{A}_1 with \mathcal{A}_2 and vice versa.

With this formula for a local minimizer at hand we prove the theorem. First, assume 0 is not optimal. As (G_F) is obtained as an equivalent formulation for (P_F) in a special case, Theorem 3.5 implies that we have a unique optimal solution x^* . Then we know that f is differentiable at the optimal solution $x^* \neq 0$. W.l.o.g. consider

the case $x^* \in \mathbb{R}_+^n$ (otherwise use the formula and exchange \mathcal{A}_1 with \mathcal{A}_2). As $x^* \neq 0$ it satisfies the local optimality condition,

$$x^* = \frac{1}{M} \cdot \left(1 - \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\left\| \sum_{A \in \mathcal{A}_1} A \right\|} \right) \cdot \sum_{A \in \mathcal{A}_1} A.$$

That $\sum_{A \in \mathcal{A}_1} A \in \mathbb{R}_+^n$ follows from $\mathcal{A}_1 \subset \mathbb{R}_+^n$. As $x^* \in \mathbb{R}_+^n$, this implies that the scalar

$$\frac{1}{M} \cdot \left(1 - \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\left\| \sum_{A \in \mathcal{A}_1} A \right\|} \right) > 0.$$

Thus,

$$\sum_{A \in \mathcal{A}_2} \|A\| < \left\| \sum_{A \in \mathcal{A}_1} A \right\|.$$

so \mathcal{A}_1 does not satisfy condition (4.14). For the case $x^* \in \mathbb{R}_-^n$ the argumentation is the same and we get that \mathcal{A}_2 does not satisfy condition (4.14). Thus, when 0 is not optimal it follows that one of the sets \mathcal{A}_i does not satisfy condition (4.14).

Now we need to prove the other direction. Assume that one of the sets \mathcal{A}_i does not satisfy condition (4.14). W.l.o.g. let this be \mathcal{A}_1 . When \mathcal{A}_1 does not satisfy (4.14), it satisfies (4.15) and it follows that \mathcal{A}_2 cannot satisfy (4.15). In particular, \mathcal{A}_2 cannot satisfy the condition from Theorem 4.10, so the optimal solution to (G_F) is not contained in $\mathbb{R}_-^n \setminus \{0\}$. As \mathcal{A}_1 satisfies (4.14) it follows again that the scalar on the right hand side of the formula

$$x^* = \frac{1}{M} \cdot \left(1 - \frac{\sum_{A \in \mathcal{A}_2} \|A\|}{\left\| \sum_{A \in \mathcal{A}_1} A \right\|} \right) \cdot \sum_{A \in \mathcal{A}_1} A$$

is positive, hence we have a local minimum at $0 \neq x^* \in \mathbb{R}_+^n$. As $\mathbb{R}_-^n \setminus \{0\}$ does not contain a minimizer for the problem, the minimizer over \mathbb{R}_+^n yields the optimal solution. Due to convexity the local optimality of x^* implies that x^* is a minimizer over the convex set \mathbb{R}_+^n and due to uniqueness of the minimizer 0 cannot be optimal. \square

With this theorem at hand, the fixed gate point problem is easily solved: Simply check if one of the two \mathcal{A}_i satisfies condition (4.14). If this is not the case, 0 is the unique optimal solution and otherwise the closed formula yields the optimal solution in the respective orthant.

Altogether we have now shown how to solve the fixed gate point center problems (G_{Med}) , (G_{max}) , (G_F) , which solve problems in tree space where the data is contained

in several pairwise incompatible orthants. This is another example that a nice reformulation of the tree space problems enables algorithms for Facility Location in \mathbb{R}^n to solve specific location problems in tree space.

5 Balance Point Algorithms for the Median Problem in Tree Space

The main problem we address in this thesis is the median problem (P_{Med}), defined in Chapter 3. We have already seen some solutions to the median problem in specific cases in Chapter 4, but here we tackle it in general: Given $\mathbf{T} = \{T_1, \dots, T_M\}$, we want to find $T^* \in \mathcal{T}_n$ such that T^* minimizes $\sum_{i=1}^M d(T_i, X)$ amongst all $X \in \mathcal{T}_n$. Until now we have seen that in some cases we may reduce the problem to some sort of Euclidean median problem variant, but in general this is tough to do. Even though one can use the canonical embedding φ into \mathbb{R}_+^N and use the intrinsic metric of $\|\cdot\|_2$ on $\text{im}(\varphi)$, this does not simplify the problem, as there are no known polynomial (in n) algorithms to solve this problem. Hence, we have to find a new ansatz to tackle the problem, as there is no straightforward reformulation to a solvable Euclidean location problem.

An existing algorithmic approach to find the median in *Hadamard spaces* has been developed in [Bac14a]. Hadamard spaces are complete spaces of global non-positive curvature, so these methods may in particular be applied to the tree space. For the median problems on Hadamard spaces Bacak shows that his proximal point algorithm (PPA) converges to a median for any given starting point. Nonetheless, in some cases convergence is quite slow. We define the PPA in Section 5.2.4 and investigate its convergence behavior in the experiment sections, Section 5.2.4 and Chapter 6 and compare the PPA to our method, we develop in the following. That the convergence rate is rather slow is not too surprising since it is the median analogue to Sturm's algorithm [Stu03]. Sturm's algorithm is an iterative scheme converging to the Fréchet mean in spaces of global non-positive curvature. Even though it converges to the optimal solution and it is known that its convergence rate is sublinear [MOP15] and hence not satisfying.

Our approach is to develop a heuristic approach that strongly uses the local Euclidean structure of the tree space to try to improve convergence behavior in comparison to Bacak's PPA. Moreover it is desirable to find an algorithm that allows for a better interpretation of the determined solution, as the PPA only yields approximate solutions and we have no measure to find out how far away this solution is from an actual median.

It seems unintuitive to develop a heuristic, when there exists an approximation

algorithm that converges to the optimal solution, but we justify why the method has practical relevance due to its advantageous use of the tree space structures.

Since the chapter is rather extensive we briefly sketch its structure: We develop a heuristic for the median problem (P_{Med}), which is based on a local improvement strategy called the Balance Point Algorithm (BPA) which operates on orthants. Since there is an exponential amount of orthants, we start by developing lower and upper bounds in Section 5.1 in order to rule out several orthants beforehand. After we have investigated the bounds we state the *Global Balance Point Heuristic* in Section 5.2, including the definition of the BPA as well as experiments. After this we investigate the convergence of the BPA in Section 5.3 before depicting some cases that justify and explain the importance of the assumptions needed for the convergence of the BPA in Section 5.4.

5.1 Bounds for the Median Problem

Bounds for the median problem in \mathcal{T}_n have already been investigated by [Coh17] in a bachelor's thesis. The lower bounds z_{L1}, z_{L2} and upper bounds z_{U1}, z_{U2} we define in the following, have been introduced in [Coh17]. The other bounds we develop in the following improve these known bounds: z_{L3} is always better than z_{L1} and z_{L2} , see Theorem 5.2, and z_{U3}, z_{U4} provide practically improved results, as can be seen in the experiments of Section 5.1.

Lower Bounds

We start by calculating lower bounds. Recall the definition of complete incompatibility, Definition 4.5. With this we introduce the notation

$$\mathbf{T}_{\mathcal{O}}^c := \{T \in \mathbf{T} : \text{Split}(T) \text{ completely incompatible with } \text{Split}(\mathcal{O})\}. \quad (5.1)$$

Now, for a fixed orthant \mathcal{O} we show that the optimal objective values of the optimization problems

$$\begin{aligned} \text{(L1)} \quad z_{L1} &:= \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} \|t - x\|_2 \\ \text{(L2)} \quad z_{L2} &:= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 \\ \text{(L3)} \quad z_{L3} &:= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 \end{aligned}$$

yield lower bounds on the objective $\min_{X \in \mathcal{O}} \sum_{i=1}^M d(X, T_i)$. We refer to the respective optimization problems by **L1, L2, L3**.

Theorem 5.1. z_{L1}, z_{L2}, z_{L3} are lower bounds for $\min_{X \in \mathcal{O}} \sum_{i=1}^M d(X, T_i)$ for any $\mathcal{O} \subset \mathcal{T}_n$.

Proof. z_{L1} For each $T \in \mathbf{T}$ it holds that $\|t - x\|_2 \leq d(T, X)$, so it also holds for the sum, i.e., the objective value z_{L1} satisfies $z_{L1} \leq \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} d(T, X)$.

z_{L2} For any $X \in \mathcal{O}$ it holds that

$$\begin{aligned} & \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 \\ &= \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} (\|t\|_2 + \|x\|_2) \\ &= \sum_{T \in \mathbf{T} \cap \mathcal{O}} d(T, X) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} d(T, X) \\ &\leq \sum_{T \in \mathbf{T} \cap \mathcal{O}} d(T, X) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} d(T, X) + \sum_{T \in \mathbf{T} \setminus (\mathcal{O} \cup \mathbf{T}_{\mathcal{O}}^c)} d(T, X) = \sum_{T \in \mathbf{T}} d(T, X) \end{aligned}$$

so it also holds for the minimum, $z_{L2} \leq \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} d(T, X)$.

z_{L3} Similarly as for z_{L2} we have that for any $X \in \mathcal{O}$ it holds that

$$\begin{aligned} & \sum_{T \in \mathbf{T} \setminus \mathbf{T}_{\mathcal{O}}^c} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 \\ &= \sum_{T \in \mathbf{T} \setminus \mathbf{T}_{\mathcal{O}}^c} \|t - x\|_2 + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} (\|t\|_2 + \|x\|_2) \\ &\leq \sum_{T \in \mathbf{T} \cap \mathcal{O}} d(T, X) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} d(T, X) = \sum_{T \in \mathbf{T}} d(T, X), \end{aligned}$$

hence, $z_{L3} \leq \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} d(T, X)$. □

Theorem 5.2. z_{L3} is the strongest of the three bounds.

Proof. First of all, note that any tree $T \in \mathbf{T} \cap \mathcal{O}$ is not in $\mathbf{T}_{\mathcal{O}}^c$. Thus, $\mathbf{T} \cap \mathcal{O} \subset \mathbf{T} \setminus \mathbf{T}_{\mathcal{O}}^c$

which implies that

$$\begin{aligned}
z_{L2} &= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 \\
&\leq \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 = z_{L3}.
\end{aligned}$$

Now, consider **L1** and **L3**. We use $\|t - x\|_2 \leq d(T, X)$ once more to get

$$\begin{aligned}
z_{L1} &= \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} \|t - x\|_2 \\
&= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + \sum_{T \in T_{\mathcal{O}}^c} \|t - x\|_2 \right) \\
&\leq \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + \sum_{T \in T_{\mathcal{O}}^c} d(T, X) \right) \\
&= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + \sum_{T \in T_{\mathcal{O}}^c} (\|x\|_2 + \|t\|_2) \right) \\
&= \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2 = z_{L3}.
\end{aligned}$$

□

The last thing we have to take care of is actually computing the bounds in order to implement them in our heuristic. Therefore we need to solve median problems in the tree space restricted to the orthant \mathcal{O} . To be able to solve these problems and compute the bounds, we will transform the problem to \mathbb{R}^N and then reduce it to a lower dimension.

The following theorem states an equivalent formulation that

Theorem 5.3. *Let $\mathbf{T} = \{T_1, \dots, T_M\}$ and \mathcal{O} be given. W.l.o.g. assume that for*

$X \in \mathcal{O}$ and for $T \in \mathcal{T}_n$ arbitrary we have the representations

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} =: \begin{pmatrix} x' \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^N, x' \in \mathbb{R}^{n-2}, \quad t =: \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \mathbb{R}^N, t' \in \mathbb{R}^{n-2}, t'' \in \mathbb{R}^{N-n+2}.$$

Then for all $T \in \mathcal{T}_n$ it holds that

$$\|t - x\|_2 = \left(\|t' - x'\|_2^2 + \|t''\|_2^2 \right)^{\frac{1}{2}}.$$

Moreover,

$$\min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} \|t - x\|_2 \Leftrightarrow \min_{x' \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \left(\|t'_i - x'\|_2^2 + \|t''_i\|_2^2 \right)^{\frac{1}{2}}$$

are equivalent optimization problems.

Proof. Choose $X \in \mathcal{O}$ and $T \in \mathcal{T}_n$ arbitrary and define $x'' := 0 \in \mathbb{R}^{N-(n-2)}$. Then we have that

$$\|t - x\|_2 = \left(\|t'_i - x'\|_2^2 + \|t''_i - x''\|_2^2 \right)^{\frac{1}{2}} = \left(\|t'_i - x'\|_2^2 + \|t''_i\|_2^2 \right)^{\frac{1}{2}}$$

by the Pythagorean theorem in \mathbb{R}^N . Hence, finding $\min_{X \in \mathcal{O}} \sum_{i=1}^M \|t_i - x\|_2$ is equivalent to finding $\min_{x' \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \left(\|t'_i - x'\|_2^2 + \|t''_i\|_2^2 \right)^{\frac{1}{2}}$. \square

Note that $\|t''_i\|_2^2$ does not depend on x' , and that the problem resembles a median problem in \mathbb{R}^{n-2} . If $t''_i = 0$ for all i , it even is a standard median problem. So we actually have a generalized version of a median problem. Thus, it is natural to try solve the problem via an adaptation of an algorithm for the standard median problem. In the following we adapt the so-called *Weiszfeld* algorithm. The Weiszfeld algorithm, introduced by [Wei37], is an iterative procedure and is often described as 'long-step' gradient descent method, where the step size is calculated implicitly in the formula for the next iteration point.

Our goal is to calculate the bounds z_{L1}, z_{L2}, z_{L3} , i.e., to solve the respective optimization problems **L1**, **L2**, **L3**. Note that **L2**, **L3** are weighted median problems, where all facilities T in the first sum have weight 1 and the facility 0 has weight $|T_{\mathcal{O}}^c|$. Thus, when we reformulate **L1**, **L2**, **L3** using Theorem 5.3 we get an optimization problem with the following structure

$$(MF) \quad \min \quad \sum_{i=1}^M w_i (\|A_i - x\|_2^2 + c_i)^{\frac{1}{2}} \\ \text{s.t.} \quad x \in \mathbb{R}^n,$$

where facilities $\mathcal{A} = \{A_1, \dots, A_M\} \subset \mathbb{R}^n$ are given with weights $w_i \geq 0$ and some 'fixed costs' $c_i \geq 0$. So when we are able to solve (MF) , then we may calculate all lower bounds.

We now follow the structure in [Kuh73]. The results are the same as in Kuhn's proof of convergence for the Weiszfeld algorithm for the Euclidean median problem, but we adapt all results and proofs to this generalized version of the median problem. The only difference are the terms

$$\begin{aligned} \eta_i(x) &:= (\|A_i - x'\|_2^2 + c_i)^{\frac{1}{2}} = ((A_{i1} - x_1)^2 + \dots + (A_{in} - x_n)^2 + c_i)^{\frac{1}{2}} \\ &\geq ((A_{i1} - x_1)^2 + \dots + (A_{in} - x_n)^2)^{\frac{1}{2}} = \|A_i - x\|_2, \end{aligned}$$

for $i = 1, \dots, M$ that depend on c_i .

We denote the objective function of (MF) by

$$g(x) = \sum_{i=1}^M w_i (\|A_i - x\|_2^2 + c_i)^{\frac{1}{2}}.$$

With the η_i terms we may rewrite g as $g(x) = \sum_{i=1}^M w_i \eta_i(x)$. From the first result on, our set of facilities is divided into two subsets that behave differently; let $\mathcal{A}^+ = \{A^i : i = 1, \dots, M \text{ s.t. } c_i > 0\}$ and $\mathcal{A}^0 = \{A^i : i = 1, \dots, M \text{ s.t. } c_i = 0\}$.

Lemma 5.4. *The objective function g is convex on \mathbb{R}^n and differentiable on $\mathbb{R}^n \setminus \mathcal{A}^0$, with*

$$\nabla g(x) = \sum_{i=1}^M \frac{w_i(x - A_i)}{\eta_i(x)} \quad \text{where existent.}$$

Proof. We start by showing convexity. g is the weighted sum of the η_i . As the weights are positive, convexity boils down to convexity of the η_i . To show that these are convex, we reformulate the term to

$$(\|A_i - x'\|_2^2 + c_i)^{\frac{1}{2}} = \left(\left\| \begin{pmatrix} A_1 - x_1 \\ \vdots \\ A_n - x_n \\ c^{\frac{1}{2}} \end{pmatrix} \right\|_2^2 \right)^{\frac{1}{2}} = \left\| \begin{pmatrix} A_1 - x_1 \\ \vdots \\ A_n - x_n \\ c^{\frac{1}{2}} \end{pmatrix} \right\|_2.$$

The right hand side is convex if and only if $h(x) = \left\| \begin{pmatrix} x \\ c \end{pmatrix} \right\|_2$ is convex, as it is only a translation. Choose $x, y \in \mathbb{R}^n$ and $\lambda \in (0, 1)$, then

$$\begin{aligned} h(\lambda x + (1 - \lambda)y) &= \left\| \begin{pmatrix} \lambda x + (1 - \lambda)y \\ c \end{pmatrix} \right\|_2 = \left\| \lambda \begin{pmatrix} x \\ c \end{pmatrix} + (1 - \lambda) \begin{pmatrix} y \\ c \end{pmatrix} \right\|_2 \\ &\leq \lambda \left\| \begin{pmatrix} x \\ c \end{pmatrix} \right\|_2 + (1 - \lambda) \left\| \begin{pmatrix} y \\ c \end{pmatrix} \right\|_2 = \lambda h(x) + (1 - \lambda)h(y) \end{aligned}$$

due to the triangle inequality, so h is convex. That implies convexity of η_i which in turn implies convexity of g .

As a composition of differentiable functions, we get that g is differentiable on $\mathbb{R}^n \setminus \mathcal{A}$, i.e., all points where $\|\cdot - A_i\|_2$ is differentiable. For the facilities in \mathcal{A}^+ we check differentiability by investigating the partial derivatives

$$\frac{\partial g}{\partial x_i}(x) = \sum_{j=1}^M \frac{w_j(x_i - A_{ji})}{\eta_j(x)}.$$

Take $x = A_k \in \mathcal{A}^+$. Then $\frac{\partial g}{\partial x_i}(x) = \sum_{j=1}^M \frac{w_j(x_i - A_{ji})}{\eta_j(x)}$. Since $c_k > 0$, and hence $\eta_k(x) > 0$, the partial derivatives are continuous functions for all i at $x = A_k$. Thus, g is totally differentiable at $A_k \in \mathcal{A}^+$ with gradient

$$\nabla g(x) = \sum_{i=1}^M \frac{w_i(x - A_i)}{\eta_i(x)}$$

as for all other $x \notin \mathcal{A}^0$. □

In order to apply a gradient descent method, we'd like to rule out the case that some facility in \mathcal{A}^0 is optimal, so that we can use the sufficient optimality criterion $\nabla g(x) = 0$ (since g is convex). The following lemma enables us to check optimality of the facilities \mathcal{A}^0 a priori.

Lemma 5.5. $A_k \in \mathcal{A}^0$ is an optimal solution to (MF) if and only if

$$\text{Test}_k := \left\| \sum_{\substack{i=1 \\ i \neq k}}^M \frac{w_i(A_k - A_i)}{\|\eta_i(A_k)\|_2} \right\|_2 \leq w_k.$$

Proof. For $A_k \in \mathcal{A}^0$, define $\tilde{g}_k(x) = \sum_{i \neq k} w_i \eta_i(x)$ and $g_k(x) = w_k \eta_k(x) = w_k \|x - A_k\|_2$, yielding $g(x) = \tilde{g}_k(x) + g_k(x)$. Now, set

$$t = - \sum_{\substack{i=1 \\ i \neq k}}^M \frac{w_i(A_k - A_i)}{\|\eta_i(A_k)\|_2} = -\nabla \tilde{g}_k(A_k),$$

the direction of steepest descent for g at A_k . For $g_k(A_k)$, consider all directions t' with $\|t'\|_2 = \|t\|_2$. Then for all $\varepsilon > 0$:

$$g_k(A_k + \varepsilon t') = w_k \|\varepsilon t'\|_2 = w_k \|\varepsilon t\|_2 = g_k(A_k + \varepsilon t),$$

so t belongs to the directions of steepest descent for g_k , so together, t is the direction of steepest descent for g at A_k . Since g is convex, local optimality is sufficient for global optimality, thus

$$A_k \text{ is optimal} \Leftrightarrow g(A_k + \varepsilon t) \geq g(A_k) \forall \varepsilon > 0 \quad (5.2)$$

We want to find an equivalent characterization for the right hand side, therefore we apply a Taylor-expansion of \tilde{g}_k at A_k in direction εt :

$$\tilde{g}_k(A_k + \varepsilon t) = \tilde{g}_k(A_k) + \varepsilon \nabla \tilde{g}_k(A_k)^t t + \mathcal{O}(\varepsilon^2).$$

Since $\nabla \tilde{g}_k(A_k) = -t$, the second term equals $-\varepsilon \|t\|_2^2$. Therewith we calculate

$$\begin{aligned} g(A_k + \varepsilon t) - g(A_k) &= \tilde{g}_k(A_k + \varepsilon t) - \tilde{g}_k(A_k) + g_k(A_k + \varepsilon t) - g_k(A_k) \\ &\Leftrightarrow g(A_k + \varepsilon t) - g(A_k) = -\varepsilon \|t\|_2^2 + w_k \|\varepsilon t\|_2 + \mathcal{O}(\varepsilon^2) \\ &\Rightarrow g(A_k + \varepsilon t) \geq g(A_k) \Leftrightarrow -\varepsilon \|t\|_2^2 + w_k \varepsilon \|t\|_2 \geq 0 \Leftrightarrow w_k \geq \|t\|_2 \end{aligned}$$

which holds for all $\varepsilon > 0$ sufficiently small. Together with (5.2) we conclude that A_k is optimal if and only if $w_k \geq \|t\|_2$. \square

So, first we can check if there exists an optimal solution $A_k \in \mathcal{A}^0$ and whenever this is not the case, we know that for any optimal solution x^* it holds that $\nabla g(x^*) = 0$. Then we rearrange terms:

$$\begin{aligned} \nabla g(x) = 0 &\Leftrightarrow \sum_{i=1}^M \frac{w_i(x - A_i)}{\eta_i(x)} = 0 \\ &\Leftrightarrow x \cdot \sum_{i=1}^M \frac{w_i}{\eta_i(x)} = \sum_{i=1}^M \frac{w_i A_i}{\eta_i(x)} \\ &\Leftrightarrow x = \frac{\sum_{i=1}^M \frac{w_i A_i}{\eta_i(x)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x)}} =: \frac{Q(x)}{P(x)} \end{aligned}$$

The idea of Andrew Vázsonyi, also known as Endre Weiszfeld, was to use this equation as iterative procedure, and set $x_{i+1} = \frac{Q(x_i)}{P(x_i)}$, for all x_i , where the objective is differentiable. This yields the mapping

$$T(x) := \begin{cases} \frac{Q(x)}{P(x)} & \text{if } x \notin \mathcal{A}^0, \\ A_i & \text{if } x = A_i \in \mathcal{A}^0. \end{cases}$$

This already describes the iterative algorithm, that we call *adapted Weiszfeld algorithm*, see Algorithm 3.

Data: $x_0 \in \mathbb{R}^n \setminus \mathcal{A}$
for $l = 1, 2, \dots$
do
 | $x_l = T(x_{l-1})$
end

Algorithm 3: The adapted Weiszfeld algorithm.

We now investigate properties of optimal solutions and $T(x)$ in more detail:

Lemma 5.6.

- i) If x is an optimal solution to (MF) , then $T(x) = x$.
- ii) If $x \notin \mathcal{A}^0$ and $T(x) = x$, then x is an optimal solution to (MF) .
- iii) Let x^* be an optimal solution to (MF) , then $x^* \in \text{conv}(\mathcal{A})$.
- iv) For $x \notin A_0$, $T(x) = x - h(x)\nabla g(x)$ with

$$h(x) = \frac{\prod_{i=1}^M \eta_i(x)}{\sum_{k=1}^M w_k \prod_{i \neq k} \eta_i(x)}.$$

Proof. **i)** and **ii)** follow directly by the definition of $T(x) = \frac{Q(x)}{P(x)}$ and their derivation.
iii): If $x^* = A_k$ for $A_k \in \mathcal{A}$ then $x^* \in \text{conv}(\mathcal{A})$. Otherwise,

$$x^* = T(x^*) = \frac{\sum_{i=1}^M \frac{w_i A_i}{\eta_i(x^*)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x^*)}},$$

so x^* is a convex combination of the A_i , because the scalars are

$$\frac{\frac{w_j}{\eta_j(x^*)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x^*)}} \geq 0 \quad \text{and} \quad \sum_{j=1}^M \frac{\frac{w_j}{\eta_j(x^*)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x^*)}} = 1.$$

iv): Note first, that $h(x) \geq 0$ and that $\nabla g(x) = xP(x) - Q(x)$. For $x \notin \mathcal{A}^0$ (we need $\eta_i(x) > 0$ for all i)

$$\begin{aligned} h(x)P(x) &= \sum_{i=1}^M \frac{w_i h(x)}{\eta_i(x)} = \sum_{i=1}^M \frac{w_i \prod_{j=1}^M \eta_j(x)}{\eta_i(x) \sum_{k=1}^M w_k \prod_{j \neq k} \eta_j(x)} \\ &= \frac{\sum_{i=1}^M w_i \prod_{j \neq i} \eta_j(x)}{\sum_{k=1}^M w_k \prod_{j \neq k} \eta_j(x)} = 1, \end{aligned}$$

Since $P(x) > 0$ this also implies $h(x) = \frac{1}{P(x)}$, and plugging everything in we get

$$T(x) = \frac{Q(x)}{P(x)} = x - x + \frac{Q(x)}{P(x)} = x - h(x)P(x)x + h(x)Q(x) = x - h(x)\nabla g(x).$$

□

Lemma 5.7. *If $T(x) \neq x$, then $g(T(x)) < g(x)$.*

Proof. $T(x) \neq x$ implies that $x \notin \mathcal{A}^0$, so

$$T(x) = \frac{\sum_{i=1}^M \frac{w_i A_i}{\eta_i(x)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x)}},$$

and in Lemma 5.6 **iii)** we have established that this is a convex combination of the $A_i \in \mathcal{A}$. This implies that $T(x)$ is the unique minimizer of $\varphi(y) = \sum_{i=1}^M \frac{w_i \eta_i(y)^2}{\eta_i(x)}$. This can be seen by rewriting $\varphi(y)$:

$$\varphi(y) = \sum_{i=1}^M \frac{w_i}{\eta_i(x)} \|y - A_i\|_2^2 + \sum_{i=1}^M \frac{w_i}{\eta_i(x)} c_i,$$

where $T(x)$ is the unique minimizer of the weighted sum of squares and the second sum is independent of y , so $T(x)$ is the unique minimizer of $\varphi(y)$. This implies

$$\varphi(T(x)) < \varphi(x) = \sum_{i=1}^M \frac{w_i}{\eta_i(x)} \eta_i(x)^2 = g(x)$$

With this we estimate

$$\begin{aligned} g(x) &> \varphi(T(x)) = \sum_{i=1}^M \frac{w_i}{\eta_i(x)} (\eta_i(T(x)) - \eta_i(x) + \eta_i(x))^2 \\ &= \sum_{i=1}^M \frac{w_i}{\eta_i(x)} (\eta_i(x)^2 + 2\eta_i(x)(\eta_i(T(x)) - \eta_i(x)) + (\eta_i(T(x)) - \eta_i(x))^2) \\ &= 2g(T(x)) - g(x) + \sum_{i=1}^M \frac{w_i}{\eta_i(x)} (\eta_i(T(x)) - \eta_i(x))^2 \\ &\Rightarrow 2g(x) > 2g(T(x)) + \sum_{i=1}^M \frac{w_i}{\eta_i(x)} (\eta_i(T(x)) - \eta_i(x))^2 \\ &\Rightarrow g(x) > g(T(x)), \end{aligned}$$

where the last implication holds since $\sum_{i=1}^M \frac{w_i}{\eta_i(x)} (\eta_i(T(x)) - \eta_i(x))^2 \geq 0$. □

The next two lemmata are slightly technical but are crucial to ensure convergence of the iterative procedure to the optimal solution rather than converging to some non-optimal facility $A_k \in \mathcal{A}^0$.

The first lemma states that when the iteration gets close to a non-optimal facility A_k , but does not *hit* it, then it is repelled and leaves a certain neighborhood of this facility after a finite amount of steps.

Lemma 5.8. *Let A_k be not optimal. Then there exists some $\delta > 0$ such that the following holds: If $x \in B_\delta(A_k)$, $x \neq A_k$, then there exists some $s \in \mathbb{N}$ satisfying $\|T^s(x) - A_k\|_2 > \delta$ and $\|T^{s-1}(x) - A_k\|_2 \leq \delta$.*

Proof.

$$\begin{aligned} T(x) - A_k &= x - h(x)\nabla g(x) - A_k = h(x) \sum_{i=1}^M \frac{w_i}{\eta_i(x)} (A_i - x) - (A_k - x) \\ &= h(x) \sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - x) + \left(\frac{h(x)w_k}{\eta_k(x)} - 1 \right) (A_k - x) \end{aligned}$$

Since A_k is not optimal, Lemma 5.5 states that

$$\left\| \sum_{i \neq k} \frac{w_i}{\eta_i(A_k)} (A_i - A_k) \right\|_2 > w_k,$$

so there exist $\delta' > 0, \varepsilon > 0$ such that

$$\left\| \sum_{i \neq k} \frac{w_i}{\eta_i(A_k)} (A_i - x) \right\|_2 \geq (1 + 2\varepsilon)w_k \quad \text{for all } x \text{ with } \|A_k - x\|_2 \leq \delta'. \quad (5.3)$$

Moreover,

$$\frac{h(x)w_k}{\eta_k(x)} = \frac{w_k \prod_{i \neq k} \eta_i(x)}{\sum_{l=1}^M w_l \prod_{i \neq l} \eta_i(x)} \rightarrow 1 \quad \text{as } x \rightarrow A_k, \text{ since } \prod_{i \neq l} \eta_i(x) \rightarrow 0 \text{ for all } l \neq k$$

because the product then contains $\eta_k(x) = \|x - A_k\|_2 \rightarrow 0$ when $x \rightarrow A_k$, and all other terms are bounded. Hence, there exists some $\delta'' > 0$, such that

$$\left| \frac{h(x)w_k}{\|x - A_k\|_2} - 1 \right| < \frac{\varepsilon}{2(1 + \varepsilon)} \quad \text{for all } x \text{ with } 0 < \|x - A_k\|_2 \leq \delta''. \quad (5.4)$$

Let $\delta = \min\{\delta', \delta''\}$, then for all $x \neq A_k$ with $\|x - A_k\|_2 \leq \delta$ we have

$$\begin{aligned}
\|T(x) - A_k\|_2 &= \left\| h(x) \sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - x) + \left(\frac{h(x)w_k}{\eta_k(x)} - 1 \right) (A_k - x) \right\|_2 \\
&\geq \left\| h(x) \sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - x) \right\|_2 - \left\| \left(1 - \frac{h(x)w_k}{\eta_k(x)} \right) (A_k - x) \right\|_2 \\
&= h(x) \left\| \sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - x) \right\|_2 - \left| 1 - \frac{h(x)w_k}{\eta_k(x)} \right| \|x - A_k\|_2 \\
&\stackrel{(5.3), (5.4)}{>} h(x)w_k(1 + 2\varepsilon) - \frac{\varepsilon}{2(1 + \varepsilon)} \|x - A_k\|_2 \\
&= \left(\frac{h(x)w_k}{\|x - A_k\|_2} - 1 + 1 \right) (1 + 2\varepsilon) \|x - A_k\|_2 - \frac{\varepsilon}{2(1 + \varepsilon)} \|x - A_k\|_2 \\
&> \left(1 - \frac{\varepsilon}{2(1 + \varepsilon)} \right) (1 + 2\varepsilon) \|x - A_k\|_2 - \frac{\varepsilon}{2(1 + \varepsilon)} \|x - A_k\|_2 \\
&= (1 + \varepsilon) \|x - A_k\|_2.
\end{aligned}$$

Since $\|x - A_k\|_2 > 0$, there exists some $t \in \mathbb{N}$ such that $(1 + \varepsilon)^t \|x - A_k\|_2 > \delta$. Using the above estimation we get $\|T^t(x) - A_k\|_2 > (1 + \varepsilon)^t \|x - A_k\|_2$, so the existence of t guarantees the existence of $s \in \mathbb{N}$ such that $\|T^s(x) - A_k\|_2 > \delta$ and $\|T^{s-1}(x) - A_k\|_2 \leq \delta$, because for $T^0(x) = x$, $0 < \|x - A_k\|_2 < \delta$ holds and the distance monotonically increases. \square

Lemma 5.9. For $A_k \in \mathcal{A}^0$ it holds that

$$\lim_{x \rightarrow A_k} \frac{\|T(x) - A_k\|_2}{\|x - A_k\|_2} = \frac{\text{Test}_k}{w_k}.$$

Proof. For $x \notin \mathcal{A}^0$

$$\begin{aligned}
T(x) &= \frac{\sum_{i=1}^M \frac{w_i}{\eta_i(x)} A_i}{\sum_{i=1}^M \frac{w_i}{\eta_i(x)}} \\
&= \frac{\sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - A_k) + A_k \sum_{i=1}^M \frac{w_i}{\eta_i(x)}}{\sum_{i=1}^M \frac{w_i}{\eta_i(x)}}
\end{aligned}$$

Thus

$$\begin{aligned}
T(x) - A_k &= \frac{\sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - A_k)}{\frac{w_i}{\eta_i(x)}}, \\
\Rightarrow \frac{T(x) - A_k}{\|x - A_k\|_2} &= \frac{\sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - A_k)}{w_k \left(1 + \frac{\|x - A_k\|_2}{w_k} \sum_{i \neq k} \frac{w_i}{\eta_i(x)}\right)} \\
\Rightarrow \lim_{x \rightarrow A_k} \frac{\|T(x) - A_k\|_2}{\|x - A_k\|_2} &= \lim_{x \rightarrow A_k} \left\| \frac{\sum_{i \neq k} \frac{w_i}{\eta_i(x)} (A_i - A_k)}{w_k \left(1 + \frac{\|x - A_k\|_2}{w_k} \sum_{i \neq k} \frac{w_i}{\eta_i(x)}\right)} \right\|_2 \\
\Leftrightarrow \lim_{x \rightarrow A_k} \frac{\|T(x) - A_k\|_2}{\|x - A_k\|_2} &= \frac{\text{Test}_k}{w_k}
\end{aligned}$$

□

Finally, we have gathered all tools to prove the convergence theorem for Algorithm 3.

Theorem 5.10. *Given $x_0 \in \mathbb{R}^n$, define $x_l = T^l(x_0)$ for $l \in \mathbb{N}$. If $(x_l)_{l \in \mathbb{N}} \cap \mathcal{A}^0 = \emptyset$, then $\lim_{l \rightarrow \infty} x_l = x^*$, where x^* is an optimal solution to (MF).*

Proof. For $l \geq 1$ we have that $x_l \in \text{conv}(\mathcal{A})$, which is compact. Hence, by Bolzano-Weierstraß there exists $\tilde{x} \in \text{conv}(\mathcal{A})$ and some subsequence $x_{l_r} \rightarrow \tilde{x}$. We need to show that $\tilde{x} = x^*$. Whenever $x_{l+1} = T(x_l)$ for some l , we have optimality by Lemma 5.6 ii), because $x_l \notin \mathcal{A}^0$ by assumption. Otherwise, Lemma 5.7 states that

$$g(x_0) > g(x_1) > \dots \geq g(x^*)$$

the x_l yield a strictly decreasing and bounded sequence. Hence $(g(x_l))_{l \in \mathbb{N}}$ converges, and using continuity of g and T we see that for the limit it holds that

$$g(\tilde{x}) = \lim_{l \rightarrow \infty} g(x_l) = \lim_{l \rightarrow \infty} g(T(x_l)) = g(T(\lim_{l \rightarrow \infty} x_l)) = g(T(\tilde{x})).$$

From $g(\tilde{x}) = g(T(\tilde{x}))$ we get $\tilde{x} = T(\tilde{x})$, because there is no decrease in g . Then, if $\tilde{x} \notin \mathcal{A}^0$, $\tilde{x} = x^*$ follows by Lemma 5.6.

If $\tilde{x} = A_k$ for some $A_k \in \mathcal{A}^0$ and A_k is optimal, we are done. So the only remaining case is $\tilde{x} = A_k$ for some $A_k \in \mathcal{A}^0$ and A_k is *not* optimal. Then Lemma 5.8 guarantees the existence of a subsequence $(x_{l_r})_{r \in \mathbb{N}}$, such that $\|T(x_{l_r}) - A_k\|_2 > \delta$ for all $r \in \mathbb{N}$. Then

$$\frac{\|T(x_{l_r}) - A_k\|_2}{\|x_{l_r} - A_k\|_2} \geq \frac{\delta}{\|x_{l_r} - A_k\|_2} \rightarrow \infty$$

because $\tilde{x} = A_k$, implying $\|x_{l_r} - A_k\|_2 \rightarrow 0$. This, however, yields a contradiction to Lemma 5.9:

$$\lim_{x \rightarrow A_k} \frac{\|T(x) - A_k\|_2}{\|x - A_k\|_2} = \frac{\text{Test}_k}{w_k}.$$

Thus $\tilde{x} \not\rightarrow A_k$ for A_k not optimal. □

With Theorem 5.10 at hand, we may now calculate the optimal solutions to the optimization problems **L1**, **L2**, **L3**, by applying Algorithm 3 to the reformulated problems that we get after applying Theorem 5.3. We complete this section by demonstrating the reformulation and application of the algorithm with a very small example in \mathcal{T}_4 .

Example 5.1.1. Consider the following five splits on $\{0, 1, 2, 3, 4\}$

$$\begin{aligned} s_1 &= (\{1, 2\}|\{0, 3, 4\}), & s_2 &= (\{1, 3\}|\{0, 2, 4\}), & s_3 &= (\{1, 4\}|\{0, 2, 3\}) \\ s_7 &= (\{1, 2, 3\}|\{0, 4\}), & s_9 &= (\{1, 3, 4\}|\{0, 2\}) \end{aligned}$$

which are labeled according to the canonical embedding in \mathcal{T}_4 , see Example 2.1.4. We have three sample trees $\mathbf{T} = \{T_1, T_2, T_3\}$ with

$$T_1 = ((s_1, s_7), (2, 3)), \quad T_2 = ((s_2, s_7), (4, 3)), \quad T_3 = ((s_3, s_8), (3, 4))$$

and want to exemplarily calculate bounds for the orthant \mathcal{O} that has the split set $\{s_1, s_7\}$.

Now let us calculate z_{L1} for ort_1 . Hence, we have to solve

$$(\mathbf{L1}) \quad z_{L1} := \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} \|t - x\|_2.$$

First we reformulate the embedded vectors according to Theorem 5.3. Note that we need to switch coordinates 2 and 7 of all canonically embedded (Definition 2.9) vectors, to meet the requirements of Theorem 5.3. Thus we have

$$\psi(T_1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \psi(T_2) = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 4 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \psi(T_3) = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 4 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 4 \\ 0 \end{pmatrix}$$

and by taking the first two coordinates of the renumbered vectors we receive

$$t'_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad t'_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad t'_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and $\|t_1''\|_2^2 = 0$, $\|t_2''\|_2^2 = 9$, $\|t_3''\|_2^2 = 25$. Now we can solve the optimization problem

$$\min_{x' \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \left(\|t_i' - x'\|_2^2 + \|t_i''\|_2^2 \right)^{\frac{1}{2}},$$

which has the same structure as (MF). When we want to fit the notation of the definition of (MF), we simply set $c_1 = 0, c_2 = 16, c_3 = 25$ and $A_i = t_i'$ for $i = 1, 2, 3$ and we may apply the adapted Weiszfeld algorithm, Algorithm 3 to find the optimal solution to the problem. Naturally, as the adapted Weiszfeld algorithm is an approximation algorithm, we only determine an approximate optimal solution and in practice stop when some precision threshold has been reached.

The reformulation of **L2** and **L3** is very similar so we will only demonstrate it for **L3** as it yields the stronger bound. Recall, that

$$(\mathbf{L3}) \quad z_{L3} := \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus T_{\mathcal{O}}^c} \|t - x\|_2 + |\mathbf{T}_{\mathcal{O}}^c| \cdot \|x\|_2 \right) + \sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|t\|_2.$$

Here it is important to mention that we omit the constant part $\sum_{T \in \mathbf{T}_{\mathcal{O}}^c} \|T\|_2$ in the optimization process, so that we are left with a weighted median problem. In comparison to the reformulation of **L1** we now additionally need to see whether $T \in \mathbf{T} \setminus T_{\mathcal{O}}^c$, before reformulating the facilities. $T_1 \in \mathcal{O}$, so $T_1 \notin T_{\mathcal{O}}^c$ and $s_7 \in \text{Split}(T_2) \cap \text{Split}(\mathcal{O})$, so $T_2 \notin T_{\mathcal{O}}^c$ as well. T_3 was chosen such that $\text{Split}(T_3)$ and $\text{Split}(\mathcal{O})$ are completely incompatible, thus, $T_3 \in T_{\mathcal{O}}^c$. Thus, when plugging in, we get the following optimization problem

$$z_{L3} = \min_{X \in \mathcal{O}} (\|t_1 - x\|_2 + \|t_2 - x\|_2 + \|x\|_2) + \|t_3\|_2,$$

We define $T_4 = 0 \in \mathcal{T}_4$ to simplify notation and receive

$$z_{L3} = \min_{X \in \mathcal{O}} \left(\sum_{i=1,2,4} \|t_i - x\|_2 \right) + \|t_3\|_2.$$

This implies that we have a different optimization problem than for z_{L1} as T_3 is 'represented' by $T_4 = 0$ in the optimization problem and $\|t_3\|_2$ is added after the optimization step. Now we apply the reformulation of Theorem 5.3 and receive

$$t_1' = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad t_2' = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad t_4' = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and $\|t_1''\|_2^2 = 0$, $\|t_2''\|_2^2 = 16$, $\|t_4''\|_2^2 = 0$. So we actually get the same representatives, but note that the constant changes, as $\|t_4''\|_2^2 = 0$. Again, to fit notation of (MF)

we set $c_1 = 0, c_2 = 9, c_4 = 0$ and $A_i = t'_i$ for $i = 1, 2, 4$ and then start the adapted Weiszfeld algorithm to find the optimal solution. Note that in contrast to z_{L1} , we now need to add $\|t_3\|_2$ after having calculated the optimal objective value for (MF).

The bounds that we calculated for the two reformulated problems using Algorithm 3 are

$$z_{L1} = 10.85, \quad z_{L3} = 12.91$$

and we see that z_{L3} is a lot stronger in this case.

Upper Bounds

Recall the definition of $\mathbf{T}_\mathcal{O}^c$ (5.1) and define the following four upper bounds:

$$\begin{aligned} \text{(U1)} \quad z_{U1} &:= \sum_{T \in \mathbf{T}} d(T, T^a), \quad \text{where } T^a \in \arg \min_{X \in \mathcal{O}} \sum_{T \in \mathbf{T}} \|t - x\|_2 \\ \text{(U2)} \quad z_{U2} &:= \sum_{T \in \mathbf{T}} d(T, T^b), \quad \text{where } T^b \in \arg \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_\mathcal{O}^c| \cdot \|x\|_2 \right) \\ \text{(U3)} \quad z_{U3} &:= \sum_{T \in \mathbf{T}} d(T, T^c), \quad \text{where } T^c \in \arg \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t - x\|_2 + |\mathbf{T}_\mathcal{O}^c| \cdot \|x\|_2 \right) \\ \text{(U4)} \quad z_{U4} &:= \sum_{T \in \mathbf{T}} d(T, T^d), \quad \text{where } T^d \in \arg \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t - x\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|x\|_2 \right) \end{aligned}$$

Note, that the optimization problems for **U1**, **U2**, **U3** on the right hand side are the optimization problems **L1**, **L2**, **L3**. Thus, we only need to calculate the objective values in order to get these bounds. Technically, **Li** are optimization problems, and the **Ui** are not. We only gave the upper bound these names to simply refer to them.

Remark 5.1.2. *It is important to note that the upper bounds depend on the choice of the minimizers T^a, T^b, T^c, T^d and that the resulting values may differ. So, in order to make them well defined we either have to take the minimum over all possible minimizers of the optimization problems, which would yield the best bound, or denote $\mathbf{U1}(T^a)$, $\mathbf{U2}(T^b)$, $\mathbf{U3}(T^c)$, $\mathbf{U4}(T^d)$, respectively. In practice we do not care for this technicality for two reasons: First of all, the minimizers on the right hand side are unique most of the time. Secondly, we still have an upper bound to work with, even if there are several minimizers and one of those yields a better bound. Trying to find the best bound is a big computational effort with very little benefit. In order to ease notation, we will write **U1** instead of $\mathbf{U1}(T^a)$ in what follows (and the same holds for **U2** - **U4**).*

Theorem 5.11. For any fixed orthant $\mathcal{O} \subset \mathcal{T}_n$ $z_{U1}, z_{U2}, z_{U3}, z_{U4}$ are upper bounds for \mathcal{O} and also global upper bounds.

Proof. All four bounds evaluate feasible solutions, namely T^a, T^b, T^c, T^d , for the median problem, naturally yielding global upper bounds. Hence they are also upper bounds for \mathcal{O} itself. \square

At first glance it looks a bit arbitrary why we choose these specific upper bounds, as the objective $\sum_{T \in \mathbf{T}} d(X, T)$ of any tree $X \in \mathcal{T}_n$ yields an upper bound. For **U1** - **U3** we already mentioned that z_{U1}, z_{U2}, z_{U3} are the objective values of the trees T^a, T^b, T^c that we receive as optimal solutions when calculating the lower bounds, i.e., by solving **L1** - **L3**. So for these bounds the computational effort reduces to simply evaluating the objective function of these trees.

An upper bound which we did not receive in this fashion is **U4**. Actually we first established

$$z_{U'} = \sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t\|_2 + \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|x - t\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|x\|_2 \right),$$

as an upper bound. It actually is an upper bound, since $d(X, T) \leq \|x\|_2 + \|t\|_2$ and thus

$$\begin{aligned} \sum_{T \in \mathbf{T}} d(X, T) &= \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|x - t\|_2 + \sum_{T \in \mathbf{T} \setminus \mathcal{O}} d(X, T) \\ &\leq \sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t\|_2 + \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|x - t\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|x\|_2 \end{aligned}$$

holds for all $X \in \mathcal{O}$, so it also holds for the minimum. We disregarded $z_{U'}$ since z_{U4} is a better bound:

Lemma 5.12. $z_{U4} \leq z_{U'}$.

Proof. By definition of z_{U4} , T^d is a minimizer of the right hand side, so we may plug

this minimizer in:

$$\begin{aligned}
& \sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t\|_2 + \min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|x - t\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|x\|_2 \right) \\
&= \sum_{T \in \mathbf{T} \setminus \mathcal{O}} \|t\|_2 + \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t^d - t\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|t^d\|_2 \\
&= \sum_{T \in \mathbf{T} \setminus \mathcal{O}} (\|t\|_2 + \|t^d\|_2) + \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t^d - t\|_2 \\
&\geq \sum_{T \in \mathbf{T} \setminus \mathcal{O}} d(T, T^d) + \sum_{T \in \mathbf{T} \cap \mathcal{O}} \|t^d - t\|_2 \\
&= \sum_{T \in \mathbf{T}} d(T, T^d).
\end{aligned}$$

□

As mentioned above, T^a, T^b, T^c are calculated when we compute the lower bounds, but T^d is not. Hence we need to solve

$$\min_{X \in \mathcal{O}} \left(\sum_{T \in \mathbf{T} \cap \mathcal{O}} \|x - t\|_2 + |\mathbf{T} \setminus \mathcal{O}| \cdot \|x\|_2 \right).$$

This can, analogously to the lower bounds, be solved by using our reformulation, Theorem 5.3, and then applying Algorithm 3.

Experiments

In this section we demonstrate that the bounds we have given in the previous subsections are easy, i.e., fast to calculate and examine how well they perform for different sizes and types of instances. All bounds have been calculated in a self implemented framework for tree space algorithms which include classes for splits, orthants, trees and several helpful functions and methods. This framework is also used for all other experiments to follow. In particular we stress that the calculation of all our lower bounds includes solving an optimization problem; In order to solve these we have implemented the Weiszfeld algorithm, as well as our adapted version of the Weiszfeld algorithm that we developed in Section 5.1.

Regarding the technical specifications, we have implemented all algorithms and methods in Python 3, the calculations have been executed on a single core of an Intel(R) Core(TM) i3-2350M and with 4GB RAM.

There are numerous interesting features of the bounds we could investigate here but since they are not of main importance, but rather a helpful tool for the heuristic

we develop in the next section, we only focus on two things. Firstly, we examine how big the gap between the bounds is to get an idea of how good our solutions for the median problem will be. Secondly, we want to find out how many orthants we can eliminate or reject by finding an upper bound for some orthant that is lower or equal than the lower bound for this orthant.

In general, however, it will be hard to depict all results of the bounds. This is due to the combinatorics of the tree space: Recall that for \mathcal{T}_n there exist $(2n - 3)!!$ orthants for which we have to calculate each of the bounds. Naturally, we do not show the values of the bounds for each orthant, as the size of tables and amount of values of different bounds is unmanageable.

So in the following we repeatedly conduct experiments for a fixed parameter setting and average the results of the bounds over all orthants and all instances. Moreover, we always pick single generic instances and extreme instances to show how the bounds look like when simply averaged over the orthants of a single instance instead of averaging over all instances.

As mentioned earlier, we calculate the bounds for $(2n - 3)!!$ orthants. The highest dimension for which we are able to do this is $n = 7$ and this is very time consuming already, so as a performance test we run an experiment on a single instance simply for the purpose of investigating the computation time, rather than interpreting the results. Otherwise we perform and compare results of the bounds in detail for $n = 4, 5$.

Generation of Data

We use the following parameters to describe and generate instances:

n	dimension of the tree space
m	number of repeated runs for a fixed parameter setting
o	number of orthants to sample the trees from
M	number of trees in an instances
W	maximal weight of splits

We will always use $W = 20$ and sample an integer in $[0, W]$ for the weight of a split using the uniform distribution. o is the value of orthants to which we restrict ourselves when sampling trees; this gives a handle on how widely spread the instances are. When $o = 1$, we obviously have only a single orthant and a resulting Euclidean location problem and when $o = (2n - 3)!!$, then we sample from all orthants of \mathcal{T}_n . The values in between, however, are far more interesting, as the trees are sampled in ,say, $o = 15$ of $105 = (10 - 3)!!$ orthants, then it is much more likely that the 15 orthants share a specific split or that one orthant contains a very high number of trees. The o orthants are a random sample from the $(2n - 3)!!$ orthants with respect to the uniform distribution.

o	\emptyset gap in %	\emptyset rej. orthants	\emptyset time per instance
5	14.02	7	0.3s
10	10.89	4.4	0.3s
15	10.10	2.5	0.3s

Table 5.1: Bounds experiment in \mathcal{T}_4 for $M = 5$ trees, averaged over $m = 10$ runs with different numbers of sample orthants o .

o	gap in %	rej. orthants	total time
5	5.71	1.1	3s
10	2.30	1.6	3.8s
15	1.32	2.9	3.9s

Table 5.2: Bounds experiment in \mathcal{T}_4 for $M = 50$ trees, averaged over $m = 10$ runs with different numbers of sample orthants o .

Finally, the M trees are generated as follows: First choose one of the o sampled orthants, again with respect to the uniform distribution, and then when the orthant is chosen, sample the weights for each of the splits of the orthant as described earlier.

Results

We start with the experiments for $n = 4$. Table 5.1 summarizes the outcomes over all $m = 10$ runs for $M = 5$ for $o \in \{5, 10, 15\}$ and Table 5.2 depicts the outcomes for $m = 10$ and $M = 50$ for $o \in \{5, 10, 15\}$.

As all experiments we conduct show common trends for the quality of bounds, the size of the gap, etc., we interpret them after the last experiment. The only thing we hint at here is that the average computation time in \mathcal{T}_4 is 0.3s or about 3.5s, respectively, which is very affordable.

From the 60 instances of the above experiments we picked out three to illustrate a ‘generic’, a ‘good’ and a ‘bad’ case. These may be found in Figure 5.1. The illustrations for each instance include the global lower bound (‘globalL’) and the best upper bound, hence the best global upper bound (‘bestUp’) we determined. The rightmost bar indicates how many of all orthants could be rejected by the bounds, i.e., the orthants in which it is clear that they do not contain an optimal solution.

The topmost instance in Figure 5.1 is a ‘generic’ instance with $M = 50$ trees in the sense that its gap and the number of orthants that are rejected are around the average values from Table 5.2. For $M = 50$, most of the instances were more or

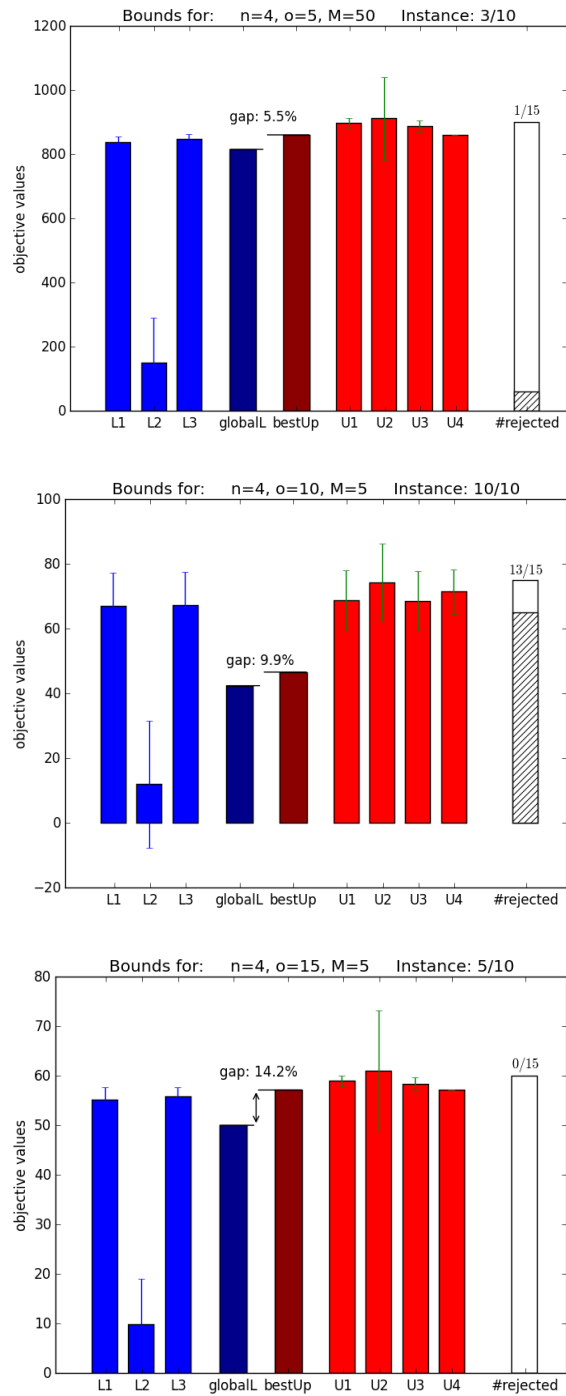


Figure 5.1: Generic and extreme instances from the bounds experiments in \mathcal{T}_4 . Both extreme cases are with $M = 5$ trees, for $M = 50$, the results were stable.

o	\emptyset gap in %	\emptyset rej. orthants	\emptyset total time
5	11.27	28.4	1.9s
25	10.41	7.4	2.1s
75	6.99	3.2	2.0s
105	9.75	11.1	2.0s

Table 5.3: Bounds experiment in \mathcal{T}_5 for $M = 5$ trees, averaged over $m = 10$ runs with different numbers of sample orthants o .

o	\emptyset gap in %	\emptyset rej. orthants	\emptyset total time
5	5.97	9.9	18.0s
25	2.18	6.8	23.8s
75	1.59	3.8	24.4s
105	1.56	4.0	23.5s

Table 5.4: Bounds experiment in \mathcal{T}_5 for $M = 50$ trees, averaged over $m = 10$ runs with different numbers of sample orthants o .

less consistently close to these average values. For $M = 5$, there are more extreme examples, which is not very surprising, since it is more likely for the trees not to be spread out too evenly across the sample orthants. The second instance in the picture depicts an extraordinary ‘good’ case, where only two out of fifteen orthants remain as candidates to contain the optimal solution. The last instance shows a rather ‘bad’ case, where the gap is high and not a single orthant could be rejected. For most of the instances where the gap is considerably higher than 10% it was usually the case that some orthants could be rejected, which is unfortunately not the case here.

Now we investigate the bounds for \mathcal{T}_5 . Here, we adapt the amount of orthants in which the trees are sampled, since there are seven times as many orthants now and we always want to have one case where the 50 trees are sampled over the whole space. Moreover, we conduct the experiments for four different values of o , since the span of values naturally becomes a lot wider.

Table 5.3 summarizes the outcomes over all $m = 10$ runs for $M = 5$ for $o \in \{5, 25, 75, 105\}$ and Table 5.4 depicts the outcomes for $m = 10$ and $M = 50$ for $o \in \{5, 25, 75, 105\}$.

Again, we showcase some of the results for specific instances in Figure 5.2.

Similarly to Figure 5.1, in Figure 5.2, the topmost instance is a ‘generic’ instance, for which the gap and the number of rejected orthants is close to the average values.

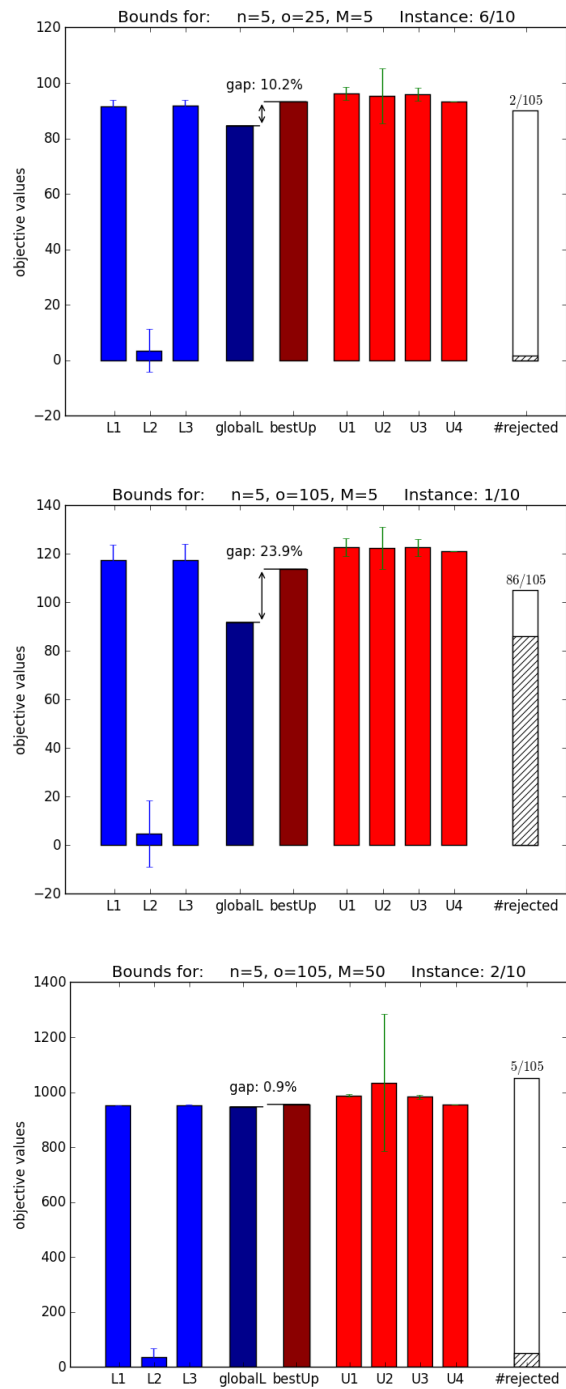


Figure 5.2: Generic and extreme instances from the bounds experiments in \mathcal{T}_5 .

o	gap in %	rej. orthants	total time
5	2.68	20	2377.6s
10395	0.71	47	3575.8s

Table 5.5: Bounds experiment in \mathcal{T}_7 for $M = 50$ trees for two exemplary instances.

The second instance in the picture shows an extraordinary ‘good’ case, where only under 20% of the orthants remain as candidates to contain the optimal solution, as 86/105 orthants were rejected. The last instance shows an extreme case in the sense that the relative gap dropped below 1%, showing that margins that are to be gained by optimization are rather small here and the tree whose objective yields the global upper bound is a pretty good solution.

As we pronounced earlier, as a performance test we want to compute all bounds on two exemplary instances, for which the results will, of course, not be representative. We note that, before starting to calculate the bounds, our framework needs to generate all splits and orthants for \mathcal{T}_7 and that there are 10395 orthants in \mathcal{T}_7 . Surely, there is a faster way to do this, but when implemented in a straightforward way, the generation of these splits and orthants for \mathcal{T}_7 already takes several minutes and the file that stores all orthants is already 2.5MB big. This is why we refrain from trying to lift this up to even higher dimensions as this would force us to use computing clusters, which is unnecessary to investigate the behavior of the bounds. The results of the two runs are shown in Table 5.5

As we may see, the computation times stack up considerably due to the high amount of orthants. Nonetheless, calculation of bounds in single orthants is in average less than 0.1s, i.e., extremely low. So in a practical setting one may calculate z_{L3} for all orthants and only calculate a few reasonable upper bounds for high dimensional problems when it is too time consuming to calculate all bounds.

Now that we have seen several experiments we can interpret the influence of dimension, number of trees and number of orthants.

The most obvious results are obtained for higher numbers of trees. For all dimensions we observe a proportional connection between the number of trees and the computation times. This is not surprising, since all lower bounds are calculated by solving optimization problems; these problems are solved by the Weiszfeld algorithm or by our adapted Weiszfeld algorithm from Section 5.1. Both algorithms depend linearly on the number of trees in each iteration, so the proportional connection was to be expected.

Another observation that we can make for different number of trees is that the gap gets distinctly lower when the number of trees is higher. This insight is very

important and intertwined with a lot of other results and interpretations:

The gap actually gets smaller with a higher number of trees and when the instance is more *spread*, i.e., sampled over a higher number of instances. This is due to the structure of the space and the problem. Due to incompatibilities of splits and the non-positive curvature it is very likely, and often the case in the experiments, that the star tree is the optimal solution to the problem for very spread instances. This then results in the optimal solutions to many of the optimization problems also being the origin in \mathbb{R}_+^{n-2} . Take for example z_{L3} and z_{U3} and assume that 0 is the minimizer for **L3**. This is not a very rare case, as the weight $|\mathbf{T}_O^c|$ of the facility at 0 is the only weight of the Wever problem that is greater than 1. If 0 is the minimizer, then the gap of these two bounds is

$$z_{U3} - z_{L3} = \sum_{i=1}^M d(0, T_i) - \left(\sum_{T \in \mathbf{T} \setminus T_O^c} \|t - 0\|_2 + |\mathbf{T}_O^c| \cdot \|0\|_2 + \sum_{T \in \mathbf{T}_O^c} \|t\|_2 \right) = 0,$$

and is thus the optimal objective for this orthant. Of course, there may be different orthants, for which the lower bounds are smaller than that and make the global lower bound smaller as well. Then the gap may not be 0 any more, but it seems to be pretty small in most of these cases.

All in all the experiments show that the bounds perform pretty well: the relative gap is around 10% or even considerably lower, which was not to be expected a priori. Moreover, the calculation of the bounds in single orthants is extremely fast and it is still affordable to calculate them for the whole \mathcal{T}_7 . Unfortunately as the number of orthants is $(2n - 3)!!$, Theorem 2.12, the calculation of the bounds will at some point take too long, no matter how fast the bounds for single orthants are implemented. Also, the number of rejected orthants varies heavily. Of course, this depends on the instance, but there is a big variance in the number of rejected orthants even for the instances that are sampled with respect to the same parameters. Thus, one may not rely on the bounds to consistently reject orthants for all cases. Nevertheless, for specific instances, where the data is very homogeneous, the bounds are able to rule out a big amount of orthants.

As for the practical use of the bounds it would make no sense to use z_{L1} and z_{L2} , due to Theorem 5.2. Thus, when we use the bounds we only calculate z_{L3} , which implies that we then only get z_{U3} as well, as we have to calculate the minimizers for z_{L1} and z_{L2} to calculate z_{U1} and z_{U2} . Then we are left with z_{U3} and z_{U4} as upper bounds. An interesting observation is that z_{U4} seems to provide a very reasonable and stable upper bound for an instance. In all experiments we performed above it very rarely happened that z_{U4} has differed for different orthants of a fixed instance. Additionally, it still yielded the best upper bound for many instances. Hence, when one needs to rigorously save time one could only calculate z_{U4} for a single orthant

and use it as the global upper bound. Nevertheless, in order to reject orthants we need to calculate z_{L3} anyways, so calculating z_{U3} is not the biggest effort as well, as we only need to calculate the objective value of the minimizer of **L3**.

5.2 The Global Balance Point Heuristic

In the following we introduce the Global Balance Point Heuristic (GBPH) for the median problem, which is based on a local improvement strategy, which is called the *Balance Point Algorithm* (BPA). The BPA may be applied to all orthants of \mathcal{T}_n but it works on a single fixed orthant at a time and requires a lot of careful handling of notation. To this end we introduce some conventions for notation to avoid confusion in the following.

Firstly, the fixed orthant on which the BPA is applied will always be denoted by \mathcal{O} . The orthant embedding $\psi_{\mathcal{O}} : \mathcal{O} \rightarrow \mathbb{R}_+^{n-2}$, compare Definition 2.13, is also always meant to be defined for \mathcal{O} , so we write $\psi = \psi_{\mathcal{O}}$. Moreover, we usually write $\|z - y\|_2 = d(Y, Z)$, when $Y, Z \in \mathcal{T}_n$ are in the same orthant, but we have that $y = \chi(Y), z = \chi(Z) \in \mathbb{R}_+^N$ are the canonical embeddings of Y, Z , compare Definition 2.9. The goal is to reformulate the median problem to $\mathbb{R}_+^{n-2} = \psi(\mathcal{O})$. Hence, we use $\psi(Y), \psi(Z) \in \mathbb{R}_+^{n-2}$ instead and write $\|\psi(Z) - \psi(Y)\|_2$. This does not matter since only the zero entries are omitted and $\|\psi(Z) - \psi(Y)\|_2 = \|z - y\|_2$ holds. Moreover, the BPA is an algorithm that involves iterates $x_j \in \mathbb{R}_+^{n-2}$, i.e., $x_j \notin \mathbb{R}_+^N$ is not the canonical embedding of a tree X_j . So for this section it is important to keep in mind, that the usual relation of capital and lowercase letters via the canonical embedding does not always hold.

Idea

Before jumping into technical definitions we want to state the idea of the Global Balance Point Heuristic that we develop in this section. The heuristic consists of two parts. The first part is a preprocessing strategy with which we discard as many orthants as possible for which we can show that they do not contain an optimal solution. Then we are left with several “candidate” orthants that may contain an optimal solution. The second part of the heuristic is an iterative local improvement procedure, the BPA, that we apply in each candidate orthant. For each orthant the BPA yields a solution and in the end we choose the best solution that has been found as the solution for the GBPH. Now we want to outline the general idea of the BPA in a fixed orthant. One step of the iteration is depicted in Figure 5.3. The main idea of the BPA is to replace facilities that are not contained in the orthant \mathcal{O} , where we apply the procedure, by representatives in the orthant. Since each orthant in \mathcal{T}_n is isometric to \mathbb{R}_+^{n-2} , we can then work within a Euclidean setting.

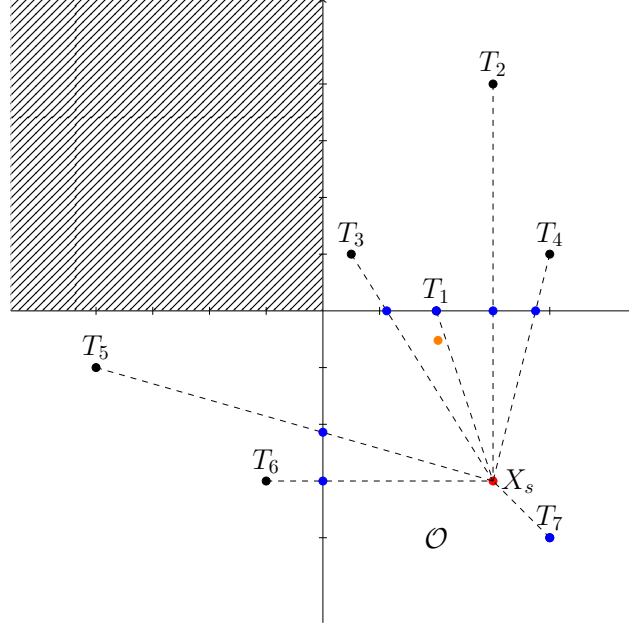


Figure 5.3: A sketch of one step of the iterative procedure in a fixed orthant \mathcal{O} with starting point X_s (red). First, choose $B_i \in \mathcal{O}$ (blue) on $\Gamma(X_s, T_i)$, that is farthest away from X_s , then calculate the median X_m (orange) of the B_i , which are all contained in \mathcal{O} .

Suppose now that $X_j \in \mathcal{O}$ is our current iterate. Then its objective value is $\sum_{i=1}^M d(X_j, T_i)$. We denote the representatives of the facilities T_i by B_i . When $T_i \in \mathcal{O}$ we simply set $B_i = T_i$. Otherwise, the geodesic $\Gamma(X_j, T_i)$, parametrized by γ , has to leave \mathcal{O} at some point $B_i = \gamma(e_i(X_j)) \in \partial\mathcal{O}$, where $e_i(X_j) \in [0, 1]$ is the maximal value such that $\gamma(e_i(X_j)) \in \mathcal{O}$. This “exit point” B_i is the representative for T_i .

With this we can divide the distance $d(X_j, T_i)$ in two parts. The first $d(T_i, B_i)$ is the part of the geodesic that is outside the \mathcal{O} and the second part $d(B_i, X_j)$ is the part inside \mathcal{O} , i.e.,

$$\sum_{i=1}^M d(X_j, T_i) = \sum_{i=1}^M d(T_i, B_i) + \sum_{i=1}^M d(B_i, X_j).$$

Note, that when $T_i \in \mathcal{O}$, we have an ‘outside’ contribution of 0 as $T_i = B_i$. Now, to improve the objective we will keep the B_i as is, hence the outside part $\sum_{i=1}^M d(T_i, B_i)$ will be constant. Recall once more that the key is, that all of the B_i are in \mathcal{O} . Hence, we can just use the Euclidean distance $\|\psi(B_i) - \psi(X_j)\|_2 = d(X_j, B_i)$ and

the isometric embedding ψ into \mathbb{R}_+^{n-2} to transform $\sum_{i=1}^M d(B_i, X_j)$ to a Euclidean median problem $\sum_{i=1}^M \|\psi(B_i) - \psi(X_j)\|_2$.

Now, the minimizer x_{j+1} of the Euclidean median problem $\sum_{i=1}^M \|\psi(B_i) - \psi(X_j)\|_2$ is either a new point $X_{j+1} = \psi^{-1}(x_{j+1})$ with $x_{j+1} \neq x_j$ or we terminate the iterative procedure when $x_{j+1} = x_j$.

5.2.1 The Balance Point Algorithm

In the previous subsection we have sketched the idea of the heuristic and especially of its central part, the BPA. We want to formalize this now and start by introducing the concept of *exit points* for the representatives of the facilities within the orthant \mathcal{O} . To this end, recall that there exists a unique geodesic connecting two points $Y, Z \in \mathcal{T}_n$ and that there is a unique parametrization with constant speed from $[0, 1] \rightarrow \mathcal{T}_n$, compare Theorem 2.21.

Definition 5.13. *Given an orthant \mathcal{O} , $T \in \mathcal{T}_n \setminus \mathcal{O}$ and $x \in \mathbb{R}_+^{n-2}$, let $\gamma_x : [0, 1] \rightarrow \mathcal{T}_n$ denote the parametrization of the geodesic from $\psi^{-1}(x)$ to T . With this, define the maps*

$$e_T : \mathbb{R}_+^{n-2} \rightarrow [0, 1], \quad e_T(x) = \max\{\lambda \in [0, 1] : \gamma_x(\lambda) \in \mathcal{O}\} \quad (5.5)$$

$$E_T : \mathbb{R}_+^{n-2} \rightarrow \mathcal{O}, \quad E_T(x) = \gamma_x(e_T(x)) \quad (5.6)$$

and we say that $E_T(x)$ is the exit point of the geodesic from $\psi^{-1}(x)$ to T in \mathcal{O} .

With this we can define the central notion of *balance points*:

Definition 5.14. *Given $\mathbf{T} = \{T_1, \dots, T_M\} \subset \mathcal{T}_n$ and an orthant \mathcal{O} , we call $X \in \mathcal{O}$ a balance point if for $x = \psi(X)$ it holds that*

$$x \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|B_i(x) - y\|_2, \quad \text{where } B_i(x) = \begin{cases} \psi(T_i) & \text{if } T_i \in \mathcal{O}, \\ \psi(E_{T_i})(x) & \text{if } T_i \notin \mathcal{O}. \end{cases}$$

We now formalize the iterative procedure described above and call it *Balance Point Algorithm* (BPA) because it terminates when a balance point as defined in Definition 5.14 is found.

Data: $\mathbf{T} = \{T_1, \dots, T_M\}$, $\mathcal{O} \subset \mathcal{T}_n$, $x_0 \in \text{int } \mathbb{R}_+^{n-2}$

Result: A balance point X_B in \mathcal{O} .

for $j = 0, 1, 2, \dots$ **do**

Set $B_i(x_j) =$	$\begin{cases} \psi(T_i) & \text{if } T_i \in \mathcal{O} \\ \psi(E_{T_i}(x_j)) & \text{if } T_i \notin \mathcal{O} \end{cases}$	as in Definition 5.14.
Find $x_{j+1} \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \ B_i(x_j) - y\ _2$		
if $x_{j+1} = x_j$ then		
return $\psi^{-1}(x_j)$		

end

Algorithm 4: The Balance Point Algorithm alternates between finding the median x_j of the facilities B_i and calculating new exit points B_i with respect to the newly calculated iterate x_j until $x_j = x_{j-1}$.

First of all, having formally described the algorithm, we prove that the objective of the iterates does not increase:

Lemma 5.15. *Let x_j, x_{j+1} be successive iterates of the BPA. Then $f(\psi^{-1}(x_{j+1})) \leq f(\psi^{-1}(x_j))$.*

Proof. Let $X_{j+1} = \psi^{-1}(x_{j+1})$ and $X_j = \psi^{-1}(x_j)$. By definition we have that $\psi^{-1}(B_i(x_j)) = E_{T_i}(x_j)$ is on the geodesic between X_j and T_i , so $\|x_j - B_i(x_j)\|_2 + d(E_{T_i}(x_j), T_i) = d(X_j, T_i)$. Moreover, recall that x_{j+1} is the minimizer of the Euclidean median problem with respect to $B_i(x_j)$ by definition. These two statements and the application of the triangle inequality (used in the first line) combined yield

$$\begin{aligned}
f(X_{j+1}) &= \sum_{i=1}^M d(X_{j+1}, T_i) \leq \sum_{i=1}^M d(X_{j+1}, E_{T_i}(x_j)) + \sum_{i=1}^M d(E_{T_i}(x_j), T_i) \\
&= \sum_{i=1}^M \|x_{j+1} - B_i(x_j)\|_2 + \sum_{i=1}^M d(E_{T_i}(x_j), T_i) \\
&= \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - B_i(x_j)\|_2 + \sum_{i=1}^M d(E_{T_i}(x_j), T_i) \\
&\leq \sum_{i=1}^M \|x_j - B_i(x_j)\|_2 + \sum_{i=1}^M d(E_{T_i}(x_j), T_i) = f(X_j),
\end{aligned}$$

what was to show. □

The question is how Algorithm 4 behaves, when there are several minimizers of

$$\min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|B_i(x_j) - y\|_2.$$

There are two cases: The first is that the minimizer is not unique but we get a solution x_{j+1} that yields a lower objective than x_j . In this case cycling is impossible, since we can never get x_j as a solution again. Hence, the interesting case is what happens when the objective stays the same:

Lemma 5.16. *Let $x_j \in \mathbb{R}_+^{n-2}$ be the current iterate of the BPA and let*

$$x_{j+1} \in M := \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|B_i(x_j) - y\|_2.$$

If $f(\psi^{-1}(x_{j+1})) = f(\psi^{-1}(x_j))$, then it holds that $B_i(x_j) = B_i(x_{j+1})$ for $i = 1, \dots, M$.

Proof. First assume that $T_i \in \mathcal{O}$. Then $B_i(x) = \psi(T_i)$ for all $x \in \mathbb{R}_+^{n-2}$, in particular for $x = x_j$ and $x = x_{j+1}$. Now consider $T_i \notin \mathcal{O}$. If $\psi^{-1}(B_i(x_j)) \notin \Gamma(\psi^{-1}(x_{j+1}), T_i)$, then

$$d(\psi^{-1}(x_{j+1}), T_i) < d(\psi^{-1}(x_{j+1}), \psi^{-1}(B_i(x_j))) + d(\psi^{-1}(B_i(x_j)), T_i)$$

and continuing the calculations of the proof of Theorem 5.15 we get a contradiction as it follows that $f(\psi^{-1}(x_{j+1})) < f(\psi^{-1}(x_j))$. Thus, $\psi^{-1}(B_i(x_j)) \in \Gamma(\psi^{-1}(x_{j+1}), T_i)$, but then $B_i(x_j) \in \partial \mathbb{R}_+^{n-2}$ is also the exit point of $\Gamma(\psi^{-1}(x_{j+1}), T_i)$, i.e., $B_i(x_j) = B_i(x_{j+1})$. \square

The essence of Lemma 5.16 is the following: When we use a deterministic algorithm \mathcal{MA} for the Euclidean median problem that takes a set of facilities as input and returns a median, then we get that

$$x_{j+1} = \mathcal{MA}(\{B_1(x_j), \dots, B_M(x_j)\}) = \mathcal{MA}(\{B_1(x_{j+1}), \dots, B_M(x_{j+1})\}) = x_{j+2}$$

and the BPA terminates. This culminates in the following:

Theorem 5.17. *If a deterministic procedure is used to solve the Euclidean median problem in each step, then the BPA does not cycle.*

5.2.2 Preprocessing

As we have pointed out earlier, the BPA only works in a fixed orthant of \mathcal{T}_n , and improves the given start solution within the orthant \mathcal{O} in each step. However, we do not know beforehand in which orthant the optimal solution T^* is contained, which brings us back to the problem of finding the “candidate” orthants.

The brute-force approach to tackle this problem is to apply this iterative method for some starting point in each orthant and then take the best solution of all orthants. However, this approach is not practical, because the number of orthants is $(2n-3)!!$,

hence exponential in n . The better approach is to tackle this problem by discarding orthants \mathcal{O}' which do not contain an optimal solution. This can be done via a preprocessing strategy.

The first part of the preprocessing is based on our results from Subsection 3.2.1. With this we know that we may discard all orthants that contain at least one split s such that $s \notin \bigcup_{T \in \mathbf{T}} \text{Split}(T)$, see Theorem 3.4. Additionally we may also discard all orthants that do not contain all splits $s \in S^* = \bigcap_{T \in \mathbf{T}} \text{Split}(T)$, see Theorem 3.3.

The second part of the preprocessing is based on the bounds on the objective function that we developed in Section 5.1. For this we use the lower bound z_{L3} and the upper bounds $z_{U1}, z_{U2}, z_{U3}, z_{U4}$ that we established in Section 5.1.

We calculate these bounds for all orthants that are not already discarded after the *split preprocessing* we described above. In the end, we check if an orthant \mathcal{O}' has a lower bound L' such that $L' \geq U^*$, where U^* is the smallest upper bound we have found. If $L' \geq U^*$ we may discard \mathcal{O}' , because it contains no optimal solution.

These two preprocessing approaches are implemented in the algorithmic scheme of our heuristic, that we formulate in the next section.

5.2.3 Pseudocode Formulation of the Heuristic

In order to simplify notation, define $\bar{n} = (2n - 3)!!$ and let the maximal orthants of \mathcal{T}_n denoted by $\mathcal{O}_1 \cup \dots \cup \mathcal{O}_{\bar{n}} = \mathcal{T}_n$. The algorithmic scheme of the GBPH is given in Algorithm 5.

Data: $\mathbf{T} = \{T_1, \dots, T_M\}$, $\mathcal{C} = \{\mathcal{O}_1, \dots, \mathcal{O}_{\bar{n}}\}$

Result: The best balance point X_B of all orthants.

Initialization: $X_B = 0 \in \mathcal{T}_n$.

Let $W := \max_{i=1, \dots, M} \left\{ \max_{s \in \text{Split}(T_i)} w_s^{T_i} \right\}$ and define $x_s := \frac{W}{3} \cdot \mathbf{1} \in \mathbb{R}_+^{n-2}$

Split Preprocessing:

Let $S_F = \bigcap_{i=1}^M \text{Split}(T_i)$ be the set of “forced splits” and

let $S_P = \bigcup_{i=1}^M \text{Split}(T_i)$ be the set of “possible splits”.

for $\mathcal{O}_i \in \mathcal{C}$ do

 | if $S_F \not\subset \text{Split}(\mathcal{O}_i)$ OR $\text{Split}(\mathcal{O}_i) \not\subset S_P$ then

 | $\mathcal{C} := \mathcal{C} \setminus \{\mathcal{O}_i\}$.

end

Bound Preprocessing:

for $\mathcal{O}_i \in \mathcal{C}$ do

 | Calculate $z_{L3}, z_{U1}, z_{U2}, z_{U3}, z_{U4}$ for \mathcal{O}_i and let

 | $X_i^l \in \mathcal{O}_i$ be the minimizer of the optimization problem of **L3**.

 | $l_i := \mathbf{L3}$, $u_i := \min\{z_{U1}, z_{U2}, z_{U3}, z_{U4}\}$

end

Set $u^* = \min_{\mathcal{O}_i \in \mathcal{C}} u_i$

for $\mathcal{O}_i \in \mathcal{C}$ do

 | if $l_i \geq u^*$ then

 | $\mathcal{C} := \mathcal{C} \setminus \{\mathcal{O}_i\}$

end

Main Loop:

for $\mathcal{O}_i \in \mathcal{C}$ do

 | Let X_i be the solution of Algorithm 4 for orthant \mathcal{O}_i with starting point

 | x^s .

 | if $f(X_i) < f(X_B)$ then

 | $X_B = X_i$

end

Output: X_B

Algorithm 5: The Global Balance Point Heuristic applies the BPA in each orthant that has not been discarded in the preprocessing and yields a balance point for each orthant. X_B is the best balance point that has been found.

5.2.4 Experiments

In this section the goal is to illustrate how the heuristic works, how good the solutions of the heuristic are and how big the computational effort is. We investigate these questions for several types of instances.

We start with a detailed analysis in \mathcal{T}_4 , where we investigate specifically constructed types of instances, for which we check if the algorithm returned the optimal solution. We then present experiments for \mathcal{T}_5 where we investigate computational effort depending on the type and size of instance and finally provide a performance test in \mathcal{T}_6 .

The Proximal Point Algorithm

As we do not know the optimal solutions to the instances we randomly generate, we need a method to compare our results. To this end, we introduce Bacak's Proximal Point Algorithm [Bac14a] in the cyclic order version. In the paper, Bacak develops algorithms to compute the median and the mean in Hadamard spaces. Since the tree space is a Hadamard space, these algorithms may be deployed here. Instead of stating the algorithms and the theorem from [Bac14a] in the general and original versions, we already translate them to our specific tree space setting. To this end, let $\mathbf{T} = \{T_1, \dots, T_M\}$ as always, let $\gamma_{(Y,Z)}$ denote the geodesic from $Y \in \mathcal{T}_n$ to $Z \in \mathcal{T}_n$ and for a given sequence $(\lambda_k)_{k \in \mathbb{N}_0}$ of positive reals define

$$t_k^\eta(X) = \min \left\{ 1, \frac{\lambda_k}{d(X, T_\eta)} \right\} \quad \text{for } \eta = 1, \dots, M, k \in \mathbb{N}_0.$$

The formal definition of the cyclic order PPA is given in Algorithm 6.

Data: $X_0 \in \mathcal{T}_n, (\lambda_k)_{k \in \mathbb{N}}$ s.t. $\sum_{i=0}^{\infty} \lambda_k = \infty, \sum_{i=0}^{\infty} \lambda_k^2 < \infty,$

for $k = 0, 1, 2, \dots$ **do**

$$X_{kM+1} := \gamma_{(X_{kM}, T_1)}(t_k^1(X_{kM}))$$

$$X_{kM+2} := \gamma_{(X_{kM+1}, T_2)}(t_k^2(X_{kM+1}))$$

⋮

$$X_{kM+M} := \gamma_{(X_{kM+M-1}, T_M)}(t_k^M(X_{kM+M-1}))$$

end

Algorithm 6: Bacak's Proximal Point Algorithm in the cyclic order version, see [Bac14a].

Even though Algorithm 6 looks a bit technical it is a surprisingly simple method: In each cycle, which consists of calculating the M iterates for a fixed k , the iterate X_{kM}

is shifted a bit towards T_1 resulting in X_{kM+1} which in turn is shifted a bit towards T_2 and so on and so forth. The shift of the iterates is actually described by moving on the geodesic between the two points. The sequence (λ_k) was chosen in such a way that it is decreasing eventually and hence the parameters t_k^n that are plugged into the geodesic are also decreasing eventually, so that there only is a infinitesimal change from the current iterate X_{kM+i-1} towards a given tree T_i for big k .

Just as for the algorithm, we also state the convergence theorem, Theorem 3.4, [Bac14a], in the tree space setting. Furthermore Theorem 3.4, [Bac14a] actually proves convergence for a class of objective functions with certain properties. In particular the theorem holds for f_{Med} so we also state it specifically for the median objective here.

Theorem 5.18 ([Bac14a]). *Let $(\lambda_k)_{k \in \mathbb{N}_0}$ be a sequence of positive reals satisfying $\sum_{i=0}^{\infty} \lambda_k = \infty$, $\sum_{i=0}^{\infty} \lambda_k^2 < \infty$. Then for any starting point $X_0 \in \mathcal{T}_n$ the sequence defined as in Algorithm 6 converges to a minimizer of f_{Med} , i.e., a median of $\{T_1, \dots, T_M\}$.*

In the following we simply refer to the cyclic order version of the PPA as PPA. We will use the PPA with the sequence $(\lambda_k)_{k \in \mathbb{N}_0}$ with $\lambda_k = \frac{1}{k+1}$ for $k \in \mathbb{N}_0$, or we use a scaled version of the sequence, i.e., $c \cdot \lambda_k$ for $k \in \mathbb{N}_0$, where $c > 0$. Independent of the scaling factor, these sequences all satisfy the requirements of Theorem 5.18, which implies that PPA converges to a median.

Nonetheless the theorem neither yields a rate of convergence nor a nice stopping criterion. Due to its cyclic, iterative nature it is really hard to gain some sort of insight how close one is to the optimal solution. For example, when already starting the algorithm at the optimal solution it will simply start to circle around it.

Setup and Implementation

The generation of the random instances for the experiments is the same as described in Section 5.1, just that we use $W = 100$ as the maximal weight of splits here. Sometimes we use a specific subset of orthants in, which the trees are sampled, but will always mention this beforehand. This is to show the behavior of the algorithm for problems with a special underlying structure.

All algorithms are implemented in Python 3, this includes the preprocessing as well as the BPA. The BPA in particular includes the calculation of geodesic distances and the parametrization of the geodesic. We implemented these routines on our own, since the interface to the Java implementation for the geodesic distance that was made available by the authors of [OP11] was not sufficient for us to calculate the parametrizations. Moreover, when testing our algorithm, we realized that there is a significant speedup in distance calculation, which is probably due to the well

implemented max-flow problem solvers in the “networkx” library [HSSC08] that we used. Max-Flow problems arise as the crucial subproblems in the calculation of the geodesic distance and the parametrization of the geodesic, see [Owe11] for more details on the algorithm.

Two more important points to mention are the choice of the starting point as well as the termination condition.

The starting point is chosen such that it is the same for all orthants of a given instance. The starting point depends on the weights of the trees of an instance and is defined as

$$x_s := \frac{W}{3} \cdot \mathbf{1} \in \mathbb{R}_+^{n-2}, \text{ with } W = \max_{i=1, \dots, M} \left\{ \max_{s \in \text{Split}(T_i)} w_s^{T_i} \right\}.$$

Obviously, there is no profound explanation to choose x_s other than that it is in the “middle” of the orthant, i.e., that it is equidistant to all hyperplanes $H_i = \{x \in \mathbb{R}^{n-2} : x_i = 0\}$. As W is the maximum weight of all splits it is clear that the values of $W \cdot \mathbf{1}$ would be far bigger than the ones of the optimal solution, so we scaled by a factor of 3. Of course, one might develop more sophisticated ways of choosing the starting point, but good choices heavily depend on the instance. Thus, we simply chose to fix some simplistic rule that determines the starting point and use it for all instances in order to get somewhat comparable results.

As the BPA is an iterative scheme we need to bound the number of steps it performs. Empirically we see that the BPA converges to a point in the orthant and does not cycle or jump, but it very rarely terminates at an actual balance point. We use two straightforward ways to terminate the BPA in a reasonable time or at a reasonable point. The first termination criterion is a cap on the number of maximal iterations, that we set to 500 for all experiments. The second termination criterion is that the BPA yields successive iterates whose distance is smaller than 10^{-9} , so we stop the algorithm when its progress gets marginal.

Note that we also increase efficiency of the preprocessing as follows: Theorem 5.2 implies that we only need to calculate z_{L3} , since it is always at least as good as z_{L1} and z_{L2} . Moreover, after having performed the bounds experiments, we decided to omit the bounds z_{U1}, z_{U2} , as z_{U1}, z_{U2}, z_{U3} have practically proven to yield rather similar results in average. We use z_{U3} as it is easy to obtain when we have solved **L3**. Additionally, we calculate z_{U4} , as this was the upper bound that fluctuated very little but showed good results, compare Section 5.1.

In order to interpret the results of the algorithms, we additionally calculate the objective of the star tree for every instance, and see if it yields a better solution than all other balance points. We did this since the star tree tends to be an optimal solution for very ‘heterogeneous’ or ‘spread’ data and we want to be able to detect,

when the star tree is better than any balance point solution and if so, if the BPA is converging towards it or towards a boundary of the orthant at least.

Just as the experiments in Section 5.1, the calculations have been executed on a single core of an Intel(R) Core(TM) i3-2350M and 4GB RAM.

Experiments in \mathcal{T}_4

In \mathcal{T}_4 , we conduct two different kinds of experiments. First, we investigate the behavior of the heuristic on small problems for specific types of instances, after that we investigate the computation times and the performance on instances of bigger sizes.

The reason to start with specific types of instances is to showcase how the preprocessing works and how the heuristic behaves in general. To this end we investigate characteristics as the computation time, the size of the gap to the calculated lower bound and the average amount of orthants that remain after the preprocessing. These trends are already present for small, easily interpretable instances with $M = 10$ sample trees.

The four types of instances we want to look at are the following: The first two types are the neighboring orthant ('NO') case and the completely incompatible ('CI') case. We have already defined these cases and developed optimal solution strategies for those in Chapter 4.

The third type of instance is another interesting special case, where data is contained in a triple of neighboring orthants, what we call the one-compatible-edge ('1C') case. In this case, we consider three orthants $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3$ with split sets S_1, S_2, S_3 , respectively, such that

- i) \mathcal{O}_1 and \mathcal{O}_2 are neighboring orthants
- ii) \mathcal{O}_2 and \mathcal{O}_3 are neighboring orthants
- iii) There exists one split $s \in S_1$ and one split $s' \in S_3$ such that s and s' are compatible. For all other pairs $\tilde{s} \in S_1, \tilde{s}' \in S_3$, \tilde{s} is incompatible with \tilde{s}' .

We have already seen such a case in Figure 2.12, when looking at examples for geodesics in \mathcal{T}_4 but we depict the situation once more in Figure 5.4.

Note that the upper left orthant in Figure 5.4 is not the image of an orthant of tree space due to the incompatibilities of the splits.

The last type of instance we look at is where trees are randomly spread in \mathcal{T}_4 , i.e., we simply sample random trees from all 15 orthants of \mathcal{T}_4 .

The results of the GBPH on these four types of instances are summarized in Table 5.6.

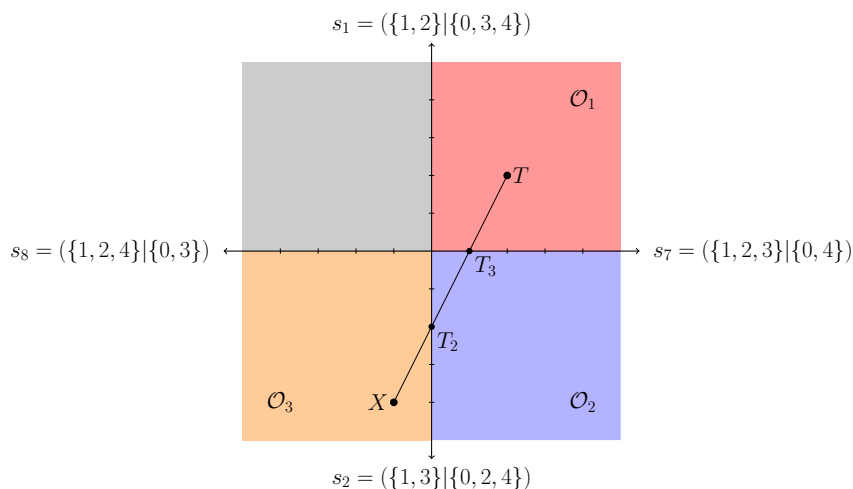


Figure 5.4: The depiction of the image of an isometric embedding of the three orthants that yield a “one-compatible-edge” case with an exemplaric geodesic.

Instance Type	\emptyset time in s	\emptyset orthants	\emptyset gap in %	star tree best
NO	12.5	1.5	12.6	0/10
CI	0.1	0	0	0/10
1C	11.9	1.8	15.7	0/10
RND	102.0	7.7	5.6	9/10

Table 5.6: Averaged results of the GBPH on 10 instances of neighboring orthant sampled sets of 10 trees in \mathcal{T}_4 .

First of all, we note some general trends of the results. One immediately sees, that the preprocessing is immensely helpful in determining the special structures of the instances. The average number of orthants that remained after the preprocessing is very low for the cases where the trees are only sampled in two or three orthants - in fact, the average number of remaining orthants is notably smaller than the number of orthants in which the trees have been sampled. Another relation the experiments emphasize is that the average computation time is strongly dependent on the average number of orthants. This is not surprising at all, since the computational effort basically reduces to the BPA: The preprocessing is actually saving a huge computational effort by discarding orthants in which the BPA has to be applied and it terminates in a split second. After the preprocessing is done, the computational effort is to apply the BPA in each orthant. So it was to be expected that

the relation between computation time and the number of remaining orthants is rather proportionate. The last general observation is, that, except for the randomly generated instances, the star tree is never the optimal solution to the problem and the GBPH solutions are better. This is due to the specific types of instances, where it is very unlikely that the star tree is the optimal solution to the problem. For the randomly generated instances, however, it is not very surprising that the star tree was optimal 9/10 times. But in seven of the nine cases the GBPH found the star tree as the optimal solution and in the two remaining cases the relative gap of the solution of the GBPH to the star tree solution was $< 10^{-9}\%$, so the randomly generated instances have been solved really good as well.

We now proceed to interpret the specific results of the experiments for each type of instance.

For the ‘NO’-case, the split preprocessing yields exactly the two neighboring orthants as remaining orthants, as it identifies the $n - 3$ common splits and then only two orthants are left that have splits that are contained in trees. The bound preprocessing is executed after this. If a lot of trees are contained in one orthant, the calculated upper bound of that particular orthant is actually lower than the upper bound of the other orthant in 50% of the cases and the other orthant gets discarded. Nonetheless, the average gap is rather high for these instances, which is probably due to the gap of the lower and upper bound in the orthant containing the optimal solution.

A very interesting case is the ‘CI’-case; it is actually solved by the preprocessing. This is because $z_{L3} = z_{U3}$, thus the optimization problem **L3** actually yields the optimal solution in each orthant. Thus, every orthant that does not contain the optimal solution is discarded and the optimal solution to the problem is the optimal solution to **L3**. The preprocessing is actually implemented such that the orthant discards itself as its lower bound is greater or equal to its upper bound and it returns the optimal solution of **L3** in the orthant with the lowest upper bound. So the GBPH finds the optimal solution in the preprocessing step which implies that 0 orthants remain in average.

The ‘1C’-case is slightly more general as the ‘NO’-case and it is not directly solvable by a Euclidean location problem, as a geodesic from \mathcal{O}_1 to \mathcal{O}_3 may actually be the cone path or a straight line, so that a simple embedding into $(\mathbb{R}^2, \|\cdot\|_2)$ may not solve this case. Even though trees are sampled within three orthants (and four different splits, which is relevant for the split preprocessing), the preprocessing reduced the number of orthants that may contain an optimal solution to 1.8 in average.

The random case does not carry any specific structure we could investigate here. Note, that the trend from the experiments for the bounds in Section 5.1 continues here, as the gap tends to get low, when the instance is very spread across the tree space. Since the number of trees was only ten, i.e., not every of the 15 orthants

contains a tree, the preprocessing still works pretty well and reduces the number of orthants by about 50% in average.

Note that fluctuations in computation times arises from the varying number of iterations the BPA performs before terminating: In each iteration of the BPA $M = 10$ geodesic parametrizations are calculated and a Euclidean median problem is solved. Hence, an orthant for which the BPA performs the maximal amount of iterations takes significantly longer than in an orthant where the BPA converges to an optimal solution in the interior and rather quickly terminates to the precision threshold.

Table 5.6 summarizes several relevant characteristics of the of the GBPH on these instances, but we have, except for the gap, no verification, how good the results are. Thus we compare the results of the GBPH with the results of the PPA in another run on the same instances (we saved the randomly generated instances for this experiment). Recall, that the PPA converges to the optimal solution. The PPA was set to terminate after 5000 cycles and was started at the star tree (which is a good heuristic starting point, as it is equally close to all orthants, and hence never too far away from an optimal solution).

It turns out that the GBPH yields a better objective value than the PPA in *all* instances. Of course, the solution of the PPA could be improved by letting it run for more cycles, but 5000 is already a respectable number of cycles and the progress of the algorithm is rather slow. We will discuss this behavior of the PPA in more detail in Section 6.2.1. Nonetheless, after 5000 cycles, the PPA yields a very good approximate solution to the median problem.

The distance between the trees the PPA and the GBPH yield was usually very low and deviated around 10^{-1} . Surprisingly, the maximal distance between the determined trees of the two methods was as high as 4.2. However, the objective of the PPA was only $\frac{620.23-620.10}{620.10} \approx 0.02\%$ higher than the objective of the GBPH, even though it was quite far away from the BPA solution. Altogether, we see that the GBPH may in fact compete with the PPA, especially in terms of solution quality, even though we have no convergence theorem for the GBPH, yet.

Now we turn to the aforementioned second experiment in \mathcal{T}_4 , concerning the computation times. Here we do not compare the results to the PPA, as we are only interested in the computation times of the algorithm.

Theoretically we would expect the computational effort of the algorithm to grow linearly in M , as the Weiszfeld-algorithm iterations grow linearly with M , as well as the number of geodesic parametrizations that need to be calculated. In order to keep computational effort reasonable, we perform the experiments for the one-compatible-edge case, where one is usually left with two to three remaining orthants after the preprocessing. The choice of the one-compatible-edge case does not also

limit the number of orthants in which we have to apply the BPA, it also offers a setting where the median will most likely not be the star tree in all cases. The results (averaged over 100 instances) of this may be found in Table 5.7.

number of trees	\emptyset time in s	\emptyset orthants	max. iterations
5	5.1	1.77	22
10	17.66	2.22	39
20	21.7	2.48	24
40	28	2.71	9

Table 5.7: Averaged results over 100 instances of the GBPH in \mathcal{T}_4 . We see monotone behavior of the computation time with respect to the number of trees per instance, which is strongly effected by the number of orthants for which the BPA reaches the maximal amount of iterations.

First of all, we recognize the general trend that the computational effort increases with the number of trees. The theoretical explanation would suggest that it should grow proportional with the number of trees. Nevertheless, the results show that there are other indicators that may disrupt a clear proportional behavior. The column to the most right contains the number of orthants for which the BPA has reached its maximal amount of 500 iterations. As explained earlier, this strongly effects the computation time, as the BPA often times stops at about iteration 40 to 50 when it reaches the precision threshold, so the number of iterations is more than 10 times higher. This explains why the factor of the computation time from the $M = 5$ tree instances to $M = 10$ tree instances is so high. Even though we calculated 100 instances for each number of trees, it seems as though there was an usual high number of ‘unlucky’ cases for 10 tree instances where the BPA reached the maximal amount of iterations.

On the other hand we see a pretty good example of the proportional behavior for the instances for 5 and 20 trees. The average computation time has increased only slightly more than by the expected factor of 4. We conjecture that this is due to the number of maximal iteration instances to be rather equal and that the slight increase comes from the additional average orthants for which the BPA has to be applied, since the preprocessing is less effective with more trees present.

Even though many of the above-mentioned arguments for the influences on the computation times are a little bit speculative it seems clear that the computational effort does not explode with an increased number of trees. Contrary to the expectation, the computational effort seems to grow a little less than proportionally. This tendency creates the hope, that bigger, real life instances may be tackled with the BPA and maybe also the GBPH. For the BPA we show the application to a real data set

in Chapter 6.

We summarize and interpret the key observations of the conducted experiments. First of all we noticed a tendency that the BPA reaches the maximal number of 500 iterations when it converges to boundary, but very rarely when it converges to a point in the interior. In the latter case it terminates due to the precision threshold. This is particularly helpful as it allows us to drastically reduce the computational effort for smaller instances by simply reducing the maximal number of iterations. This way, the boundary solutions that yield bad objective values anyways are not calculated too precisely. The downside of this would be that the result of the GBPH may get worse when the optimal solution is contained in the boundary of an orthant, as the BPA may perform less iterations and thus yields a worse result.

In the cases where the BPA terminated due to the precision threshold and has found a solution in the interior of the orthant, this balance point was always the best balance point. Combined with the fact that the GBPH always yielded a better solution than the PPA in the first experiment, we might conjecture that the GBPH actually yields an approximate optimal solution in this case. Hence, there is hope that the BPA converges to an optimal solution, when starting it in the orthant that contains a median in its interior. This empirical observation has led to a convergence analysis of the BPA, when started in the orthant of the optimal solution, see Section 5.3.

Experiments in \mathcal{T}_5 and \mathcal{T}_6

In this section we investigate the GBPH in \mathcal{T}_5 . The behavior of the GBPH did not change in comparison to \mathcal{T}_4 in preliminary test runs, it simply took a lot longer, as the number of orthants is 7 times higher. Thus we restrict ourselves to only two parameter setting with 5 instances each, but we consider somewhat bigger instances to see how the heuristic performs on somewhat bigger problems.

sample	orthants	trees	\emptyset time in s	\emptyset orthants	\emptyset gap in %
5		20	367.3	15.4	7.36
20		50	2503	78.6	1.2

Table 5.8: Results of the GBPH on in \mathcal{T}_5 for two differently spreading sample settings, both averaged over 5 instances.

We sum up the main insights of the experiments, as for \mathcal{T}_4 , but keep it short as the general trends coincide. The gap is especially low, when the star tree is found to be the best solution. This happens frequently, when the data is widely spread across

\mathcal{T}_5 ; 20 sample orthants of the 105 total orthants already suffice to depict this trend. On the other hand, the gap tends to be high when the solution of the GBPH is contained in the interior of an orthant.

As for the computational effort, we see that the preprocessing still plays a major role in keeping it reasonable, as it rules out many of the 105 orthants. The computation time in a single orthant is not too high, as the BPA only takes about 20 to 30 seconds in average to determine an optimal solution in an orthant. So it definitely takes longer than in \mathcal{T}_4 for a similar amount of trees, but the main problem is not the computation time of the BPA in a single orthant rather than the sheer amount of orthants in which it has to be applied. We continue to investigate this trend of computational effort in \mathcal{T}_6 .

We have just seen that, depending on the spread or heterogeneity of the sample trees, the GBPH can already take a rather long time to solve an instance. Since \mathcal{T}_6 has $9 \cdot 7 \cdot 5 \cdot 3 = 945$ orthants, Theorem 2.12, which is 9 times as many as \mathcal{T}_5 , we only solve two instances with 50 trees each, as the repetition of these experiments and averaging over them takes too long.

As we have 945 orthants and consider 50 trees, we start to challenge the algorithm in terms of computational effort.

Similarly to \mathcal{T}_5 , we take one widely spread and one rather ‘centered’ instance. The trees of the first instance are randomly sampled from all orthants and the 50 trees of the other instance are sampled in 5 randomly picked orthants.

Instance	remaining orthants	time in s	gap in %
spread	373	19100	0.66
centered	16	520	9.45

Table 5.9: Results of the GBPH on a widely spread sample set of 50 trees in \mathcal{T}_6 and of a more homogeneous data set of 50 trees.

After having conducted the experiments in \mathcal{T}_4 and \mathcal{T}_5 we do not have too much to add in terms of interpretation here, especially as the interpretation of two randomly generated data sets is not too reliable. We rather take note of the continuous trends, that have already been sketched in the other experiments. Again, the gap for the spread data set is exceptionally low. The BPA has determined the star tree as the best solution in a lot of orthants and the lower bounds for the orthants coincide with the objective value of the star tree, when 0 is the minimizer of **L3**. This happens quite frequently, but the global lower bound is the minimum of the lower bounds of all orthants and for some orthants 0 is not the the minimizer of **L3**, thereby creating a gap.

Moreover, we hint at the computation times which are on the one hand already quite time consuming with 5.3h for the spread instance. On the other hand this is still acceptable, as this translates to an average of about 50s of computation time for the BPA per orthant. This is still reasonable as for a real data set as in Chapter 6 one also only solves a single instance, so with enough time the GBPH should also terminate for data sets in $n = 7$. For $n = 8$, however, it gets close to impossible to even generate all orthants and do the preprocessing on them as the combinatorics start to kick in heavily at this point. But as we will see for the real data set, the BPA, in comparison to the GBPH, still works pretty well for $n = 8$ and should continue to do so for slightly larger n .

Evaluation of the GBPH

As a small resumee we point out one major advantage and one disadvantage of the GBPH. The disadvantage is that it is simply too slow for tree spaces \mathcal{T}_n with $n \geq 7$. The exponential amount of orthants could only be somewhat controlled by the preprocessing and the number of remaining orthants will be overwhelming for higher dimensions if the data sets do not have a very specific structure. Hence, it is necessary to combine the GBPH with additional heuristic rules or other methods. One possibility is to apply heuristic rules that determine which orthants may yield good solutions. This could for example be an orthant rating function, where an orthant gets a high rating if its splits are contained in many of the sample trees. Another idea is to radically change the approach and only use the BPA for a very specific orthant, namely the orthant that contains the solution of the PPA after applying it for a prespecified amount of cycles. We apply this *hybrid* algorithmic strategy in Chapter 6 on a real data set. Naturally, both of the above-mentioned strategies might yield cases, where the BPA is not applied in the orthant containing the optimal solution.

So the disadvantage of the GBPH is its limitation to tree spaces of rather low dimensions. The advantage of the GBPH is the BPA: the BPA did not take unreasonably long to converge when only applied in a single orthant, so it is likely that it still works well in higher dimensions. Moreover, it determined approximate optimal solutions (when applied in the orthant that contains them) for all cases where we verified it via the PPA. This is why we meticulously investigate the convergence behavior of the BPA, when applied in an orthant that contains an optimal solution, in the following section.

5.3 Local Convergence of the BPA

In Subsection 5.2.1 we introduced the BPA and have established that it does not cycle and that the objective of its iterates does not increase. In this section we have two goals: In Subsection 5.3.2 we show that the sequence of iterates actually yields a cluster point and that the sequence of the objective of the iterates converges. The second goal is to show optimality under certain conditions. If the BPA is started in an orthant that contains an optimal solution and if the cluster point X^* of the iterates satisfies certain, mild conditions, we can prove optimality of X^* , i.e., that X^* is a global minimizer of f . This is investigated in Subsection 5.3.3. In order to get these results we use the framework of the *Block-wise Coordinate Descent Method*, that we introduce in Subsection 5.3.1 where we also reformulate BPA in such a way that we can apply this method.

Note that we talk about *local* convergence of the BPA: naturally we can only get convergence to an optimal solution if we apply the BPA in an orthant $\mathcal{O} \subset \mathcal{T}_n$ that contains an optimal solution to (P) , the median problem in tree space.

5.3.1 The Block-wise Coordinate Descent Method

The block-wise coordinate descent method is a known algorithmic scheme that exists in several variants. Here, we adopt the setup from [War63].

Given Euclidean spaces E^1, \dots, E^n , let X_1, \dots, X_n be compact and convex sets in E^1, \dots, E^n , respectively. Moreover let $X = X_1 \times \dots \times X_n$ and $\varphi : X \rightarrow \mathbb{R}$ be a convex function. The goal is to find a global minimizer of φ , to which end [War63] introduces the block-wise coordinate descent method, see Algorithm 7.

Algorithm 7 returns $u^{(\ell)}$ when no further progress is made. These points are called *stationary points*:

Definition 5.19. *A point $u \in X$ that satisfies*

$$u_i \in \arg \min_{x \in X_i} \varphi \left((u_0, u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_M)^t \right) \quad \text{for } i = 1, \dots, n \quad (5.7)$$

is called a stationary point of φ .

In the following theorem we combine the central statements concerning the convergence of the block-wise coordinate descent method to a global minimizer:

Theorem 5.20 ([War63]). *Let X and φ be as above. Additionally, suppose that the following two assumptions hold:*

Data: $u^{(0)} = (u_1^{(0)}, u_2^{(0)}, \dots, u_n^{(0)}) \in X = X_1 \times \dots \times X_n$.

Result: A *stationary* point.

for $\ell = 0, 1, 2, \dots$ **do**

$$u_1^{(\ell+1)} = \arg \min_{x \in X_1} \varphi \left((x, u_2^{(\ell)}, \dots, u_n^{(\ell)})^t \right),$$

$$u_2^{(\ell+1)} = \arg \min_{x \in X_2} \varphi \left((u_1^{(\ell+1)}, x, u_3^{(\ell)}, \dots, u_n^{(\ell)})^t \right),$$

\vdots

$$u_n^{(\ell+1)} = \arg \min_{x \in X_n} \varphi \left((u_1^{(\ell+1)}, u_2^{(\ell+1)}, \dots, u_{n-1}^{(\ell+1)}, x)^t \right).$$

if $u^{(\ell+1)} = u^{(\ell)}$ **then**

return $u^{(\ell)}$

end

Algorithm 7: The block-wise coordinate descent method.

W1 φ is continuously differentiable in some neighborhood of every stationary point x of φ .

W2 For every $k = 1, \dots, n$, and every $\bar{x}_j \in X_j$ for $j = 1, \dots, n$, $j \neq k$ denote $\bar{x}_{-k} := (\bar{x}_1, \dots, \bar{x}_{k-1}, \bar{x}_{k+1}, \dots, \bar{x}_n)$. Then the function

$$\varphi_k(\bar{x}_{-k}) : X_k \rightarrow \mathbb{R}, \quad \varphi_k(\bar{x}_{-k})(x_k) = \varphi(\bar{x}_1, \dots, \bar{x}_{k-1}, x_k, \bar{x}_{k+1}, \dots, \bar{x}_n)$$

is strictly convex.

Then, any stationary point \tilde{x} of φ is a minimizer of φ over X . Moreover, when x' is an accumulation point of $(u^{(\ell)})_{\ell \in \mathbb{N}_0}$ (as in Algorithm 7), then x' is a minimizer of φ over X .

We use notations and definitions close to [War63]. Unfortunately we will see later, when defining our specific φ , that we neither have strict convexity nor do we have continuous differentiability over the whole set of definition. As this is required for Theorem 5.20, we cannot apply it. Thus, the remainder of the section is dedicated to adapting the analysis of the proofs of the convergence from [War63] with the prerequisites we have.

We first check, that we may restrict ourselves to a compact subset of the tree space to search for an optimal solution. Then, a single orthant is also restricted to a compact subset, which is isometric to a Euclidean compact subset as required for the X_i .

Theorem 5.21. *Let $\mathbf{T} = \{T_1, \dots, T_M\}$ be an instance of the median problem (P) in tree space. Then for any optimal solution $X^* = \{(s_1, \dots, s_k), (w_1, \dots, w_k)\}$ to (P) it holds that $w_i \leq U$ for $i = 1, \dots, k$, where $U = \max_{i=1, \dots, M} \{\max_{s \in \text{Split}(T_i)} w_s^{T_i}\}$.*

Proof. Assume that $X = ((s_1, \dots, s_l), (w_1, \dots, w_l)) \in \mathcal{T}_n$ and that at least one split s_{i_0} satisfies $w_{i_0} > U$. Now, define the tree $X' = ((s_1, \dots, s_l), (w'_1, \dots, w'_l))$, where the weights $w'_j = w_j$ for $j \neq i_0$ and $w'_{i_0} = U$. We want to show that $d(T_i, X') < d(T_i, X)$ for $i = 1, \dots, M$ implying that $f(X') = \sum_{i=1}^M d(T_i, X') < \sum_{i=1}^M d(T_i, X) = f(X)$ and X cannot be optimal. So let $i \in \{1, \dots, M\}$ be arbitrary.

First assume that s_{i_0} is a double compatible split of X and T_i , recalling Definition 2.18, and that $(\mathcal{A}, \mathcal{B})$ is the minimal support for the geodesic. Then $(\mathcal{A}, \mathcal{B})$ is also the minimal support for X' and T_i because the subtrees for which the support is determined does not depend on double compatible splits. Then the length formula from Theorem 2.21 applied to the two geodesics yields

$$\begin{aligned} d(X, T_i) &= \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^X - w_s^{T_i})^2} \\ &> \sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s^{X'} - w_s^{T_i})^2} = d(X', T_i), \end{aligned}$$

because all terms are the same except for the one corresponding to s_{i_0} in the sum for the double compatible splits. Since $w_{s_{i_0}}^{T_i} \leq U$ and $w_{s_{i_0}}^X = w_{i_0} > U$ it follows that $(w_{s_{i_0}}^X - w_{s_{i_0}}^{T_i})^2 > (U - w_{s_{i_0}}^{T_i})^2 = (w_{s_{i_0}}^{X'} - w_{s_{i_0}}^{T_i})^2$, which implies the strict inequality.

The other case is that s_{i_0} is not a double compatible split. Then let $\Gamma(X, T_i)$ be given by the path that corresponds to the sequence of trees $(Y_0, Y_1, \dots, Y_{k-1}, Y_k)$ with $Y_0 = X, Y_k = T_i$. Recall the canonical embedding $\chi: \mathcal{T}_n \rightarrow \mathbb{R}_+^N$ (from Definition 2.9) of trees into Euclidean space and define $y^j := \chi(Y_j)$ for $j = 0, \dots, k$. Then

$$d(X, T_i) = \sum_{j=1}^k \|y^j - y^{j-1}\|_2.$$

Our goal is to construct a path from X' to T_i that is shorter than that. To this end let $j_0 \geq 1$ be the smallest index such that $s_{i_0} \notin \text{Split}(Y_{j_0})$, i.e., $s_{i_0} \in \text{Split}(Y_j)$ for $j \leq j_0 - 1$. Then define the path (Y'_0, \dots, Y'_k) as follows: For $j \geq j_0$ set $Y'_j = Y_j$. For $j \leq j_0 - 1$ we define the trees Y'_j through their embeddings into \mathbb{R}_+^N , i.e., $Y'_j := \chi^{-1}(y'_j)$ with

$$y'_j := \begin{cases} y_{j_l}, & \text{if } l \neq k_0 \\ \frac{U}{w_{i_0}} y_{j_{i_1}}, & \text{if } l = k_0 \end{cases},$$

where k_0 is the coordinate of \mathbb{R}_+^N corresponding to s_{i_0} . Since $w_{i_0} > U$ it follows that $y'_{k_0} < y_{k_0}$ and moreover $y'_{k_0} > 0$ since $y_{k_0} > 0$. This yields

$$\begin{aligned} \|y'_j - y'_{j-1}\|_2 &= \sqrt{\sum_{l=1}^N (y'_{j_l} - y'_{j-1_l})^2} \\ &= \sqrt{\sum_{l=1}^{k_0-1} (y_{j_l} - y_{j-1_l})^2 + \left(\frac{U}{w_{i_0}}y_{j_{k_0}} - \frac{U}{w_{i_0}}y_{j-1_{k_0}}\right)^2 + \sum_{l=k_0+1}^N (y_{j_l} - y_{j-1_l})^2} \\ &< \sqrt{\sum_{l=1}^N (y_{j_l} - y_{j-1_l})^2} = \|y_j - y_{j-1}\|_2 \end{aligned}$$

for $j = 1, \dots, j_0$, because of the strict inequality

$$\left(\frac{U}{w_{i_0}}y_{j_{k_0}} - \frac{U}{w_{i_0}}y_{j-1_{k_0}}\right)^2 = \left(\frac{U}{w_{i_0}}\right)^2 \cdot (y_{j_{k_0}} - y_{j-1_{k_0}})^2 < (y_{j_{k_0}} - y_{j-1_{k_0}})^2.$$

Moreover $\|y'_j - y'_{j-1}\|_2 = \|y_j - y_{j-1}\|_2$ for $j = j_0 + 1, \dots, k$. Since $j_0 \geq 1$ there is at least one strict inequality so we get

$$d(X, T_i) = \sum_{j=1}^k \|y_j - y_{j-1}\|_2 > \sum_{j=1}^k \|y'_j - y'_{j-1}\|_2 \geq d(X', T_i),$$

because (Y'_0, \dots, Y'_k) is a path from X' to T_i and the geodesic distance is the infimum over the lengths of all paths from X' to T_i .

This shows $d(X, T_i) > d(X', T_i)$ for $i = 1, \dots, M$ since i was arbitrary. Thus

$$\sum_{i=1}^M d(X, T_i) > \sum_{i=1}^M d(X', T_i),$$

so X cannot be optimal. □

Now, in order to reformulate the BPA in a way that it fits into the framework of the block-wise coordinate descent method we introduce a new objective function. It depends on auxiliary variables, that model the representatives of facilities outside the orthants (compare Figure 5.3) that are determined by several auxiliary maps that we introduce later. Note that the representatives of the facilities will be the weight vectors of the exit points, see Definition 5.13, which we need for the BPA.

We consider elements u of $\mathbb{R}_+^{n'}$ with $n' = (n - 2) \cdot (M + 1)$ and denote them as a vector of *blocks*

$$u = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_M \end{pmatrix}, \quad \text{where each } u_i \in \mathbb{R}_+^{n-2}.$$

Recall, that by definition,

$$\mathbb{R}_+^{n'} = \underbrace{\mathbb{R}_+ \times \dots \times \mathbb{R}_+}_{n' \text{ times}},$$

but as we consider blocks here and do not want to extensively use the space consuming vector notation above, we interpret it as

$$\mathbb{R}_+^{n'} = \underbrace{\mathbb{R}_+^{n-2} \times \dots \times \mathbb{R}_+^{n-2}}_{M+1 \text{ times}},$$

so that we may write

$$u = (u_0, u_1, \dots, u_M) \in \mathbb{R}_+^{n-2} \times \dots \times \mathbb{R}_+^{n-2} = \mathbb{R}_+^{n'},$$

which is frequently used in the following.

In the notation of the block-wise coordinate descent method, each block u_i corresponds to an element in $X_i = \mathbb{R}_+^{n-2}$ and we have $M + 1$ of such blocks. Theorem 5.21 states that for an optimal solution $X^* = ((s_1, \dots, s_l), (w_1, \dots, w_l))$ in tree space, it holds that $w_i \leq U$ for $i = 1, \dots, l$. That implies that when we are in a maximal orthant \mathcal{O}^* that contains the optimal solution X^* , we may restrict ourselves to $[0, U]^{n-2} \subset \psi_{\mathcal{O}^*}$, since the optimal solution has to be contained in that subset. On a side note, the proof of Theorem 5.21 actually shows that it is possible to restrict to this compact subset for any maximal orthant \mathcal{O}' .

With this restriction to the compact subset we define our new objective function for the blockwise-coordinate descent method on the compact set $[0, U]^{n'}$:

$$\varphi : [0, U]^{n'} \rightarrow \mathbb{R}, \quad \varphi(u) = \sum_{i=1}^M d(\psi^{-1}(u_i), T_i) + \sum_{i=1}^M \|u_i - u_0\|_2. \quad (5.8)$$

Thereby we also restrict the auxiliary variables corresponding to the blocks $1, \dots, M$ to $[0, U]^{n-2}$, though this is no real restriction: Similar as to the proof of Theorem 5.21 one easily sees that the length w_i of the splits s_i of any point

$$Y = ((s_1, \dots, s_l), (w_1, \dots, w_l)) \in \Gamma(X, T_i)$$

satisfies

$$w_i \leq \max_{Z \in \{X, T_i\}} \left\{ \max_{s \in \text{Split}(Z)} \{w_s^Z\} \right\} \leq U,$$

when $X \in \psi^{-1}([0, U]^{n-2})$.

Hence, for the optimal solution X^* we also have that the representatives B_i , with $\psi^{-1}(B_i) \in \Gamma(X, T_i)$, on the boundary of the orthant only have weights of splits that are all bounded by U .

Of course, φ is not a randomly chosen function, but it rose by a specific construction idea that comes from the block-wise-coordinate descent approach. In each step of the method one fixes all blocks but one and searches for the minimizer for the variable block. So it is important to note that for all u_i with $\psi^{-1}(u_i) \in \Gamma(\psi^{-1}(u_0), T_i)$ and any $x \in \mathbb{R}_+^{n-2}$ the relation $d(\psi^{-1}(x), T_i) = d(\psi^{-1}(x), \psi^{-1}(u_i)) + d(\psi^{-1}(u_i), T_i) = \|x - u_i\|_2 + d(\psi^{-1}(u_i), T_i)$ holds, i.e., the triangle inequality is satisfied with equality. For all points that are not on the geodesic the triangle equality is strict. This is why points on the geodesic will automatically be chosen later, when minimizing for some block $i = 1, \dots, M$.

The following lemma describes this relation between f and φ .

Lemma 5.22. *For all $u = (u_0, \dots, u_M) \in \mathbb{R}_+^{n'}$ it holds that $\varphi(u) \geq f(\psi^{-1}(u_0))$ and equality holds if and only if for all $i \in \{1, \dots, M\}$ it holds that $\psi^{-1}(u_i) \in \Gamma(\psi^{-1}(u_0), T_i)$.*

Proof. First we show that the inequality holds. For $i = 1, \dots, M$ we have that $d(\psi^{-1}(u_0), T_i) \leq d(\psi^{-1}(u_0), \psi^{-1}(u_i)) + d(\psi^{-1}(u_i), T_i) = \|u_0 - u_i\|_2 + d(\psi^{-1}(u_i), T_i)$. So for the sum this inequality also holds and we get

$$f(\psi^{-1}(u_0)) = \sum_{i=1}^M d(\psi^{-1}(u_0), T_i) \leq \sum_{i=1}^M (\|u_0 - u_i\|_2 + d(\psi^{-1}(u_i), T_i)) = \varphi(u).$$

For the geodesic distance, the triangle inequality $d(X, T) \leq d(X, Y) + d(Y, T)$ holds with equality if and only if $Y \in \Gamma(X, T)$, because otherwise we would get a contradiction to the uniqueness of the geodesic $\Gamma(X, T)$. This implies that $f(\psi^{-1}(u_0)) = \varphi(u)$ holds if and only if $\psi^{-1}(u_i) \in \Gamma(\psi^{-1}(u_0), T_i)$ for all $i = 1, \dots, M$. \square

After we have introduced φ we check if the prerequisites of Theorem 5.20 hold. One important assumption is the onvexity of φ . Naturally it makes sense to hope for the block-wise coordinate descent method to converge to a global minimum if our objective function is convex; otherwise, the most likely outcome is to simply get stuck in a local minimum. In order to prove convexity of φ , we establish the notion of convexity on Hadamard-spaces.

Definition 5.23 ([Bac14b]). *Let (\mathcal{H}, d) be a Hadamard-space. We say that a function $f : \mathcal{H} \rightarrow (-\infty, \infty]$ is convex, if for each geodesic $\gamma : [0, 1] \rightarrow \mathcal{H}$, the function $f \circ \gamma$ satisfies*

$$f(\gamma(t)) \leq (1 - t) \cdot f(\gamma(0)) + t \cdot f(\gamma(1)),$$

for every $t \in (0, 1)$. We say f is strictly convex if the inequality is strict, whenever $\gamma(0) \neq \gamma(1)$.

Lemma 5.24 ([Bac14b]). *Let (\mathcal{H}, d) be a Hadamard-space. For any point $x_0 \in \mathcal{H}$, $d(\cdot, x_0)$ is a convex function.*

We come back to convexity of φ . The function φ consists of two different kinds of summands, which we separate in order to analyze them. Therefore we introduce for $i = 1, \dots, M$

$$\begin{aligned} g_i : \mathbb{R}_+^{n'} &\rightarrow \mathbb{R}, & g_i(u) &:= d(\psi^{-1}(u_i), T_i) \\ h_i : \mathbb{R}_+^{n'} &\rightarrow \mathbb{R}, & h_i(u) &:= \|u_i - u_0\|_2 \end{aligned}$$

and we may decompose φ into the parts $\varphi(u) = \sum_{i=1}^M g_i(u) + \sum_{i=1}^M h_i(u)$.

Proposition 5.25. *φ is a convex function.*

Proof. φ is given as the sum of the h_i and g_i and we prove the convexity of these functions. For any $i \in \{1, \dots, M\}$ we show that h_i is convex in $u = (u_0, \dots, u_M)$: Take $\lambda \in (0, 1)$ and $x, y \in \mathbb{R}_+^{n'}$. Then

$$\begin{aligned} h_i(\lambda x + (1 - \lambda)y) &= \|\lambda x_i + (1 - \lambda)y_i - \lambda x_0 - (1 - \lambda)y_0\|_2 \\ &\leq \|\lambda x_i - \lambda x_0\|_2 + \|(1 - \lambda)y_i - (1 - \lambda)y_0\|_2 \\ &= \lambda h_i(x) + (1 - \lambda)h_i(y). \end{aligned}$$

Now we show that g_i is convex in $u = (u_0, \dots, u_M)$. Take $\lambda \in (0, 1)$ and $x, y \in \mathbb{R}_+^{n'}$. Moreover, let $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ be the parametrization of $\Gamma(\psi^{-1}(x_i), \psi^{-1}(y_i))$. In a single orthant, the distance between two points is the Euclidean distance of the weight vectors and this implies that the geodesics are Euclidean shortest paths, i.e., line segments. Since $\psi^{-1}(x_i)$ and $\psi^{-1}(y_i)$ are in the same orthant, this implies that the weight vector of a point on the geodesic is simply the convex combination of x_i and y_i , $\psi(\gamma(1 - \lambda)) = \lambda x_i + (1 - \lambda)y_i$. Note, that $\gamma(0) = \psi^{-1}(x)$, so we have $\lambda x_i + (1 - \lambda)y_i$ and not $(1 - \lambda)x_i + \lambda y_i$. Now applying ψ^{-1} on both sides of $\psi(\gamma(1 - \lambda)) = \lambda x_i + (1 - \lambda)y_i$ yields $\gamma(1 - \lambda) = \psi^{-1}(\lambda x_i + (1 - \lambda)y_i)$. With this

$$\begin{aligned} g_i(\lambda x + (1 - \lambda)y) &= d(\psi^{-1}(\lambda x_i + (1 - \lambda)y_i), T_i) = d(\gamma(1 - \lambda), T_i) \\ &\leq \lambda d(\gamma(0), T_i) + (1 - \lambda)d(\gamma(1), T_i) \\ &= \lambda d(\psi^{-1}(x_i), T_i) + (1 - \lambda)d(\psi^{-1}(y_i), T_i) = \lambda g_i(x) + (1 - \lambda)g_i(y), \end{aligned}$$

where we applied convexity of $d(\cdot, T_i)$ on \mathcal{T}_n , which was established in Lemma 5.24. Thus, all summands of φ are convex and so is φ . \square

In order to prove optimality of our algorithm we need the directional derivatives of φ . These are easy to find using the total differentiability that we establish on a certain region of $\text{int}(\mathbb{R}_+^{n'})$.

Proposition 5.26. *φ is differentiable on $\text{int}(\mathbb{R}_+^{n'}) \setminus (V \cup W)$, where $V = \{u \in \mathbb{R}_+^{n'} : \exists i \in \{1, \dots, M\} : \psi^{-1}(u_i) = T_i\}$ and $W = \{u \in \mathbb{R}_+^{n'} : \exists i \in \{1, \dots, M\} : u_i = u_0\}$.*

Proof. As a sum of functions, φ is differentiable at all points $u \in \text{int}(\mathbb{R}_+^{n'})$ where each summand is differentiable, so we consider the functions g_i and h_i , $i = 1, \dots, M$ separately.

$h_i(u) = \|u_i - u_0\|_2$ is differentiable at $u \in \text{int}(\mathbb{R}_+^{n'})$ except for $W_i = \{u \in \text{int}(\mathbb{R}_+^{n'}) : u_i = u_0\}$. That implies that the sum of the h_i is differentiable on $\text{int}(\mathbb{R}_+^{n'})$ except on $\cup_{i=1}^M W_i = W$.

Now consider the g_i . Let us start with the easy case of $T_i \in \mathcal{O}$. Then $g_i(u) = d(\psi^{-1}(u_i), T_i) = \|u_i - \psi(T_i)\|_2$, which is differentiable for all $u_i \neq \psi(T_i)$. So in this case g_i is differentiable on $\text{int}(\mathbb{R}_+^{n'}) \setminus V_i$ with $V_i = \{u \in \text{int}(\mathbb{R}_+^{n'}) : u_i = \psi(T_i)\}$.

Suppose now that $T_i \notin \mathcal{O}$, so $\psi^{-1}(u_i)$ and T_i are not in the same orthant. To avoid double indexing, let $x = u_i \in \mathbb{R}_+^{n-2}$ and $\psi^{-1}(x) = X \in \mathcal{T}_n$ with $X = ((s_1, \dots, s_{n-2}), (x_1, \dots, x_{n-2}))$. Recall the parametrization of the geodesic and the maximal set of splits C that is compatible with $\text{Split}(X)$ and $\text{Split}(T_i)$. We prove continuous differentiability of the partial derivatives but we need to distinguish between two different types of coordinates x_k .

First, consider a component x_k , $k \in \{1, \dots, n-2\}$ such that the corresponding split $s_k \notin C$. Let $A_j(x)$ be the subset of a support ¹

$$(\mathcal{A}(x), \mathcal{B}) = ((A_1(x), \dots, A_k(x)), (B_1, \dots, B_k))$$

of the geodesic $\Gamma(X, T_i)$ that contains s_k . We write $A_l(x)$, to stress that the support and the values of $\|A_l(x)\|_2$ depend on x . In particular, $A_j(x)$ depends on x_k , whereas the other sets A_l , $l \neq j$ do not, as they do not contain s_k . Then we use the distance

¹[MOP15] have considered derivatives of the geodesic distance and have shown that the derivative is independent of the support that is chosen. This holds, because in the case that there are multiple possible supports one has that the fractions $\frac{\|A_i\|_2}{\|B_i\|_2} = \frac{\|A_{i+1}\|_2}{\|B_{i+1}\|_2}$ are identical, which leaves the derivative unchanged, as one may see in the resulting derivative.

formula from Theorem 2.21:

$$\begin{aligned}
\frac{\partial g_i}{\partial x_k}(x) &= \frac{\partial(d(\psi(x), T_i))}{\partial x_k}(x) = \\
&\frac{\partial \left(\sqrt{\sum_{l=1}^k (\|A_l(x)\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s \psi(x) - w_s(T_i))^2} \right)}{\partial x_k}(x) = \\
&\frac{1}{2d(\psi(x), T_i)} \cdot \frac{\partial \left(\sum_{l=1}^k (\|A_l(x)\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s \psi(x) - w_s(T_i))^2 \right)}{\partial x_k}(x) = \\
&\frac{1}{2d(\psi(x), T_i)} \cdot \frac{\partial (\|A_j(x)\|_2^2 + 2\|A_j(x)\|_2 \cdot \|B_j\|_2)}{\partial x_k}(x) = \\
&= \frac{2x_k \left(1 + \frac{\|B_j\|_2}{\|A_j(x)\|_2} \right)}{2d(\psi(x), T_i)} = \frac{x_k \left(1 + \frac{\|B_j\|_2}{\|A_j(x)\|_2} \right)}{d(\psi(x), T_i)}
\end{aligned}$$

Hence, the partial derivative of g_i at $u_i = x$ in direction x_k exists and is continuous if $\psi^{-1}(u_i) \neq T_i$.

For the second case consider a component x_k that corresponds to a split $s_k \in C$. Let the length of a split s in T_i be denoted by $w_s^{T_i}$ and for X denote w_s^X analogously. This means, that for $s_k \in C$, $w_{s_k}^X = x_k$ and for all other splits $s \in \text{Split}(X) \setminus \{s_k\}$, $w_s^X = x_j$ for some $j \neq k$, which is independent from x_k . Then, once again, the partial derivative for coordinate x_k can be derived via the distance formula from Theorem 2.21:

$$\begin{aligned}
\frac{\partial g_i}{\partial x_k}(x) &= \frac{\partial(d(\psi(x), T_i))}{\partial x_k}(x) = \\
&\frac{\partial \left(\sqrt{\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s \psi(x) - w_s(T_i))^2} \right)}{\partial x_k}(x) = \\
&\frac{1}{2d(\psi(x), T_i)} \cdot \frac{\partial \left(\sum_{l=1}^k (\|A_l\|_2 + \|B_l\|_2)^2 + \sum_{s \in C} (w_s \psi(x) - w_s(T_i))^2 \right)}{\partial x_k}(x) \\
&= \frac{2w_{s_k}(\psi(x)) - 2w_{s_k}(T_k)}{2d(\psi(x), T_i)} = \frac{x_k - w_{s_k}(T_k)}{d(\psi(x), T_i)}
\end{aligned}$$

Again, the partial derivative of g_i at $u_i = x$ in direction x_k exists and is continuous if $\psi(u_i) \neq T_i$. Thus we have that for all cases, the partial derivative at $x = u_i$ exists and is continuous if $\psi(u_i) \neq T_i$, i.e., g_i is differentiable on the set $\text{int}(\mathbb{R}_+^{n'}) \setminus V_i$ with $V_i := \{u \in \text{int}(\mathbb{R}_+^{n'}) : \psi(u_i) \neq T_i\}$. Then the sum of the g_i is differentiable on all of $\text{int}(\mathbb{R}_+^{n'})$ except for $\cup_{i=1}^M V_i =: V$. Altogether, φ is differentiable on $\text{int}(\mathbb{R}_+^{n'}) \setminus (V \cup W)$. \square

Proposition 5.26 shows that the facilities T_i that are in orthant \mathcal{O} yield non-differentiability of φ . Because we want to check for optimality of a point using the gradient, we need to take care of these specifically. The following Lemma gives us an easily verifiable necessary condition for such a facility to be optimal. Recall, that in this whole section, \mathcal{O} is a fixed orthant that contains an optimal solution in which we model the BPA and for which ψ is defined, in particular \mathcal{O} , is not arbitrary.

Lemma 5.27. *Let $k \in \{1, \dots, M\}$ be such that $T_k \in \mathcal{O}$. Moreover, define $B_i = P_{T_i}(T_k)$ for $i = 1, \dots, M$. Then, if*

$$\text{Test}_k = \left\| \sum_{\substack{i=1 \\ i \neq k}}^M \frac{B_k - B_i}{\|B_k - B_i\|_2} \right\|_2 > 1,$$

where Test_k is defined as in Theorem 3.1, T_k is not optimal for (P).

Proof. We show that from optimality of T_k for (P) it follows that $\text{Test}_k \leq 1$. First note that, if T_k is optimal for (P), then B_k has to be optimal for the problem

$$\begin{aligned} \min \quad & \sum_{i=1}^M \|x - B_i\|_2 \\ \text{s.t.} \quad & x \in \mathbb{R}_+^{n-2}. \end{aligned}$$

Otherwise, if B_k is not optimal, then we can estimate for an optimal solution x

$$\begin{aligned} \sum_{i=1}^M d(\psi^{-1}(x), T_i) &\leq \sum_{i=1}^M \|x - B_i\|_2 + \sum_{i=1}^M d(\psi^{-1}(B_i), T_i) \\ &< \sum_{i=1}^M \|B_k - B_i\|_2 + \sum_{i=1}^M d(\psi^{-1}(B_i), T_i) = \sum_{i=1}^M d(T_k, T_i), \end{aligned}$$

which contradicts optimality of T_k .

Now, Theorem 3.1 implies that B_k is optimal to the unweighted median problem if and only if

$$\text{Test}_k = \left\| \sum_{\substack{i=1 \\ i \neq k}}^M \frac{B_k - B_i}{\|B_k - B_i\|_2} \right\|_2 \leq 1.$$

□

The problematic cases are where the algorithm stops at some facility $\psi^{-1}(T_k)$, where we can't check for optimality using the gradient. Nevertheless, in practice we may start the algorithm with several different starting points. When it always hits the facility $\psi^{-1}(T_k)$, this is an indication, that it is an optimal solution: Otherwise, if

$\psi^{-1}(T_k)$ was not the optimal solution then there exists a neighborhood \mathcal{N} around the optimal solution for which the objective value is less than for $\psi^{-1}(T_k)$, due to continuity of the objective. Since the BPA may not increase the objective, this guarantees that it cannot converge to $\psi^{-1}(T_k)$ for any starting point $y \in \mathbb{R}_+^{n'}$ such that $y \in \mathcal{N}$. This idea is later (Corollary 5.53) used to construct a convex neighborhood of the optimal solution for which the BPA converges.

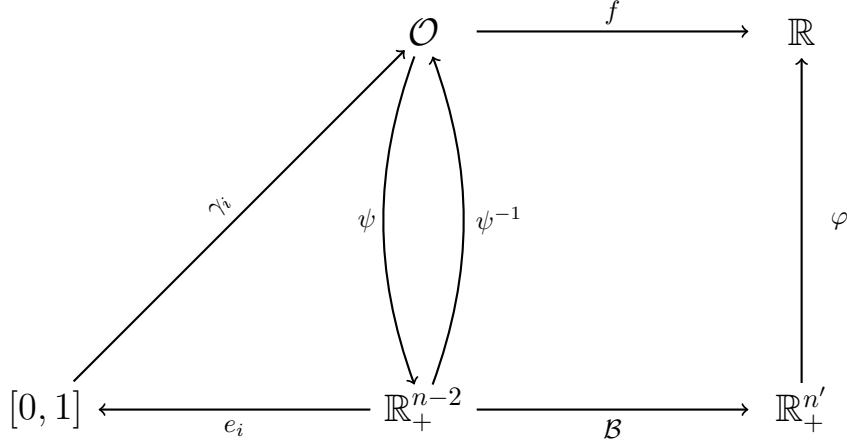


Figure 5.5: Commutative diagram for the most important involved maps for modeling the BPA with the block-wise coordinate descent method with the auxiliary map \mathcal{B} , where $\mathcal{B}(u_0) = (u_0, P_{T_1}(u_0), \dots, P_{T_M}(u_0))^t$.

After we have seen several features of the objective function φ we now come back to modeling the BPA using φ as objective for the block-wise coordinate descent method. We want to have representatives of all our facilities $\mathbf{T} \subset \mathcal{T}_n$ in the orthant \mathcal{O} . The easy ones are the $T_i \in \mathcal{O}$, where we can simply use $\psi(T_i)$ as a Euclidean representative. Now we define the representatives for $T_i \in \mathbf{T} \setminus \mathcal{O}$, so let i be an index such that $T_i \notin \mathcal{O}$. For $x \in \mathbb{R}_+^{n-2}$ let $\gamma_i(x) : [0, 1] \rightarrow \mathcal{T}_n$ denote the parametrization of the geodesic $\Gamma(\psi^{-1}(x), T_i)$ from $\psi^{-1}(x)$ to T_i . Then, we use the map (5.5), see Definition 5.13, and set $e_i(x) = e_{T_i}(x)$. This implies $\gamma_i(x)(\lambda) \in \mathcal{O}$ for $0 \leq \lambda \leq e_i(x)$ and $\gamma_i(x)(\lambda) \notin \mathcal{O}$ for $1 \geq \lambda > e_i(x)$.

Thus, it always holds that $\gamma_i(x)(e_i(x)) \in \partial\mathcal{O}$. Note, that whenever $x \in \text{int } \mathbb{R}_+^{n-2}$, it holds that $\Gamma(\psi^{-1}(x), T_i) \cap \partial\mathcal{O} = \{\gamma_i(x)(e_i(x))\}$, compare Figure 5.6 for an illustration of different 'cases' of exit points.

Now we define a map that unifies the *choice* of representatives for all $T_i \in \mathbf{T}$:

$$P_{T_i} : \mathbb{R}_+^{n-2} \rightarrow \mathbb{R}_+^{n-2}, \quad P_{T_i}(x) = \begin{cases} \psi(\gamma_i(x)(e_i(x))), & \text{if } T_i \notin \mathcal{O} \\ \psi(T_i), & \text{if } T_i \in \mathcal{O} \end{cases} \quad (5.9)$$

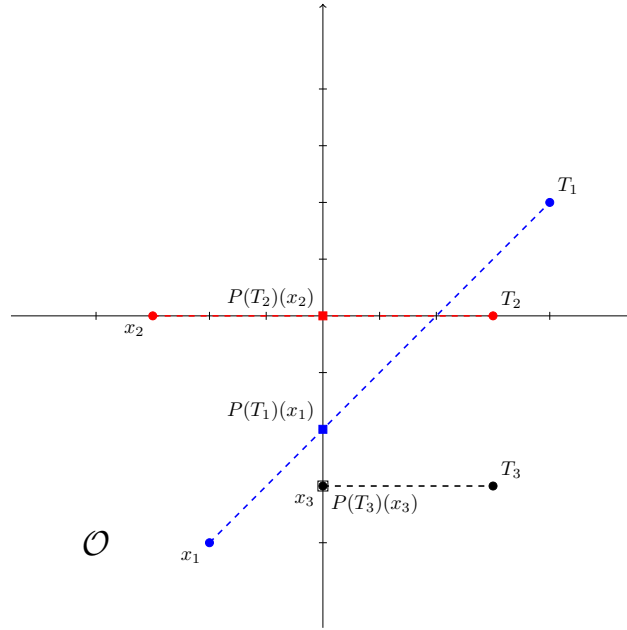


Figure 5.6: An example for several exit points (marked by squares) and a non-unique intersection with the boundary of \mathcal{O} (dashed red line).

Note, that we emphasize the term “choice” here. The problem is that we have to make a choice for the auxiliary variables for blocks $i = 1, \dots, M$ instead of having unique minimizers.

At the beginning of the section we noted that [War63] makes two specific assumptions in order to show that his algorithm converges to the optimal solution. In particular he supposes that **W2** holds that guarantees that the minimizers for each of the corresponding minimization subproblems, i.e., for each block, are unique. Our specific φ , (5.8), does not meet the assumption **W2**:

Take any point $u \in [0, U]^{(n-2) \cdot (M+1)}$. Then we define

$$\varphi_i(u_{-i}) : [0, U]^{n-2} \rightarrow \mathbb{R}, \quad \varphi_i(u_{-i})(y) = \varphi(u_0, \dots, u_{i-1}, y, u_{i+1}, \dots, u_M)$$

for $i = 0, 1, \dots, M$ as in **W2**.

It actually holds that all $x, y \in \psi(\Gamma(\psi^{-1}(u_0), T_i))$ are minimizers, because

$$\begin{aligned}
\varphi_i(u_{-i})(x) &= \varphi(u_0, \dots, u_{i-1}, x, u_{i+1}, \dots, u_M) \\
&= d(\psi^{-1}(x), T_i) + \|x - u_0\|_2 + \sum_{l \neq i} d(\psi^{-1}(u_l), T_l) + \sum_{l \neq i} \|u_l - u_0\|_2 \\
&= d(\psi^{-1}(u_0), T_i) + \sum_{l \neq i} d(\psi^{-1}(u_l), T_l) + \sum_{l \neq i} \|u_l - u_0\|_2 \\
&= d(\psi^{-1}(y), T_i) + \|y - u_0\|_2 + \sum_{l \neq i} d(\psi^{-1}(u_l), T_l) + \sum_{l \neq i} \|u_l - u_0\|_2 \\
&= \varphi(u_0, \dots, u_{i-1}, y, u_{i+1}, \dots, u_M) = \varphi_i(u_{-i})(y),
\end{aligned}$$

implying that the $\varphi_i(u_{-i})$ have no unique minimizers and are thus also not strictly convex. The strict convexity is not actually required for the proofs of [War63], he just demands it in order to ensure uniqueness of minimizers.

Nonetheless, we want to apply the iterative method from [War63], so we have to adapt the analysis to our situation. Because we do not have unique minimizers of φ with respect to its block coordinates, i.e., no unique solutions for the subproblems, we have to work around this in all proofs always keeping in mind that we made a specific *choice* for the minimizers with (5.9).

That we do not have unique minimizers is bad news for the analysis of the method. On the other hand, from an algorithmic point of view, the minimization steps for coordinates $1, \dots, M$ are trivial in our case: for a fixed i any point $x \in \mathbb{R}_+^{n-2}$ with $\psi^{-1}(x) \in \Gamma(u_0^{(\ell+1)}, T_i)$ is optimal. This is where we need to make a specific choice for the minimizer and we take the exit points $P_{T_i}(u_0^{(\ell+1)})$. We have to show convergence of the method with respect to this specific choice and cannot simply apply Warga's theorem.

Before we formulate the BPA in the framework of the block-wise coordinate descent method, we reformulate the block minimization problems by plugging in our explicit φ . Recall that minimization is performed with respect to a single block u_i , where $(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n)$ are fixed. The general versions of the minimization problems are:

$$\begin{aligned}
&\arg \min_{x \in X_1} \varphi((x, u_2, \dots, u_n)), \\
&\arg \min_{x \in X_2} \varphi((u_1, x, u_3, \dots, u_n)), \\
&\quad \vdots \\
&\arg \min_{x \in X_n} \varphi((u_1, u_2, \dots, u_{n-1}, x)).
\end{aligned}$$

Now, recall that $\varphi(u) = \sum_{i=1}^M d(\psi^{-1}(u_i), T_i) + \sum_{i=1}^M \|u_i - u_0\|_2$ and $X_i = \mathbb{R}_+^{n-2}$ for our blocks $i = 0, \dots, M$. Plugging this in we receive

$$\begin{aligned} & \arg \min_{x \in \mathbb{R}_+^{n-2}} \left(\sum_{i=1}^M d(\psi^{-1}(u_i), T_i) + \sum_{i=1}^M \|u_i - x\|_2 \right), \\ & \arg \min_{x \in \mathbb{R}_+^{n-2}} \left(d(\psi^{-1}(x), T_1) + \sum_{i=2}^M d(\psi^{-1}(u_i), T_i) + \|u_1 - u_0\|_2 + \sum_{i=2}^M \|u_i - u_0\|_2 \right), \\ & \quad \vdots \\ & \arg \min_{x \in \mathbb{R}_+^{n-2}} \left(\sum_{i=1}^{M-1} d(\psi^{-1}(u_i), T_i) + d(\psi^{-1}(x), T_M) + \sum_{i=1}^{M-1} \|u_i - u_0\|_2 + \|x - u_0\|_2 \right). \end{aligned}$$

Leaving out constants we get

$$\begin{aligned} & \arg \min_{x \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|x - u_i\|_2, \\ & \arg \min_{x \in \mathbb{R}_+^{n-2}} (d(\psi^{-1}(x), T_1) + \|x - u_0\|_2) \\ & \quad \vdots \\ & \arg \min_{x \in \mathbb{R}_+^{n-2}} (d(\psi^{-1}(x), T_M) + \|x - u_0\|_2). \end{aligned}$$

Then plugging in our choice P_{T_i} for the minimization problem with respect to the blocks $i = 1, \dots, M$ we are able to describe the BPA in the framework of the block-wise coordinate descent method, see Algorithm 8.

With the definition of φ and the choice of P_{T_i} we prove a simple, but important connection between balance points and stationary points:

Theorem 5.28. *$X \in \mathcal{O}$ is a balance point of f if and only if*

$$v = (\psi(X), P_{T_1}(\psi(X)), \dots, P_{T_M}(\psi(X)))$$

is a stationary point of φ .

Proof. Let $v = (\psi(X), P_{T_1}(\psi(X)), \dots, P_{T_M}(\psi(X)))$ for some $X \in \mathcal{O}$ be a stationary point. Then the stationary point condition (5.7) for our first block, $i = 0$, implies

$$\psi(X) \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \left(\sum_{i=1}^M \|y - P_{T_i}(\psi(X))\|_2 \right),$$

Data: $\mathbf{T} = \{T_1, \dots, T_M\}$, $x_0 \in \text{int } \mathbb{R}_+^{n-2}$.

Result: A stationary point.

Initialization $u^{(0)} = (x_0, P_{T_1}(u_0), \dots, P_{T_M}(u_0))^t$.

for $\ell = 0, 1, 2, \dots$ **do**

$$u_0^{(\ell+1)} = \arg \min_{x \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|x - u_i^{(\ell)}\|_2,$$

$$u_1^{(\ell+1)} = P_{T_1}(u_0^{(\ell+1)})$$

⋮

$$u_M^{(\ell+1)} = P_{T_M}(u_0^{(\ell+1)})$$

if $u^{(\ell+1)} = u^{(\ell)}$ **then**

return $u^{(\ell)}$

end

Algorithm 8: The block-wise coordinate descent method with the specific functions to model the BPA and the choice of $P_{T_i}(u_0)$ as minimizers for the blocks $i = 1, \dots, M$.

so X is a balance point by definition.

Now, let X be a balance point. Again, by definition,

$$\psi(X) \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \left(\sum_{i=1}^M \|y - P_{T_i}(\psi(X))\|_2 \right).$$

This implies that the stationary point condition (5.7) holds for $i = 0$. For $i = 1, \dots, M$, the stationary point condition holds as well, as $v_i = P_{T_i}(\psi(X)) \in \Gamma(X, T_i)$ and is thus a minimizer of $(d(\psi^{-1}(y), T_i) + \|y - \psi(X)\|_2)$, the optimization problem with respect to block i . □

Note, for the block-wise coordinate descent method in general, one is allowed to pick any starting point in X and converges to a minimizer of φ over X . Here we define a specific starting point $u^{(0)}$, whose 0-th coordinate x_0 is in the interior of \mathbb{R}_+^{n-2} . This is to model the BPA as it was defined in Algorithm 4 and it also avoids some unfortunate cases, where the BPA directly stops at a balance point at the boundary. Why boundary points are problematic will be described in detail in Section 5.4.

Moreover, we have set

$$u_0^{(\ell+1)} = \arg \min_{x \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|x - u_i^{(\ell)}\|_2.$$

But when the $u_i^{(\ell)}, i = 1, \dots, M$ are collinear, the minimizer is not unique, see Theorem 3.1. Later, we will rule out non-uniqueness, so that this does not matter for the proof of convergence. Until then, we choose a deterministic algorithm \mathcal{MA} that solves the Euclidean median problem, so that $u_0^{(\ell+1)}$ is well-defined by setting $u_0^{(\ell+1)} = \mathcal{MA}(u_1^{(\ell)}, \dots, u_M^{(\ell)})$.

With the formulation of the BPA in the block-wise coordinate descent method framework at hand we prove that this version actually is the same as the BPA, i.e., that Algorithm 4 and Algorithm 8 yield the same result.

Proposition 5.29. *Assume we are given the input data $\mathbf{T} = \{T_1, \dots, T_M\}$ and $x_0 \in \text{int } \mathbb{R}_+^{n-2}$. Let $u^{(\ell)}$ be the output of Algorithm 8 and $X^{(\ell)}$ be the output of Algorithm 4, where both methods use the same deterministic algorithm \mathcal{MA} to solve the Euclidean median problem. Then $\psi(X^{(\ell)}) = u_0^{(\ell)}$.*

Proof. The algorithms follow the exact same steps: For the given starting point x_0 , the minimizer $u_0^{(1)} = x_1$ of $\sum_{i=1}^M \|y - P_{T_i}(x_0)\|_2$ is calculated by \mathcal{MA} for both algorithms, then the exit points are calculated with respect to $u_0^{(1)} = x_1$ and the procedure is repeated. Hence, for each $j = 1, \dots, M$ it holds that $u_0^{(j)} = x_j$. When $x_l = x_{l-1}$, the BPA (Algorithm 4) stops. This implies that $u_0^{(l)} = u_0^{(l-1)}$, and that in turn implies that $u_i^{(l)} = P_{T_i}(u_0^{(l)}) = P_{T_i}(u_0^{(l-1)}) = u_i^{(l-1)}$ for $i = 1, \dots, M$. Thus we have $u^{(l)} = u^{(l-1)}$ and Algorithm 8 stops as well. The same argumentation can be made when Algorithm 8 stops, to get that Algorithm 4 stops as well.

This proves that the algorithms terminate in the same loop, and that $x_l = u_0^{(l)} = u_0^{(l-1)} = x_{l-1}$ holds. Algorithm 4 then returns $\psi^{-1}(x_{l-1})$, and $\psi(\psi^{-1}(x_{l-1})) = u_0^{(l-1)}$, which is the 0-th coordinate of the returned $u^{(l-1)}$ of Algorithm 8. \square

The following lemma is the analogue to Lemma 5.15 and points out the relation of the objective functions f and φ for the iterates of the algorithm. It is an immediate consequence of Proposition 5.29, as the algorithms perform the same iteration steps.

Lemma 5.30. *Let $u^{(\ell)}, u^{(\ell+1)}$ be successive iterates of Algorithm 8. Then*

$$f(\psi^{-1}(u_0^{(\ell+1)})) = \varphi(u^{(\ell+1)}) \leq \varphi(u^{(\ell)}) = f(\psi^{-1}(u_0^{(\ell)})).$$

With Proposition 5.29 we have shown that we are able to model the BPA in the framework of the block-wise coordinate descent method. Nonetheless, there are still a lot of problems to tackle - the block-wise coordinate descent method only yields

a subsequence converging to the optimal solution if we are able to work around the missing prerequisites: We have to show that the choice of P_{T_i} as optimal solution for the respective optimization problems does not create problems in the analysis as well as that the non-uniqueness of the Euclidean median problem for a given set of collinear points does not destroy convergence.

5.3.2 Convergence to a Stationary Point

After we have formulated the BPA as a variant of the block-wise coordinate descent method, compare Algorithm 8, we show that this algorithm yields a sequence whose cluster points are optimal solutions of φ under certain conditions. As mentioned at the end of Section 5.3.1, we need to adapt the analysis done by [War63] to our setting.

In the following we prove two important but very technical propositions (Proposition 5.39, Proposition 5.41), that are required in order to prove Theorem 5.44 that states that all cluster points of the iterates the BPA yields are stationary points. Since it is not straightforward to prove these two propositions we have to prove several auxiliary results beforehand. We present a graph of implications in Figure 5.7 that shows the structure of the results of this subsection:

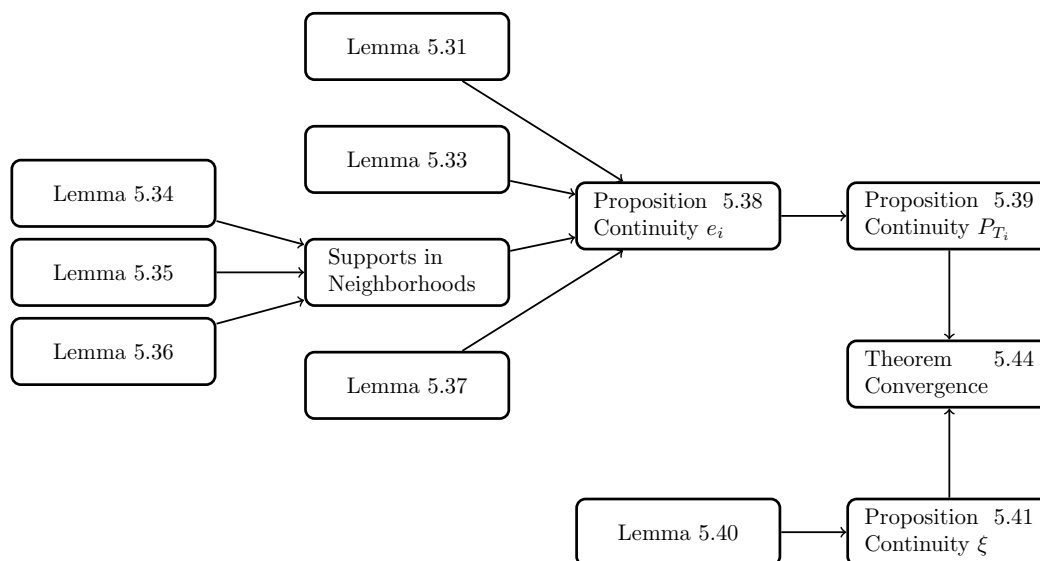


Figure 5.7: Results of Subsection 5.3.2 in a graph.

In both propositions we prove continuity of the maps that define the next iteration points $u_i^{(\ell+1)}$ for $i = 0, 1, \dots, M$. We start with continuity of P_{T_i} for $i = 1, \dots, M$, which is very tricky to prove. Before we present the proof of continuity we first need

to gain more insights about parametrizations of geodesics and how they change in certain neighborhoods.

Lemma 5.31. *Let $T \in \mathcal{T}_n \setminus \mathcal{O}$ and $x \in \partial\mathbb{R}_+^{n-2}$. Then X is the exit point, i.e., $\Gamma(X, T) \cap \mathcal{O} = \{X\}$ if and only if there exists a split $s \in \text{Split}(T) \setminus \text{Split}(\mathcal{O})$ such that $\text{Split}(X) \cup \{s\}$ is a compatible set of splits.*

Proof. First, let $s \in \text{Split}(T) \setminus \text{Split}(\mathcal{O})$ be compatible with all splits in $\text{Split}(X)$. Note that this is only possible when $|\text{Split}(X)| < n - 2$, which follows from $x \in \partial\mathbb{R}_+^{n-2}$. Otherwise we would get a set of compatible splits of cardinality greater than $n - 2$, which is impossible. Let $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ be the parametrization of the geodesic from X to T as in Theorem 2.21. Then $s \in C$ and the length of the split s changes linearly from 0 for $\gamma(0)$ to w_s^T , the weight of split s in T , for $\gamma(1)$. Hence for all $\lambda > 0$ it follows that $w_s^{\gamma(\lambda)} > 0$. Since $s \notin \text{Split}(\mathcal{O})$ it follows that $\gamma(\lambda) \notin \mathcal{O}$, so X is the exit point of the geodesic from \mathcal{O} .

For the other direction let X be the exit point, i.e., $\Gamma(X, T) \cap \mathcal{O} = \{X\}$. Theorem 2.21 implies that every tree on the geodesic from X to T only contains splits from $\text{Split}(X) \cup \text{Split}(T)$. If there was no split in $\text{Split}(T) \setminus \text{Split}(\mathcal{O})$ that was compatible with all splits in $\text{Split}(X)$, this would imply that a split of X has to be removed before adding another split of $\text{Split}(T)$. More formally, this means that $C \subset \text{Split}(X)$ and then the first leg of the geodesic $\gamma([0, \frac{\|A_1\|_2}{\|A_1\|_2 + \|B_1\|_2}])$ will have the splits $A_1 \cup \dots \cup A_k \cup C \subset \text{Split}(X) \subset \text{Split}(\mathcal{O})$, where $A_1 \neq \emptyset$ and $\|A_1\|_2 > 0$. Thus, $\{X\} = \{\gamma(0)\} \subsetneq \gamma([0, \frac{\|A_1\|_2}{\|A_1\|_2 + \|B_1\|_2}]) \subset \mathcal{O}$, which is a contradiction. \square

A useful concept that we use in the following proof is to partition a fixed orthant \mathcal{O} into sets that describe the minimal supports from all trees in \mathcal{O} to some given tree $T \notin \mathcal{O}$. The concept was introduced in [MOP15], Definition 3.3:

Definition 5.32. *Given a source tree $T \in \mathcal{T}_n$, a maximal orthant $\mathcal{O} \subseteq \mathcal{T}_n$, a support $(\mathcal{A}, \mathcal{B})$ and a set C of double compatible splits, let $\mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}, C)$ be the closure of the set of trees $X \in \mathcal{O}$ for which the geodesic joining X to T has support $(\mathcal{A}, \mathcal{B})$ satisfying (P2) and (P3) (see Definition 2.20) with strict inequalities and for which C is the set of double compatible splits of T and X . A previstal facet is any nonempty set $\mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}, C)$ of this form.*

In contrast to the original definition in [MOP15], we have added the set of double compatible splits C , because we handle these in another way than they are handled in [MOP15], where double compatible splits are taken care of in support pairs A_i, B_i with artificially introduced negative ratios $\frac{\|A_i\|_2}{\|B_i\|_2}$. This is only a technicality and we receive the same previstal facets, just with different notation.

Note, that a support satisfying (P2) with inequalities always is a minimal support as introduced in Definition 2.24. For trees in the interior of a previstal facet it is additionally required that (P3) holds with strict inequalities. This implies that $(\mathcal{A}, \mathcal{B})$ is

the unique support for the geodesic $\Gamma(X, T)$ for the trees $X \in \text{int}(\mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}, C))$. By definition a previstal facet contains all trees in the closure of this set of trees with the unique support $(\mathcal{A}, \mathcal{B})$ from X to T . Thus, it also contains all trees X for which $(\mathcal{A}, \mathcal{B})$ parametrizes the geodesic $\Gamma(X, T)$, but for which some inequalities in **(P2)** or **(P3)** are satisfied with equality.

It is important to note that the closure of some previstal facets also contains trees in the boundaries of the orthant \mathcal{O} with split set S , i.e., $\mathcal{O} = \mathcal{O}(S)$. In other words, a previstal facet $\mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}, C)$ may contain trees $Y \in \partial\mathcal{O}$ so that $\text{Split}(Y) \subsetneq S$. Then, $(\mathcal{A}, \mathcal{B})$ is not a support for such a tree Y , as some set $A_i \subset \text{Split}(\mathcal{O})$ contains a split \bar{s} that is not contained in $\text{Split}(Y)$ and logically this split cannot be removed from $\text{Split}(Y)$, as the support sequence $(\mathcal{A}, \mathcal{B})$ describes. We will establish what the supports look like for these specific cases in the following, but first we provide an easy example of previstal facets in \mathcal{T}_4 to get used to the concept.

Example 5.3.1. *We consider the same setup as in Example 2.2.3 with the following four splits for $n = 4$:*

$$\begin{aligned} s_1 &= (\{1, 2\}|\{0, 3, 4\}), & s_2 &= (\{1, 2, 3\}|\{0, 4\}), \\ s_3 &= (\{1, 3\}|\{0, 2, 4\}), & s_4 &= (\{1, 3, 4\}|\{0, 2\}) \end{aligned}$$

Now, let $\text{Split}(\mathcal{O}) = \{s_3, s_4\}$ and $T = ((s_1, s_2), (2, 2)) \in \mathcal{T}_4$. Then \mathcal{O} consists of two previstal facets:

$$\begin{aligned} \mathcal{V}_1 &= \mathcal{V}(T, \mathcal{O}; (\{s_3, s_4\}), (\{s_1, s_2\}), \emptyset) \\ \mathcal{V}_2 &= \mathcal{V}(T, \mathcal{O}; (\{s_4\}, \{s_3\}), (\{s_2\}, \{s_1\}), \emptyset) \end{aligned}$$

The situation is depicted in Figure 5.8, where one tree and its geodesic to T are given for each previstal facet. It is important to note that the trees in $\mathcal{V}_1 \cap \mathcal{V}_2$ may be parametrized using both supports. Even though they seem to yield different results, they actually do not. This is because some legs of the geodesic (in the example the one from X to T passing through the orthant having splits s_2, s_3) disappear because at the boundary $\frac{\|A_1\|_2}{\|B_1\|_2} = \frac{\|A_2\|_2}{\|B_2\|_2}$ holds, which implies that

$$\left[\frac{\|A_1\|_2}{\|A_1\|_2 + \|B_1\|_2}, \frac{\|A_2\|_2}{\|A_2\|_2 + \|B_2\|_2} \right) = \emptyset,$$

i.e., the leg passing through this orthant contains no points and the other two legs are joined at its boundary, which is the origin in this case.

In Figure 5.9 we show that the number of previstal facets that partition an orthant changes with the chosen orthant and that a previstal facet may even yield the whole orthant, as it is often the case in \mathcal{T}_4 . Suppose now that instead of $s_3 = (\{1, 3\}|\{0, 2, 4\})$ we have $s'_3 = (\{1, 4\}|\{0, 2, 3\})$ and all other splits remain

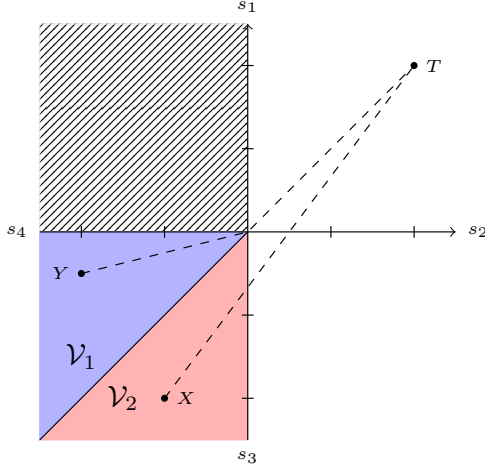


Figure 5.8: The orthant with splits $\{s_3, s_4\}$ is partitioned into two previstal facets with respect to $T = ((s_1, s_2), (2, 2))$.

the same. We now consider the orthant \mathcal{O}' with $\text{Split}(\mathcal{O}') = \{s'_3, s_4\}$. In this case all splits of T are incompatible with all splits in $\text{Split}(\mathcal{O}')$, hence there is only one previstal facet

$$\mathcal{V}'_1 = \mathcal{V}(T, \mathcal{O}'; (\{s'_3, s_4\}), (\{s_1, s_2\}), \emptyset).$$

Lemma 5.33. *Let a tree $T \in \mathcal{T}_n$ and any maximal orthant \mathcal{O} be given. Then \mathcal{O} is covered by a finite number of previstal facets.*

Proof. Proposition 3.6 in [MOP15] states that the so-called *vistal facets* cover \mathcal{T}_n . The vistal facets are the images of the previstal facets after squaring the coordinates of the weight vector of a tree. As the tree space orthants have non-negative weight vectors, this is a bijective map and it follows that the previstal facets cover \mathcal{T}_n as well.

So it only remains to show that the number of previstal facets covering \mathcal{O} is finite. To this end, consider a support $(\mathcal{A}, \mathcal{B}) = ((A_1, \dots, A_k), (B_1, \dots, B_k))$ with the set of double compatible splits C for $\Gamma(X, T)$. Since the A_i and B_i partition the split sets of the trees, $A_1 \cup \dots \cup A_k \cup C = \text{Split}(X)$ and $B_1 \cup \dots \cup B_k \cup C = \text{Split}(T)$, there is only a finite amount of such partitions. This implies that the number of different combinations of supports $(\mathcal{A}, \mathcal{B})$ and double compatible splits C is finite and thus, the number of previstal facets is finite as well. \square

Lemma 5.34. *Let $T \in \mathcal{T}_n$ and a maximal orthant $\mathcal{O} \subset \mathcal{T}_n$ be given and consider a non-empty previstal facet $\mathcal{W} = \mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}, C)$, with $\mathcal{A} = (A_1, \dots, A_k)$, $\mathcal{B} = (B_1, \dots, B_k)$.*

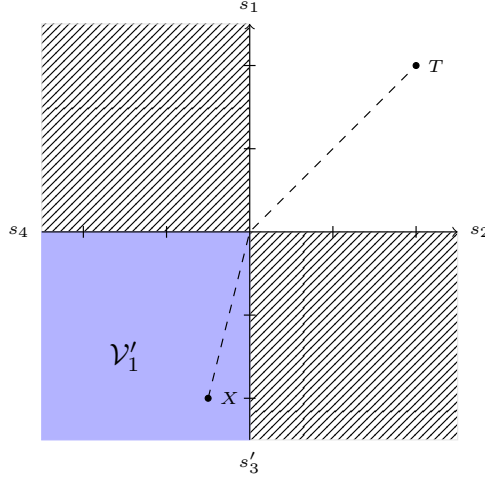


Figure 5.9: The orthant with splits $\{s'_3, s_4\}$ is consists of one previstal facet with respect to T .

Consider $X \in (\partial\mathcal{O}) \cap \mathcal{W}$ and define $\mathcal{A}^X = (A_1^X, \dots, A_k^X)$, where $A_i^X = A_i \cap \text{Split}(X)$. If $A_1^X \neq \emptyset$, then $(\mathcal{A}^X, \mathcal{B})$ is a support for $\Gamma(X, T)$.

Proof. We show that **(P1)**, **(P2)**, **(P3)** hold. Since $A_i^X \subset A_i$ it follows that $B_1 \cup \dots \cup B_{i-1}$ is compatible with A_i^X for each $i = 2, \dots, k \Rightarrow$ **(P1)**.

For **(P2)** we use that $X \in (\partial\mathcal{O}) \cap \mathcal{W}$, which implies that $X \in \partial\mathcal{W}$. This holds because $\text{Split}(X) \subsetneq \text{Split}(\mathcal{O})$, hence $(\mathcal{A}, \mathcal{B}), C$ cannot describe the geodesic from X to T , i.e., X is not in the interior of \mathcal{W} . Thus, there exists a sequence $X^1, X^2, \dots \subset \text{int}(\mathcal{W})$ such that $X^n \rightarrow X$. We have $\mathcal{W} \subset \mathcal{O}$, so we may apply our isometry $\psi : \mathcal{O} \rightarrow \mathbb{R}_+^{n-2}$ to get that for $x^n := \psi(X^n)$ we have $x^n \rightarrow x = \psi(X)$. In particular this implies that $x_j^n \rightarrow x_j$ for $j = 1, \dots, n-2$ and $x_j^n \rightarrow 0$ when coordinate j corresponds to a split $s \in \text{Split}(\mathcal{O}) \setminus \text{Split}(X)$. For $s \in \text{Split}(X)$ we may also write this as $w_{X^n}^s \rightarrow w_X^s$ and for the other splits $s \in \text{Split}(\mathcal{O}) \setminus \text{Split}(x)$, $w_{X^n}^s \rightarrow 0$. Then it holds for all $i = 1, \dots, k$ that

$$\begin{aligned} \|A_i^\eta\|_2 &:= \|A_i(X^\eta)\|_2 = \sqrt{\sum_{s \in A_i} w_{X^\eta}^s{}^2} = \sqrt{\sum_{s \in A_i \cap \text{Split}(X)} w_{X^\eta}^s{}^2 + \sum_{s \in A_i \setminus \text{Split}(X)} w_{X^\eta}^s{}^2} \quad (5.10) \\ &\rightarrow \sqrt{\sum_{s \in A_i^X} w_X^s{}^2 + 0} = \|A_i^X\|_2. \end{aligned}$$

Now, for each $\eta \in \mathbb{N}$ we have that $(\mathcal{A}, \mathcal{B})$ satisfies **(P2)** with inequalities, i.e., for

$i = 1, \dots, k - 1,$

$$\frac{\|A_i^\eta\|_2}{\|B_i\|_2} < \frac{\|A_{i+1}^\eta\|_2}{\|B_{i+1}\|_2},$$

and thus

$$\frac{\|A_i^X\|_2}{\|B_i\|_2} \leq \frac{\|A_{i+1}^X\|_2}{\|B_{i+1}\|_2}$$

for $i = 1, \dots, k - 1. \Rightarrow$ **(P2)**.

For **(P3)** we can use the exact same argumentation with the sequence X^η . **(P3)** holds with strict inequality for each partition $C_1^\eta \cup C_2^\eta$ of A_i^η , $D_1 \cup D_2$ of B_i with C_2 being compatible with D_1 , i.e.,

$$\frac{\|C_1^\eta\|_2}{\|D_1\|_2} < \frac{\|C_2^\eta\|_2}{\|D_2\|_2}.$$

Then it follows that for $c_i := \lim_{\eta \rightarrow \infty} \|C_i^\eta\|_2$ for $i = 1, 2$ we have

$$\frac{c_1}{\|D_1\|_2} \leq \frac{c_2}{\|D_2\|_2}$$

holds with equality in the limit. \Rightarrow **(P3)**. □

For the preceding lemma we supposed that $A_1^X \neq \emptyset$ holds. Now we investigate what happens when $A_j^X = \emptyset$ for some j .

Lemma 5.35. *Assume T, \mathcal{O}, X are given and \mathcal{A}^X defined as in Lemma 5.34. If $A_l^X = \emptyset$ for some $l \in \{1, \dots, k\}$, then $A_j^X = \emptyset$ for all $j = 1, \dots, l$.*

Proof. Again we use that $X \in \partial\mathcal{W}$ and let $X^1, X^2, \dots \subset \mathcal{W}$ be the sequence converging to X . In the proof of Lemma 5.34 (5.10), we have established that $\|A_i^\eta\|_2 \rightarrow \|A_i^X\|_2$ for $i = 1, \dots, k$, which holds for l in particular.

Now assume that there exists some $j < l$ such that $A_j^X \neq \emptyset$. That implies that $\|A_j^X\|_2 = c > 0$ and $\|A_j^\eta\|_2 \rightarrow c$, so there exists some $\eta' \in \mathbb{N}$ such that for $\eta > \eta'$ $\|A_j^\eta\|_2 > \frac{c}{2}$. Moreover, we have that $A_l^X = \emptyset$, so $\|A_l^\eta\|_2 \rightarrow 0$, hence there exists some $\tilde{\eta} \in \mathbb{N}$ such that for all $\eta > \tilde{\eta}$ it holds that $\|A_l^\eta\|_2 < \frac{c \cdot \|B_l\|_2}{2 \cdot \|B_j\|_2}$, which is equivalent to $\frac{c}{2} > \frac{\|A_l^\eta\|_2 \cdot \|B_j\|_2}{\|B_l\|_2}$, because $\|B_l\|_2, \|B_j\|_2 > 0$. Then for all $\eta > \max\{\eta', \tilde{\eta}\}$ we have that

$$\frac{\|A_j^\eta\|_2}{\|B_j\|_2} > \frac{c}{2 \cdot \|B_j\|_2} > \frac{\|A_l^\eta\|_2 \cdot \|B_j\|_2}{\|B_j\|_2 \cdot \|B_l\|_2} = \frac{\|A_l^\eta\|_2}{\|B_l\|_2}$$

with $j < l$ which is a contradiction to $(\mathcal{A}, \mathcal{B})$ being the support for X^η , because **(P2)** does not hold. □

Lemma 5.35 implies that there exists a maximal index $j_{\max} \in \{0, 1, \dots, k\}$ s.t. $A_i^X = \emptyset$ for $i \leq j_{\max}$. It is important to note that $j_{\max} = k$ is possible, however, this implies that $\text{Split}(X) \subset \text{Split}(T)$ and the support is trivial in this case.

Lemma 5.36. *In the situation of Lemma 5.34, let $j_{\max} \in \{1, \dots, k\}$ be the maximal index such that $A_{j_{\max}}^X = \emptyset$. Then for $i = 1, \dots, j_{\max}$, $B_i \subset C^X$ (the set of double compatible splits for T and X , depending on X) and (\tilde{A}^X, \tilde{B}) with $\tilde{\mathcal{A}}^X = (A_{j_{\max}+1}^X, \dots, A_k^X)$, $\tilde{\mathcal{B}} = (B_{j_{\max}+1}, \dots, B_k)$ is a support for the geodesic $\Gamma(X, T)$.*

Proof. If $j_{\max} = k$, then the minimal support $(\mathcal{A}, \mathcal{B})$ for $\Gamma(X, T)$ is trivial, $(\mathcal{A}, \mathcal{B}) = (\emptyset, \emptyset)$, since $\text{Split}(X) \subset \text{Split}(T)$ holds, which implies $\text{Split}(X) \subset C^X$. So suppose in the following that $j_{\max} < k$. Now, since $A_i^X = \emptyset$ for $i = 1, \dots, j_{\max}$, it follows that the splits in $B_1, \dots, B_{j_{\max}}$ are immediately added, because $B_1, \dots, B_{j_{\max}}$ are compatible with $A_{j_{\max}+1}, \dots, A_k$ and thus also with $A_{j_{\max}+1}^X, \dots, A_k^X \Rightarrow B_1 \cup \dots \cup B_{j_{\max}}$ are double compatible splits and are contained in C^X . For $i > j_{\max}$ we know that $A_i^X \neq \emptyset$ due to Lemma 5.35 and the choice of j_{\max} . The proof that (\tilde{A}^X, \tilde{B}) satisfies **(P1)**, **(P2)**, **(P3)** is the same as in the proof of Lemma 5.34. \square

With Lemmas 5.34, 5.35 and 5.36 we have established how supports in specific neighborhoods behave. This knowledge is crucial for the proof that the parameter $\lambda \in [0, 1]$ where a geodesic $\Gamma(X, T)$ parametrized by $\gamma_X : [0, 1] \rightarrow \mathcal{T}_n$ behaves continuously when X changes. In order to prove the continuity this parameter we need the following lemma.

Lemma 5.37. *For $t \geq t_{\min} > 0$ consider the family of functions*

$$P_{t_{\min}} = \{p_t\}_{t \geq t_{\min}} \quad \text{with } p_t : \mathbb{R}_+ \rightarrow [0, 1], \quad p_t(u) = \frac{u}{u+t}.$$

Then $P_{t_{\min}}$ is uniformly equicontinuous on \mathbb{R}_+ .

Proof. We need to show that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $\forall p_t \in P_{t_{\min}}, \forall u, v \in \mathbb{R}_+$ with $|u - v| < \delta$ it holds that $|p_t(u) - p_t(v)| \leq \varepsilon$.

In order to show that we prove that the functions are continuously differentiable on the domain \mathbb{R}_+ and thereby receive a Lipschitz-constant for each t :

$$p'_t(u) = \frac{t}{(u+t)^2} \quad \text{and} \quad \max_{u \in \mathbb{R}_+} p'_t(u) = \frac{1}{t} \quad \text{is attained at } u = 0.$$

Thus, we get a Lipschitz-constant of $\frac{1}{t}$ for each p_t . Now, let an $\varepsilon > 0$ be given and set $\delta = \varepsilon \cdot t_{\min}$. Then for all $t \geq t_{\min}$ and all $u, v \in \mathbb{R}_+$ with $|u - v| < \delta$ it holds that

$$|p_t(u) - p_t(v)| \leq \delta \cdot \frac{1}{t} \leq \varepsilon \cdot \frac{1}{t_{\min}} \leq \varepsilon,$$

which was to show. \square

Proposition 5.38. $e_i(x)$ is continuous.

Proof. Let $x \in \mathbb{R}_+^{n-2}$ and as always $\psi^{-1}(X) \in \mathcal{O}$. Let $\varepsilon > 0$ be given. First we specify $\delta > 0$ for which $|e_i(x) - e_i(y)| < \varepsilon$ holds for all $y \in B_\delta(x)$. δ is chosen as the minimum of three different values that guarantee different conditions. We define

$$\delta := \frac{1}{2} \cdot \min\{\varepsilon \cdot t_{\min}, d_v, x_{\min}\},$$

where

$$x_{\min} = \min_{i=1, \dots, n-2} \{x_i : x_i > 0\},$$

$$t_{\min} = \min_{j=1, \dots, n-2} \{t_{i_j} : t_{i_j} > 0\},$$

$$d_v = \min\{d(X, \mathcal{V}_i) : \mathcal{V}_i \subset \mathcal{O} \text{ is a previstal facet that does not contain } X\}.$$

For any $y \in B_\delta(x)$ we have that $\delta \leq \frac{x_{\min}}{2}$. This implies

$$|x_i - y_i| = \sqrt{(x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^{n-2} (x_i - y_i)^2} = \|x - y\|_2 < \delta \leq \frac{x_{\min}}{2}.$$

Suppose now that $y_i < \frac{x_{\min}}{2}$. Then $x_i \geq x_{\min} > y_i$ and with this

$$|x_i - y_i| = x_i - y_i \geq x_{\min} - y_i > x_{\min} - \frac{x_{\min}}{2} = \frac{x_{\min}}{2} \geq \delta > |x_i - y_i|,$$

which yields a contradiction. Hence, $y_i \geq \frac{x_{\min}}{2}$ for all i with $x_i > 0$ and this proves that each split in $\text{Split}(X)$ is also contained in Y . So $\delta \leq \frac{x_{\min}}{2}$ ensures that $\text{Split}(X) \subset \text{Split}(Y)$.

$\delta < d_v$ ensures that Y is not contained in a facet that does not also contain X and t_{\min} is the lower bound for the family of functions p_t as in Lemma 5.37, that we apply for $e_i(x) > 0$ because then $e_i(x) = \frac{\|A_1\|_2}{\|A_1\|_2 + \|B_1\|_2}$ and $\|B_1\|_2 \geq t_{\min}$ by definition, since $B_1 \subset \text{Split}(T)$.

We divide the proof into three parts for different locations of X . Lemma 5.33 states that the previstal facets cover \mathcal{O} , but the inequalities for (P2) and (P3) are only strict, when a tree is contained in the interior of such a facet. That is why distinguish the following three cases:

- (1) $X \in \text{int}(\mathcal{W})$ for some previstal facet \mathcal{W}
- (2) $X \in \text{int}(\mathcal{O})$ but \nexists a previstal facet \mathcal{V} such that $X \in \text{int} \mathcal{V}$
- (3) $X \in \partial \mathcal{O}$

(1): If X is contained in $\text{int}(\mathcal{W})$, then d_v guarantees that $Y \in \text{int}(\mathcal{W})$ as well, because $d(X, Y) = \|x - y\|_2 < \delta < d_v \leq d(X, \mathcal{V}) = \inf_{Z \in \mathcal{V}} d(X, Z)$ for any previstal facet \mathcal{V} other than \mathcal{W} . In order for the above to hold, we have to show that $d_v > 0$. Due to their definition we know that previstal facets are closed so when $X \in \mathcal{O}$ and $\mathcal{V} \subset \mathcal{O}$ such that $X \notin \mathcal{V}$, then $d(X, \mathcal{V}) = \inf_{Z \in \mathcal{V}} d(X, Z) > 0$. Moreover there only exists a finite amount of previstal facets in \mathcal{O} , it is for example bounded by the amount of all possible supports $(\mathcal{A}, \mathcal{B})$. Thus the minimum over this finite amount of distances $d_v = \min\{d(X, \mathcal{V}_i) : \mathcal{V}_i \subset \mathcal{O} \text{ is a previstal facet that does not contain } X\}$ is greater than 0.

Now, let $\mathcal{W} = \mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B})$, and since $X, Y \in \mathcal{W}$, $(\mathcal{A}, \mathcal{B})$ is the support for $\Gamma(X, T_i)$, $\Gamma(Y, T_i)$ respectively, where we denote the dependence of the parametrization on the weights in \mathcal{A} by $\mathcal{A}(x) = (A_1(x), \dots, A_k(x))$, $\mathcal{A}(y) = (A_1(y), \dots, A_k(y))$ respectively.

By definition $\|A_1(X)\|_2 = \sqrt{\sum_{s \in A_1} w_s^{X^2}}$ and

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n-2} x_i^2} = \sqrt{\sum_{s \in \text{Split}(X)} w_s^{X^2}}.$$

This also implies that

$$\sqrt{\sum_{i=1}^k \|A_i(x)\|_2^2} = \|x\|_2,$$

and moreover we have that

$$\delta > d(X, Y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^{n-2} (w_s^X - w_s^Y)^2}.$$

This implies that $|w_s^X - w_s^Y| < \delta$ for each split s in $\text{Split}(\mathcal{O})$, where $w_s^X = 0, w_s^Y$ if s is not in X or Y respectively.

Now, using the reversed triangle inequality and our inequalities established above we get

$$\begin{aligned} \left| \|A_1(x)\|_2 - \|A_1(y)\|_2 \right| &\leq \|A_1(x) - A_1(y)\|_2 = \sqrt{\sum_{s \in A_1} (w_s^X - w_s^Y)^2} \\ &\leq \sqrt{\sum_{i=1}^k \sum_{s \in A_i} (w_s^X - w_s^Y)^2} = \|x - y\|_2 \leq \delta \end{aligned}$$

Now we apply the uniform equicontinuity of the $\{p_t\}_{t \geq t_{\min}}$ where we use the map $p_{\|B_1\|_2}$ and $\|B_1\|_2 \geq t_{\min} > 0$ by definition. We have that $\left| \|A_1(x)\|_2 - \|A_1(y)\|_2 \right| \leq \delta$

so it follows that

$$|p_{\|B_1\|_2}(A_1(x)) - p_{\|B_1\|_2}(A_1(y))| < \varepsilon$$

and then

$$\begin{aligned} \varepsilon &> |p_{\|B_1\|_2}(A_1(x)) - p_{\|B_1\|_2}(A_1(y))| \\ &= \left| \frac{\|A_1(x)\|_2}{\|A_1(x)\|_2 + \|B_1\|_2} - \frac{\|A_1(y)\|_2}{\|A_1(y)\|_2 + \|B_1\|_2} \right| = |e_i(x) - e_i(y)| \end{aligned}$$

what was to show.

(2) Now, let $X \in \text{int}(\mathcal{O})$ but \nexists a previstal facet \mathcal{V} such that $X \in \text{int}\mathcal{V}$. We have seen above that the amount of previstal facets of \mathcal{O} is finite. Since previstal facets cover \mathcal{O} , there exist previstal facets \mathcal{V}_i such that $X \in \bigcap_{i=1}^{n_x} \partial\mathcal{V}_i$ with $n_x \geq 1$. For all other previstal facets \mathcal{V} with $X \notin \mathcal{V}$ we know again that $d(X, Y) < \delta < d(X, \mathcal{V})$, so $Y \in \bigcup_{i=1}^{n_x} \partial\mathcal{V}_i$ as well. That implies that there exists a $\mathcal{V}' = \mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B}) \in \{\mathcal{V}_1, \dots, \mathcal{V}_{n_x}\}$ such that $X \in \mathcal{V}'$ and $Y \in \mathcal{V}'$. Thus, both have support $(\mathcal{A}, \mathcal{B})$ where weights only differ in $\mathcal{A}(x), \mathcal{A}(y)$. The same argument as in **(1)** shows that then $|e_i(x) - e_i(y)| < \varepsilon$.

(3) The last remaining case is $X \in \partial\mathcal{O}$. In this case, we need Lemmas 5.34, 5.36 to get supports for X and Y . Just as in **(2)**, the choice of $\delta < d_v$ ensures that there exists a previstal facet $\mathcal{V}' = \mathcal{V}(T, \mathcal{O}; \mathcal{A}, \mathcal{B})$ that contains X and Y . We use Lemmas 5.34, 5.36 to get supports for X and Y respectively. $\text{Split}(X) \subset \text{Split}(Y)$ thereby guarantees that $A_i^X \subset A_i^Y$ for $i = 1, \dots, k$. We consider three cases:

- i) $A_1^X = \emptyset, A_1^Y = \emptyset$
- ii) $A_1^X = \emptyset, A_1^Y \neq \emptyset$
- iii) $A_1^X \neq \emptyset, A_1^Y \neq \emptyset$

The first case, i), is the easiest, since the splits $A_1 \subset \text{Split}(\mathcal{O})$ are neither contained in $\text{Split}(X), \text{Split}(Y)$. Moreover, $B_1 \cup A_2 \dots \cup A_k$ is a compatible set of splits with $B_1 \subset \text{Split}(T)$. Since $A_1^X = \emptyset, A_1^Y = \emptyset$ that implies that B_1 belongs to the set of double compatible splits and these splits are added immediately. Thus, $B_1 \subset \text{Split}(Z)$ for each $Z \neq X \in \Gamma(X, T_i)$ and $e_i(x) = 0$. The same holds for Y , thus $e_i(y) = 0$ and $|e_i(x) - e_i(y)| = 0 < \varepsilon$.

In the second case, ii), we have $e_i(x) = 0$, but $e_i(y) > 0$, because $A_1^Y \neq \emptyset$. $\Rightarrow e_i(y) = \frac{\|A_1^Y\|_2}{\|A_1^Y\|_2 + \|B_1\|_2}$. Now, all splits that are in A_1 are not in $\text{Split}(X)$, so the coordinates in \mathbb{R}_+^{n-2} that correspond to these splits are 0. Together with $\|x - y\|_2 < \delta$ this yields $\|A_1^Y\|_2 < \delta$ and thus $|0 - \|A_1^Y\|_2| = \|A_1^Y\|_2 < \delta$, so we can now use continuity of

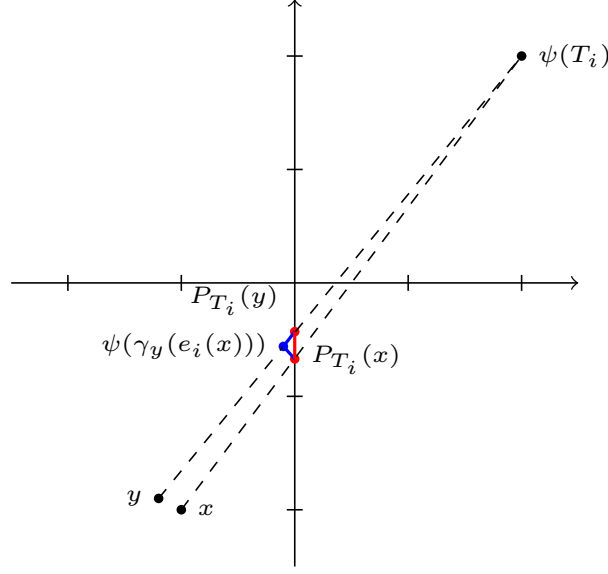


Figure 5.10: Illustration of the bound used in the proof of Proposition 5.39.

$p_{\|B_1\|_2}$, as in (1) :

$$|e_i(x) - e_i(y)| = \left| \frac{0}{0 + \|B_1\|_2} - \frac{\|A_1^Y\|_2}{\|A_1^Y\|_2 + \|B_1\|_2} \right| = |p_{\|B_1\|_2}(0) - p_{\|B_1\|_2}(A_1(y))| < \varepsilon.$$

In the last case, iii), we have that $e_i(x) = \frac{\|A_1^X\|_2}{\|A_1^X\|_2 + \|B_1\|_2} > 0$ and $e_i(y) = \frac{\|A_1^Y\|_2}{\|A_1^Y\|_2 + \|B_1\|_2} > 0$. Again, from $\|x - y\|_2 < \delta$ we get that $|\|A_1^X\|_2 - \|A_1^Y\|_2| < \delta$ and $|e_i(x) - e_i(y)| < \varepsilon$ follows from continuity of $p_{\|B_1\|_2}$ as in the other cases. \square

Next we prove that P_{T_i} is continuous, which is one of the two key propositions that we need for the proof of convergence for Theorem 5.44. The idea for this is depicted in Figure 5.10. Using the triangle inequality we bound the distance of $P_{T_i}(x)$ and $P_{T_i}(y)$ (red) by the sum of the lengths of the two blue line segments, whose lengths we can control.

Proposition 5.39. *For $i = 1, \dots, M$, P_{T_i} is continuous.*

Proof. In the case that $T_i \in \mathcal{O}$, we have that $P_{T_i}(y) = \psi(T_i)$ for all $y \in \mathbb{R}_+^{n-2}$, which is constant, hence continuous. Let now $T_i \notin \mathcal{O}$, so we prove continuity of $P_{T_i}(y) = \psi(\gamma_i(y)(e_i(y)))$. For this proof we make use of two properties of our geodesic and its parametrization:

- (i) For a fixed tree $T \in \mathcal{T}_n$ and a fixed parameter $t \in (0, 1)$ the map $\nu_t : \mathcal{T}_n \rightarrow \mathcal{T}_n$, $\nu_t(X) = \gamma_X(t)$ is continuous, where γ_X is the parametrization of the geodesic from X to T (Lemma 1.2.2. in [Bac14b]).

(ii) For the parametrization $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ of a geodesic from X_1 to X_2 it holds that

$$d(\gamma(s), \gamma(t)) = |s - t| \cdot d(X_1, X_2) \quad \forall s, t \in [0, 1],$$

see Theorem 2.23.

Now we choose $x \in \mathbb{R}_+^{n-2}$ and prove continuity of P_{T_i} at x : Let $\varepsilon > 0$ be given. Due to (i), there exists some $\delta_1 > 0$ such that $d(\nu_{e_i(x)}(X), \nu_{e_i(x)}(Y)) < \frac{\varepsilon}{2}$ for all Y with $d(X, Y) < \delta_1$. In a similar fashion we choose $\delta_2 > 0$, which exists due to the continuity of $e_i(x)$ that was established in Lemma 5.38: Let $\delta_2 > 0$ be such that $|e_i(x) - e_i(y)| < \frac{\varepsilon}{4d(T_i, X)}$ for all y with $\|x - y\|_2 < \delta_2$. Now we combine these estimates. Let $\delta = \min\{\delta_1, \delta_2, d(X, T_i)\}$ and let $y \in B_\delta(x) \cap \mathbb{R}_+^{n-2}$. Then

$$\begin{aligned} & \|P_{T_i}(x) - P_{T_i}(y)\|_2 = \|\psi(\gamma_X(e_i(x))) - \psi(\gamma_Y(e_i(y)))\|_2 \\ & \leq \|\psi(\gamma_X(e_i(x))) - \psi(\gamma_Y(e_i(x)))\|_2 + \|\psi(\gamma_Y(e_i(x))) - \psi(\gamma_Y(e_i(y)))\|_2 \\ & \stackrel{*}{=} d(\nu_{e_i(x)}(X), \nu_{e_i(x)}(Y)) + |e_i(x) - e_i(y)| \cdot d(Y, T_i) \\ & < \frac{\varepsilon}{2} + |e_i(x) - e_i(y)| \cdot (d(X, T_i) + \delta) \\ & \leq \frac{\varepsilon}{2} + |e_i(x) - e_i(y)| \cdot 2d(X, T_i) \\ & < \frac{\varepsilon}{2} + \frac{\varepsilon}{4d(X, T_i)} \cdot 2d(X, T_i) \\ & = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

where we use (ii) at the equality marked with *. This proves continuity of $P_{T_i}(x)$ for an arbitrary $x \in \mathbb{R}_+^{n-2}$. \square

Figure 5.11 depicts the structure of the subsection once more, where we marked all results that are proven until now. As we can see, all results in the upper part of the figure have successfully been proven and the continuity of the P_{T_i} is established. We now need to prove the continuous behaviour of ξ , which is the map yielding the minimizers of the Euclidean median problems in block 0 for given facilities u_1, \dots, u_M .

Since the Euclidean median problem may not have a unique solution when the given points are collinear, compare Example 5.3.2, we even have a problem with defining this map. Therefore, we make slight assumptions to rule out this unfortunate case and thereby force uniqueness of the minimizers of the Euclidean median problems for the relevant cases. We start with a lemma that will later on ensure uniqueness on a compact neighborhood.

Lemma 5.40. *Let $m \geq 3$ and $p_1, \dots, p_m \in \mathbb{R}^d$ be such that p_1, \dots, p_m are not collinear. Then there exists a $\delta > 0$ such that for all $v = (v_1, \dots, v_m)^t \in \mathbb{R}^{d \cdot m}$, $v \in B_\delta((p_1, \dots, p_m)^t)$ it holds that v_1, \dots, v_m are not collinear.*

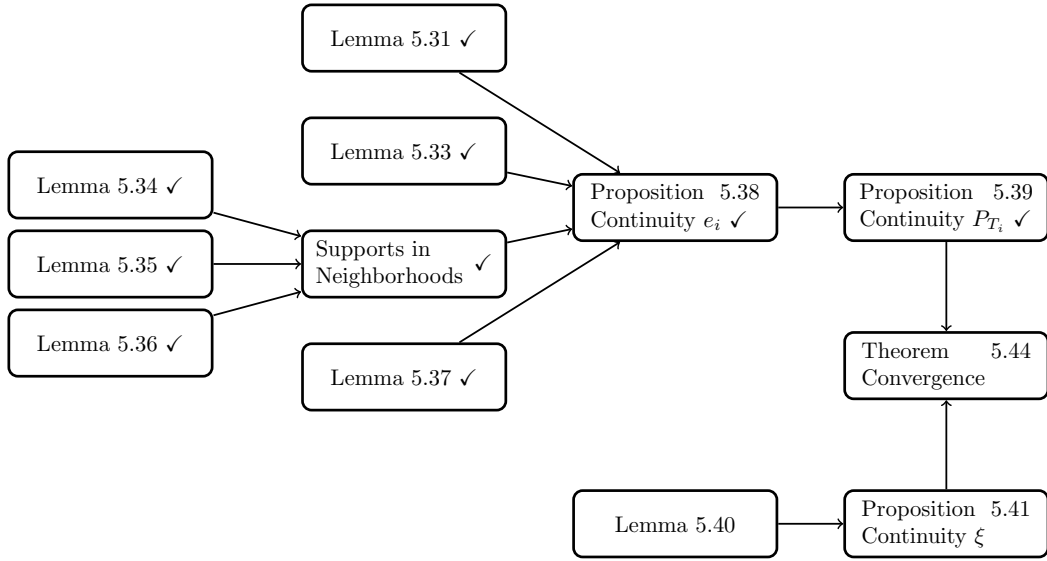


Figure 5.11: Progress of proven results of Subsection 5.3.2.

Proof. Suppose w.l.o.g. that the points p_1, p_2 satisfy

$$\|p_1 - p_2\|_2 \geq \|p_i - p_j\|_2 \quad \text{for } 1 \leq i < j \leq m.$$

Since p_1, \dots, p_m are not collinear, there exists at least one p_i , $i \geq 3$ that does not lie on the line through p_1 and p_2 . Thus, the triangle inequality is strict:

$$\|p_1 - p_2\|_2 < \|p_1 - p_i\|_2 + \|p_i - p_2\|_2.$$

Let $d = \|p_1 - p_i\|_2 + \|p_i - p_2\|_2 - \|p_1 - p_2\|_2 > 0$ and define $\varepsilon := \frac{d}{4}$, so that

$$\|p_1 - p_2\|_2 + 3\varepsilon = \|p_1 - p_2\|_2 + \frac{3}{4}d < \|p_1 - p_2\|_2 + d = \|p_1 - p_i\|_2 + \|p_i - p_2\|_2. \quad (5.11)$$

Now, since the norm $\|\cdot\|_2 : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a continuous function, there exists $\delta_1 > 0$ such that for all $(v_1, v_2)^t \in B_{\delta_1}((p_1, p_2)^t)$, it holds that $\|v_1 - v_2\|_2 \leq \|p_1 - p_2\|_2 + \varepsilon$. Analogously there exists such a $\delta_2 > 0$ for $(v_1, v_i)^t \in B_{\delta_2}((p_1, p_i)^t)$ and $\delta_3 > 0$ for $(v_2, v_i)^t \in B_{\delta_3}((p_2, p_i)^t)$. Then, setting $\delta' = \min\{\delta_1, \delta_2, \delta_3\} > 0$ we get that for any v_1, v_2, v_i satisfying $(v_1, v_2) \in B_{\delta'}(p_1, p_2)$, $(v_1, v_i) \in B_{\delta'}(p_1, p_i)$ and $(v_2, v_i) \in B_{\delta'}(p_2, p_i)$ it holds that

$$\|v_1 - v_2\|_2 \leq \|p_1 - p_2\|_2 + \varepsilon < \|p_1 - p_i\|_2 + \|p_i - p_2\|_2 - 2 \cdot \varepsilon \leq \|v_1 - v_i\|_2 + \|v_i - v_2\|_2, \quad (5.12)$$

where the strict inequality is implied by (5.11). This proves that $v_i \notin \{v_1 + \lambda \cdot (v_2 - v_1) : \lambda \in \mathbb{R}\}$, so v_1, \dots, v_m are not collinear.

Now we need to show that there also exists a δ ball about $p = (p_1, \dots, p_m)$ so that v_1, \dots, v_m are not collinear when $v = (v_1, \dots, v_m) \in B_\delta(p)$. We may simply choose $\delta = \delta'$: For $v \in B_{\delta'}(p)$ and any pair of indices $i < j \in \{1, \dots, m\}$ it holds that

$$\delta'^2 > \|v - p\|_2^2 = \left\| \begin{pmatrix} v_i \\ v_j \end{pmatrix} - \begin{pmatrix} p_i \\ p_j \end{pmatrix} \right\|_2^2 + \sum_{l \in I_{i,j}} (v_l - p_l)^2 \geq \left\| \begin{pmatrix} v_i \\ v_j \end{pmatrix} - \begin{pmatrix} p_i \\ p_j \end{pmatrix} \right\|_2^2$$

where $I_{i,j} = \{1, \dots, m\} \setminus \{i, j\}$. Thus,

$$\left\| \begin{pmatrix} v_i \\ v_j \end{pmatrix} - \begin{pmatrix} p_i \\ p_j \end{pmatrix} \right\|_2 < \delta'$$

and hence Equation (5.12) holds. \square

We use Lemma 5.40 to get uniqueness of minimizers: The solution to the median problem with facilities v_1, \dots, v_M is unique if and only if v_1, \dots, v_M are not collinear, due to the strict convexity of the objective in this case (as stated in [VZ00]).

Thus, we get uniqueness of minimizers of the Euclidean median problems with respect to the points (v_1, \dots, v_M) on some δ -ball, say $B_\delta(u)$. Then, on this ball we may define the map

$$\xi : B_\delta(u) \rightarrow \mathbb{R}_+^{n-2}, \quad \xi((v_1, \dots, v_M)) := \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - v_i\|_2 \quad (5.13)$$

and we want to show its continuity.

Proposition 5.41. *Assume that for all $u_0 \in \mathbb{R}_+^{n-2}$ it holds that $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear. Then for any $u_0 \in \mathbb{R}_+^{n-2}$ there exists a $\delta > 0$ such that for all $v = (v_1, \dots, v_M) \in B_\delta(u)$ with $u = (P_{T_1}(u_0), \dots, P_{T_M}(u_0))$ it holds that*

$$\min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - v_i\|_2$$

has a unique minimizer. Then

$$\xi : B_\delta(u) \rightarrow \mathbb{R}_+^{n-2}, \quad \xi((v_1, \dots, v_M)) := \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - v_i\|_2$$

is continuous.

Proof. Take $u_0 \in \mathbb{R}_+^{n-2}$ arbitrary. The prerequisite implies that $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear and Lemma 5.40 states that there exists a $\delta > 0$ such that for all

$v = (v_1, \dots, v_M)^t \in B_\delta(u)$ it holds that v_1, \dots, v_M are not collinear. [VZ00] states that non-collinearity implies that there exists a unique minimizer of

$$\min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - v_i\|_2$$

due to the strict convexity of the objective in this case. So for all $v \in B_\delta(u)$ we have a unique minimizer, so we may define the map

$$\xi : B_\delta(u) \cap \mathbb{R}_+^{n-2} \rightarrow \mathbb{R}_+^{n-2}, \quad \xi((v_1, \dots, v_M)) := \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - v_i\|_2.$$

Our goal is to show that ξ is continuous. To this end we want to apply Theorem 3.2 of [War83], that states the following:

Given a continuous map $g : V \times W$, where $V \subset \mathbb{R}^{d_1}$ is a compact and $W \subset \mathbb{R}^{d_2}$ arbitrary, define $X(w)$ as the set of minimizers of $g(\cdot, w)$ over V for each $w \in W$. Then the set-valued map $\bar{\xi} : W \rightrightarrows V, \bar{\xi}'(w) := X(w)$ is upper semicontinuous.

The notion of upper semicontinuity is defined as follows: A set-valued map $F : X \rightrightarrows Y$ is called upper-semicontinuous at x if and only if for all neighborhoods \mathcal{U} of $F(x)$, there exists a $\delta > 0$ such that for all $x' \in B_\delta(x)$ it holds that $F(x') \subset \mathcal{U}$. Here, $B_\delta(x)$ is the ball centered at x with respect to the distance defined on X .

However, if the images $F(x)$ are singletons for all x , this notion of upper semicontinuity translates into the usual continuity of the point-valued map $F(x)$.

Now we want use this theorem to show continuity for ξ . To this end we first have to define the corresponding map g and the sets V, W . In order to define V , we recall Theorem 5.21, that states that may restrict the orthant \mathbb{R}_+^{n-2} to the set $[0, U]^{n-2}$ when searching for an optimal solution. With this we set $V := [0, U]^{n-2}$, which is compact. For the set of parameters W , on which the minimizers $X(w)$ depend, we use $W := B_\delta(u)$ and define

$$g : V \times W \rightarrow \mathbb{R}, \quad g(v, w) = \sum_{i=1}^M \|v - w_i\|_2 \quad \text{where } w = (w_1, \dots, w_M).$$

g is continuous because $d_2 : \mathbb{R}^{2 \cdot (n-2)} \rightarrow \mathbb{R}, d_2(x) := \|x_1 - x_2\|_2$ with $x = (x_1, x_2)^t \in \mathbb{R}^{2 \cdot (n-2)}$ is continuous. We have shown above that for a fixed $w \in W = B_\delta(u)$, $X(w) = \arg \min_{v \in [0, U]^{n-2}} \sum_{i=1}^M \|v - w_i\|_2$ has a unique minimizer and that $X(w) = \xi(w)$ holds by definition of ξ . Then we have verified that all requirements for Theorem 3.2 of [War83] are satisfied and it follows that ξ is continuous. \square

In the following example we clarify why the condition on the points $P_{T_i}, i = 1, \dots, M$ not to lie on a line segment is necessary.

Example 5.3.2.

$$\text{Let } u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \text{ with } u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, u_4 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

and consider the median problem $P(u) = \min_{x \in \mathbb{R}^2} \sum_{i=1}^4 \|x - u_i\|_2$. Now for any $\delta \in (0, 1)$ define

$$\tilde{u}(\delta) = \begin{pmatrix} \tilde{u}_1(\delta) \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \text{ with } \tilde{u}_1(\delta) = \begin{pmatrix} 1 \\ 1 + \frac{\delta}{2} \end{pmatrix},$$

$$\hat{u}(\delta) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \hat{u}_4(\delta) \end{pmatrix} \text{ with } \hat{u}_4(\delta) = \begin{pmatrix} 4 \\ 1 + \frac{\delta}{2} \end{pmatrix}.$$

Figure 5.12 depicts the minimizers of $P(u)$ in red, the minimizer of $P(\hat{u})$ in blue and the minimizer of $P(\tilde{u})$ in orange:

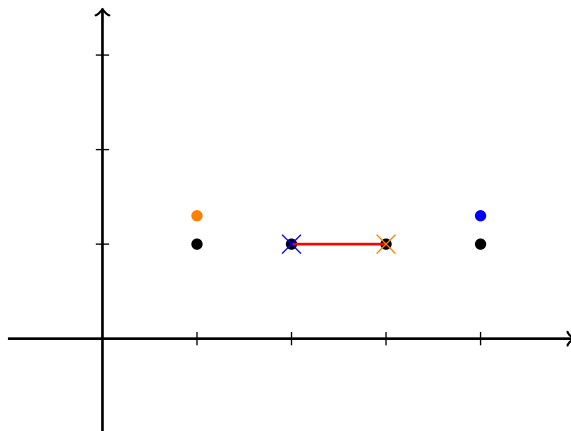


Figure 5.12: Example for the non-continuous behavior of minimizers in the case of non-unique minimizers.

The Euclidean median problem with four given facilities, is well studied (see [Pla05] for a survey). In our cases at hand we have that one point is contained in the triangle spanned by three other points, which is the so called 'absorbed case'. In this case, the point inside the triangle is the optimal solution. Thus, the minimizer for $P(\tilde{u}(\delta))$ is $\tilde{x} = u_3$ and the unique minimizer for $P(\hat{u}(\delta))$ is $\hat{x} = u_2$. The minimizers

are independent of δ , so the distance between the minimizers is $\|\hat{x} - \tilde{x}\|_2 = 1$ for all $\delta \in (0, 1)$.

But by definition we have $\tilde{u}, \hat{u} \in B_\delta(u)$, so we see that there is no continuous behavior of the minimizers with respect to parameter u in this case in the 'usual', not set-valued, sense: The distance to the minimizers \tilde{x} and \hat{x} gets arbitrarily small when δ tends to zero, but the distance between the image of the minimizers is 1 as explained above.

We want to stress that we actually do have upper semi-continuity in the sense of [War83]. But what we require in the proof for Theorem 5.41 is that we have a unique minimizer for the problem, since we cannot use a set of minimizers as ξ is not a set-valued map. The example above depicts a case where the minimizer is not unique and where there is no choice of a minimizer for $P(u)$ on the line such that ξ is continuous in the end. Hence, we have to rule out these cases via the prerequisite of the theorem.

An easily verifiable, sufficient condition for the prerequisites of Proposition 5.41 is non-collinearity of the facilities that are in \mathcal{O} .

Lemma 5.42. *Suppose that $|\mathbf{T} \cap \mathcal{O}| \geq 3$ and that $\psi(\mathbf{T} \cap \mathcal{O})$ are not collinear. Then $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear for any $u_0 \in \mathbb{R}_+^{n-2}$.*

Proof. For $T_i \in \mathbf{T} \cap \mathcal{O}$ it holds that $P_{T_i}(u_0) = \psi(T_i)$ for all $u_0 \in \mathbb{R}_+^{n-2}$. Hence, $\psi(\mathbf{T} \cap \mathcal{O}) \subset \{P_{T_1}(u_0), \dots, P_{T_M}(u_0)\}$ for all $u_0 \in \mathbb{R}_+^{n-2}$ and as $\psi(\mathbf{T} \cap \mathcal{O})$ is not collinear, it follows that $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear. \square

Before we finally prove the convergence of a subsequence of the BPA to a stationary point, we provide a useful lemma required in the proof.

Lemma 5.43. *Let $(u^{(\ell)})_{\ell \in \mathbb{N}}$ be the sequence of iterates of the BPA and let $(u^{(\ell_m)})_{m \in \mathbb{N}}$ be a subsequence of $(u^{(\ell)})_{\ell \in \mathbb{N}}$ converging to $\bar{u} = (\bar{u}_0, \bar{u}_1, \dots, \bar{u}_M)$. Then*

$$\bar{u}_i = P_{T_i}(\bar{u}_0) \quad \text{for } i = 1, \dots, M.$$

Proof. When $u^{(\ell_m)} \rightarrow \bar{u}$ it follows that we in particular have convergence for the blocks

$$u_i^{(\ell_m)} \rightarrow \bar{u}_i \quad \text{for } i = 0, 1, \dots, M. \quad (5.14)$$

Proposition 5.39 states that P_{T_i} is continuous for $i = 1, \dots, M$. Now, use (1) for $i = 0$, i.e.,

$$u_0^{(\ell_m)} \rightarrow \bar{u}_0$$

Then continuity of the P_{T_i} implies that

$$P_{T_i}(u_0^{(\ell_m)}) \rightarrow P_{T_i}(\bar{u}_0). \quad (5.15)$$

Due to the definition of the sequence $(u^{(\ell)})_{\ell \in \mathbb{N}}$ it holds that

$$u_i^{(\ell_m)} = P_{T_i}(u_0^{(\ell_m)}).$$

Using (5.14) for $i = 1, \dots, M$ we get $u_i^{(\ell_m)} \rightarrow \bar{u}_i$ and plugging in $u_i^{(\ell_m)} = P_{T_i}(u_0^{(\ell_m)})$ on the left hand side yields

$$P_{T_i}(u_0^{(\ell_m)}) \rightarrow \bar{u}_i. \quad (5.16)$$

Combining (5.15) and (5.16) it follows that $\bar{u}_i = P_{T_i}(\bar{u}_0)$ for $i = 1, \dots, M$. \square

Now we are finally able to prove convergence of a subsequence of the iterates of the BPA.

Theorem 5.44. *Assume that for all $u_0 \in \mathbb{R}_+^{n-2}$ it holds that $P_{T_1}(u_0), \dots, P_{T_M}(u_M)$ are not contained in a line segment. Moreover, let $W := [0, U]^{n'}$, where*

$$U = \max_{i=1, \dots, M} \left\{ \max_{s \in \text{Split}(T_i)} w_s^{T_i} \right\}.$$

Then, for any starting point $u^{(0)} \in W$, let $u^{(\ell)}$ be defined as in Algorithm 8. Then $(u^{(\ell)})_{\ell \in \mathbb{N}}$ has a cluster point and any cluster point of $u^{(\ell)}$ is a stationary point.

Proof. Before we start with the actual proof we rule out the case $M = 2$, due to the assumption: When only two trees T_1, T_2 are given, then $\{P_{T_1}(u), P_{T_2}(u)\}$ is contained in a line segment for any u , so we may assume $M \geq 3$ in the following, which is of importance when applying Lemma 5.40.

In order to find a convergent subsequence of $(u^{(\ell)})_{\ell \in \mathbb{N}}$ we show that $(u^{(\ell)})_{\ell \in \mathbb{N}} \subset W$ holds. By definition of U , the weights of all P_{T_i} are less or equal U . Since $u^{(0)} \in W$ it holds in particular that $u_0 \in [0, U]^{n-2}$. Now, since the weights of all splits $s \in \text{Split}(Y)$ of a tree Y on a geodesic $\Gamma(X, T)$ are always linear combinations of the split lengths of the input trees T and X , (compare the parametrization in Theorem 2.21), it follows that $w_s^Y \in [0, \max\{\max_{s \in \text{Split}(X)} w_s^X, \max_{s \in \text{Split}(T)} w_s^T\}]$. In particular, this holds for the exit points $\psi^{-1}(P_{T_i}(y)) \in \Gamma(\psi^{-1}(y), T_i)$: if $y \in [0, U]^{n-2}$, it holds that $P_{T_i}(y) \in [0, U]^{n-2}$ as all split lengths of T_i are less or equal U . Moreover, when $P_{T_i}(y) \in [0, U]^{n-2}$ for $i = 1, \dots, M$ then $\text{conv}(P_{T_1}(y), \dots, P_{T_M}(y)) \subset [0, U]^{n-2}$ and thus

$$\arg \min_{x \in \mathbb{R}_+^{n-2}} \left(\sum_{i=1}^M \|x - P_{T_i}(y)\|_2 \right) \in \text{conv}(P_{T_1}(y), \dots, P_{T_M}(y)) \subset [0, U]^{n-2}.$$

Applying these results iteratively we get the following: for a starting point $u^{(0)} \in W$ it holds that $u_0^{(0)} \in [0, U]^{n-2}$; this implies $u_i^{(0)} = P_{T_i}(u_0^{(0)}) \in [0, U]^{n-2}$ for $i = 1, \dots, M$ which in turn implies that $u_0^{(1)} \in [0, U]^{n-2}$ and so on and so forth. Hence, the iterates

satisfy $u^{(\ell)} \in W$ for $\ell = 0, 1, 2, \dots$. So $(u^{(\ell)})_{\ell \in \mathbb{N}}$ is a sequence in the compact set W and thus it has at least one cluster point.

Now, let \bar{u} be a cluster point of $u^{(\ell)}$ and let $(u^{(\ell_m)})_{m \in \mathbb{N}}$ be the subsequence that converges to \bar{u} . Our goal is to show that \bar{u} is a stationary point and we assume to the contrary that it was no stationary point. Then we have

$$u^{(\ell_m)} \neq u^{(\ell_{m+1})}, m \in \mathbb{N} \quad (5.17)$$

because otherwise one $u^{(\ell_{m'})}$ would be a stationary point and

$$u^{(\ell_{m'})} = u^{(\ell_{m'+1})} = \dots = u^{(\ell_{(m'+1)})} = \dots = \dots = \bar{u}.$$

The tricky part now is to construct a neighborhood of \bar{u} such that ξ , see (5.13), is well-defined and continuous.

To this end we first apply Lemma 5.43, that shows that $\bar{u}_i = P_{T_i}(\bar{u}_0)$ for $i = 1, \dots, M$. Then the prerequisite ensures that $P_{T_1}(\bar{u}_0), \dots, P_{T_M}(\bar{u}_0)$ are not collinear, and thus $\bar{u}_1, \dots, \bar{u}_M$ are not collinear. Now, Lemma 5.40 provides a $\delta' > 0$ such that for any $v \in B_{\delta'}(\bar{u})$ we have that v_1, \dots, v_M are not collinear. Choosing $0 < \delta < \delta'$ we have that this holds for the closed ball as well which we denote by $D = \{v \in \mathbb{R}^n : \|v - u\|_2 \leq \delta\}$. Now we define another compact set $C' = \overline{\{u^{(\ell_m)} : m \in \mathbb{N}\}}$, which is the closure of our subsequence iterates. Since \bar{u} is the cluster points of our subsequence, the closure implies that $\bar{u} \in C'$ as well. The convergence of $u^{(\ell_m)}$ to \bar{u} implies that there exists some $m' \in \mathbb{N}$ such that $\|u - u^{(\ell_m)}\|_2 < \delta$ for all $m \geq m'$, i.e., $u^{(\ell_m)} \in D$ for $m \geq m'$. Hence we have that $\overline{\{u^{(\ell_m)} : m \geq m'\}} \subset D$ and define $C = \overline{\{u^{(\ell_m)} : m \geq m'\}}$, which is compact.

Now, we introduce an auxiliary map that describes the sequence of iterates of the BPA:

$$g : B_{\delta}(u) \rightarrow W, \quad g(w) = \begin{pmatrix} \xi(w_1, \dots, w_M) \\ P_{T_1}(\xi(w_1, \dots, w_M)) \\ \vdots \\ P_{T_M}(\xi(w_1, \dots, w_M)) \end{pmatrix}.$$

Note, that g is only well defined, when ξ is well-defined, i.e., when there is a unique choice for the minimizer of the Euclidean median problem. This is ensured as we have just shown that we have uniqueness of minimizers on $D = B_{\delta}(u)$. This implies that g is a well-defined map on $C \subset D$. Moreover, g is continuous on C due to Proposition 5.39 and Proposition 5.41, as it is a composition of continuous functions.

Using g we now construct a continuous map F for which $F(u^{(\ell_m)}) \rightarrow -\infty$, i.e., we construct a continuous map that is unbounded on a compact set if the assumption that \bar{u} is no stationary point holds.

Continuity of g implies that $F : C \rightarrow \mathbb{R}$, $F(w) = \varphi(g(w)) - \varphi(w)$ is continuous, because φ is continuous as well. Furthermore $F(w) < 0$ holds if and only if $\varphi(g(w)) < \varphi(w)$: On C , minimizers to the Euclidean median problems are unique, so we know that for any $w \in C$

$$\sum_{i=1}^M \|g(w)_0 - w_i\|_2 < \sum_{i=1}^M \|w_0 - w_i\|_2. \quad (5.18)$$

With this it follows that for any $w \in C$

$$\begin{aligned} \varphi(g(w)) &= \sum_{i=1}^M d(\psi^{-1}(g(w)_i), T_i) + \sum_{i=1}^M \|g(w)_0 - g(w)_i\|_2 = \sum_{i=1}^M d(\psi^{-1}(g(w)_0), T_i) \\ &\leq \sum_{i=1}^M d(\psi^{-1}(w_i), T_i) + \sum_{i=1}^M \|g(w)_0 - w_i\|_2 \\ &< \sum_{i=1}^M d(\psi^{-1}(w_i), T_i) + \sum_{i=1}^M \|w_0 - w_i\|_2 = \varphi(w), \end{aligned}$$

so $F(w) < 0$.

Since C is compact and F is continuous, F attains its minimum $c = \min_{v \in C} F(v)$ on C , which has to be smaller than 0, because $F(w) < 0$ for all $w \in C$. Thus $F(w) \leq c < 0$ and therewith

$$F(u^{(\ell_m)}) = \varphi(g(u^{(\ell_m)})) - \varphi(u^{(\ell_m)}) = \varphi(u^{(\ell_{m+1})}) - \varphi(u^{(\ell_m)}) \leq c \text{ for } m = 1, 2, \dots$$

Moreover, $(\varphi(u^{(\ell)}))_{\ell \in \mathbb{N}}$ is a non-increasing sequence and $\ell_{m+1} \geq \ell_m + 1$, so

$$\varphi(u^{(\ell_{m+1})}) - \varphi(u^{(\ell_m)}) \leq \varphi(u^{(\ell_{m+1})}) - \varphi(u^{(\ell_m)}) \leq c \text{ for } m = 1, 2, \dots$$

Applying the inequality for ℓ_m, ℓ_{m-1} successively until we arrive at ℓ_1 we receive

$$\varphi(u^{(\ell_m)}) \leq \varphi(u^{(\ell_1)}) + (m-1) \cdot c.$$

So, when $m \rightarrow \infty$, $\varphi(u^{(\ell_m)}) \rightarrow -\infty$, which is a contradiction as φ is a continuous function and is bounded over the compact set W .

Thus, the assumption that \bar{u} is no stationary point yields a contradiction. \square

Proposition 5.45. *Let \bar{u} be a cluster point of the iterates $u^{(\ell)}$ of Algorithm 8. Then $\lim_{\ell \rightarrow \infty} \varphi(u^{(\ell)})$ exists and $\varphi(u^{(\ell)}) \rightarrow \varphi(\bar{u})$.*

Proof. Let $(u^{(\ell_m)})_{m \in \mathbb{N}}$ be the subsequence converging to \bar{u} . Since φ is continuous, this implies that $(\varphi(u^{(\ell_m)}))_{m \in \mathbb{N}}$ converges to $\varphi(\bar{u})$.

First assume, that there exists some $u^{(\ell')}$ such that $\varphi(u^{(\ell')}) < \varphi(\bar{u})$. But then there exists some $m' \in \mathbb{N}$ such that $\ell_m \geq \ell'$ for all $m \geq m'$ and since $(\varphi(u^{(\ell)}))$ is non-increasing that yields

$$\varphi(u^{(\ell_m)}) \leq \varphi(u^{(\ell')}) < \varphi(\bar{u}), \quad \forall m \geq m'$$

which is a contradiction. Hence, $\varphi(u^\ell) \geq \varphi(\bar{u})$.

Now let $\varepsilon > 0$ be given. Then there exists $m' \in \mathbb{N}$ such that

$$\|\varphi(\bar{u}) - \varphi(u^{(\ell_m)})\|_2 < \varepsilon \quad \forall m \geq m'.$$

Lemma 5.30 states that φ is non-increasing and since $\varphi(u^\ell) \geq \varphi(\bar{u})$ we have that $\|\varphi(\bar{u}) - \varphi(u^{(\ell_{m'})})\|_2 \geq \|\varphi(\bar{u}) - \varphi(u^{(\ell)})\|_2$ for $\ell \geq \ell_{m'}$, so

$$\|\varphi(\bar{u}) - \varphi(u^{(\ell)})\|_2 \leq \|\varphi(\bar{u}) - \varphi(u^{(\ell_{m'})})\|_2 < \varepsilon \quad \forall \ell \geq \ell_{m'}.$$

□

Note that Proposition 5.45 holds even without the further assumptions of Theorem 5.44, it only requires the existence of a cluster point and the sequence not to increase.

The reason why we proved it after stating Theorem 5.44 is that the theorem does not prove actual convergence of the iterates $u^{(\ell)}$ to some u^* for the whole sequence. However, we can ensure that the sequence of objective values converges and in the next section we will prove that stationary points yield optimal solutions under some assumptions - so when we prove optimality of a cluster point of $u^{(\ell)}$ we know that the whole sequence converges to the optimal objective.

Obviously it would be more convenient, when

5.3.3 Optimality of Stationary Points and Balance Points

Theorem 5.44 in Subsection 5.3.2 guarantees us that we end up in a stationary point, when the stated assumption holds. This was the first goal that we proclaimed at the start of the section. But eventually we want to find a minimizer of φ - so we want to prove that, under slight assumptions, the stationary points the BPA yields are optimal. Luckily, φ is a convex function so it suffices to show that we have a local optimum. To this end, we want to calculate the directional derivatives of φ . We have seen in Proposition 5.26 that φ is not differentiable on our whole set of definition. So once more, we have to check one of the assumptions of [War63] by hand, namely if φ is differentiable at the stationary point for which we want to show optimality.

Unfortunately, when \bar{u} is a stationary point that is an accumulation point of the sequence of iterates $u^{(\ell)}$ of the BPA, then φ is not differentiable at \bar{u} : Lemma 5.43

implies that $\bar{u}_i = P_{T_i}(\bar{u}_0) \in \partial(S)$. But the geodesic distance d is not differentiable at the boundary of an orthant, so φ is not differentiable at \bar{u} .

This observation once more complicates the matter, as we cannot prove optimality for the stationary points the BPA yields by showing that all directional derivatives at a stationary point are ≥ 0 . But, a crafty trick allows us to still prove optimality in a similar fashion. To this end, choose a fixed $\varepsilon \in (0, 1)$ and define

$$R_{T_i}(u_0) = \varepsilon u_0 + (1 - \varepsilon)P_{T_i}(u_0) \quad \text{for } i = 1, \dots, M. \quad (5.19)$$

Then we have that for any $u_0 \in \text{int}(S)$ it holds that

$$\varphi(u_0, P_{T_1}(u_0), \dots, P_{T_M}(u_0)) = \varphi(u_0, R_{T_1}(u_0), \dots, R_{T_M}(u_0)),$$

as $R_{T_i}(u_0) \in \Gamma(\psi^{-1}(u_0), T_i)$.

So when we are able to prove optimality for $(u_0, R_{T_1}(u_0), \dots, R_{T_M}(u_0))$, then it follows that $(u_0, P_{T_1}(u_0), \dots, P_{T_M}(u_0))$ is optimal as well. The next result shows that φ is differentiable at $(u_0, R_{T_1}(u_0), \dots, R_{T_M}(u_0))$ under a mild condition.

Lemma 5.46. *Let $u_0 \in \text{int } \mathbb{R}_+^{n-2}$ and a fixed $\varepsilon \in (0, 1)$ be given. Then φ is differentiable at $u = (u_0, R_{T_1}(u_0), \dots, R_{T_M}(u_0))$, if $\psi(u_0) \notin \{T_1, \dots, T_m\}$.*

Proof. Assume $u_0 \in W$. That implies that there exists some $k \in \{1, \dots, M\}$ such that $u_0 = u_k$. Additionally we have $u_k = \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$, hence $u_0 = \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$. There are two cases now that both lead to the same contradiction. Firstly, if $T_k \in \text{int } \mathcal{O}$, then $P_{T_k}(u_0) = \psi(T_k) \neq u_0$ due to the assumption $\psi^{-1}(u_0) \notin \{T_1, \dots, T_m\}$. Thus $u_0 \neq \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$. Secondly, if $T_k \notin \text{int } \mathcal{O}$, then $P_{T_k}(u_0) \in \partial \mathcal{O}$. Since $u_0 \in \text{int } \mathcal{O}$ this implies again $u_0 \neq P_{T_k}(u_0)$ and we receive the same contradiction.

Assume $u_0 \in V$. That implies that there exists some $k \in \{1, \dots, M\}$ such that $\psi^{-1}(u_k) = T_k$. Since $\psi : \mathcal{O} \rightarrow \mathbb{R}_+^{n-2}$, this already implies that $T_k \in \mathcal{O}$. We have $u_k = \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$ and together with $\psi^{-1}(u_k) = T_k$ this yields $\psi(T_k) = \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$. Since u_0 is in $\text{int } \mathbb{R}_+^{n-2}$, this implies that $T_k \in \text{int } \mathcal{O}$, so $P_{T_k}(u_0) = \psi(T_k)$. The prerequisite states $u_0 \notin \{\psi^{-1}(T_1), \dots, \psi^{-1}(T_M)\}$ so we have $u_0 \neq \psi(T_k) = P_{T_k}(u_0)$. Thus $\psi(T_k) \neq \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0)$.

These two results show that $u \notin V \cup W$. Moreover, when $u_0 \in \text{int } S$, then $u_k = \varepsilon u_0 + (1 - \varepsilon)P_{T_k}(u_0) \in \text{int } \mathbb{R}_+^{n-2}$ for all $k \in \{1, \dots, M\}$ and that implies $u \in \text{int } (\mathbb{R}_+^{n'})$. Finally, Proposition 5.26 implies that φ is differentiable at u . \square

As we mentioned [War63] proves the optimality of stationary points using differentiability. Lemma 5.46 has given us a point $v \in \mathbb{R}_+^{n'}$ with the same objective value as the BPA stationary point, φ is differentiable at v . But in order to use the optimality criterion we still need to show that v is a stationary point.

Proposition 5.47. *Let \bar{u} be a stationary point determined by the BPA. Then*

$$v = (\bar{u}_0, R_{T_1}(\bar{u}_0), \dots, R_{T_M}(\bar{u}_0))$$

is a stationary point of φ as well.

Proof. First, recall that Lemma 5.43 implies $\bar{u} = (\bar{u}_0, P_{T_1}(\bar{u}_0), \dots, P_{T_M}(\bar{u}_0))$. By the definition of stationary points, Definition 5.19, a point $u \in \mathbb{R}_+^{n'}$ is a stationary point if

$$u_i \in \arg \min_{x \in \mathbb{R}_+^{n-2}} \varphi \left((u_0^{(\ell+1)}, u_1^{(\ell+1)}, \dots, u_{i-1}^{(\ell+1)}, x, u_{i+i}^{(\ell)}, \dots, u_M^{(\ell)})^t \right) \quad \text{for } i = 0, 1, \dots, M. \quad (5.20)$$

In the following we deduce that v is a stationary point by using that u is a stationary point. We first verify (5.20) for the easy cases of $i = 1, \dots, M$. We know that $P_{T_i}(\bar{u}_0) \in \Gamma(\psi^{-1}(\bar{u}_0), T_i)$ and as $R_{T_i}(\bar{u}_0) = \varepsilon \bar{u}_0 + (1 - \varepsilon)P_{T_i}(\bar{u}_0)$ we have $R_{T_i}(\bar{u}_0) \in \Gamma(\psi^{-1}(\bar{u}_0), T_i)$ as well. Thus

$$v_i = R_{T_i}(\bar{u}_0) \in \arg \min_{x \in \mathbb{R}_+^{n-2}} \varphi \left((u_0^{(\ell+1)}, u_1^{(\ell+1)}, \dots, u_{i-1}^{(\ell+1)}, x, u_{i+i}^{(\ell)}, \dots, u_M^{(\ell)})^t \right),$$

as

$$\begin{aligned} & \arg \min_{x \in \mathbb{R}_+^{n-2}} \varphi \left((u_0^{(\ell+1)}, u_1^{(\ell+1)}, \dots, u_{i-1}^{(\ell+1)}, x, u_{i+i}^{(\ell)}, \dots, u_M^{(\ell)})^t \right) \\ &= \{x \in \mathbb{R}_+^{n-2} : \psi^{-1}(x) \in \Gamma(\psi^{-1}(\bar{u}_0), T_i)\}. \end{aligned}$$

Now we need to check the condition for $i = 0$. Since \bar{u} is a stationary point it holds that

$$\bar{u}_0 \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - P_{T_i}(\bar{u}_0)\|_2. \quad (5.21)$$

Condition (5.20) for $i = 0$ is

$$v_0 \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - R_{T_i}(\bar{u}_0)\|_2. \quad (5.22)$$

Because $v_0 = \bar{u}_0$, we need to show that

$$\bar{u}_0 \in \arg \min_{y \in \mathbb{R}_+^{n-2}} \sum_{i=1}^M \|y - R_{T_i}(\bar{u}_0)\|_2.$$

We prove that \bar{u}_0 is a minimizer for two different cases. First, assume that $\bar{u}_0 = \bar{u}_k$ for some $k \in \{1, \dots, M\}$. Since $\bar{u}_k = P_{T_k}(\bar{u}_0)$, this implies that the facility $P_{T_k}(\bar{u}_0)$ is optimal. Theorem 3.1 (ii) implies that then

$$\left\| \sum_{\substack{i=1 \\ i \neq k}}^M \frac{\bar{u}_0 - P_{T_i}(\bar{u}_0)}{\|\bar{u}_0 - P_{T_i}(\bar{u}_0)\|_2} \right\|_2 \geq 1.$$

By definition, $R_{T_i}(\bar{u}_0) = \varepsilon u_0 + (1 - \varepsilon)P_{T_i}(\bar{u}_0)$, so

$$\frac{\bar{u}_0 - P_{T_i}(\bar{u}_0)}{\|\bar{u}_0 - P_{T_i}(\bar{u}_0)\|_2} = \frac{\bar{u}_0 - R_{T_i}(\bar{u}_0)}{\|\bar{u}_0 - R_{T_i}(\bar{u}_0)\|_2}.$$

Applying Theorem 3.1 (ii) to the right hand side now implies that $\bar{u}_0 = v_0$ is optimal to (5.22), what was to show.

The other case is that $\bar{u}_0 \notin \{P_{T_1}(\bar{u}_0), \dots, P_{T_m}(\bar{u}_0)\}$. As the $P_{T_i}(\bar{u}_0)$ are the facilities of the Euclidean median problem in (5.21), it follows that $g(y) := \sum_{i=1}^M \|y - P_{T_i}(\bar{u}_0)\|_2$ is differentiable at \bar{u}_0 and the gradient is

$$\nabla g(y) = \sum_{i=1}^M \frac{y - P_{T_i}(\bar{u}_0)}{\|y - P_{T_i}(\bar{u}_0)\|_2}.$$

Due to the optimality of \bar{u}_0 , $\nabla g(\bar{u}_0) = 0$. Then

$$0 = \nabla g(\bar{u}_0) = \sum_{i=1}^M \frac{\bar{u}_0 - P_{T_i}(\bar{u}_0)}{\|\bar{u}_0 - P_{T_i}(\bar{u}_0)\|_2} = \sum_{i=1}^M \frac{\bar{u}_0 - R_{T_i}(\bar{u}_0)}{\|\bar{u}_0 - R_{T_i}(\bar{u}_0)\|_2}$$

implies that $\bar{u}_0 = v_0$ is optimal to the optimization problem in (5.22), which finishes the proof. \square

Finally we have constructed a specific stationary point at which φ is differentiable. So we have checked that all prerequisites of the proof of optimality from [War63] hold and now only present a last auxiliary lemma before proving optimality.

To this end, we introduce additional 'sub'functions of φ . Recall that $n' = (M + 1) \cdot (n - 2)$ and for $i = 0, \dots, M$, $u \in \mathbb{R}_+^{n'}$ denote $(u_i, u_{-i}) \in \mathbb{R}_+^{n-2} \times \mathbb{R}^{M \cdot (n-2)}$, where $u_{-i} = (u_0, \dots, u_{i-1}, u_{i+1}, \dots, u_M)$. With this notation define for a given u_{-i}

$$\varphi_i(\cdot, u_{-i}) : \mathbb{R}_+^{n-2} \rightarrow \mathbb{R}, \quad \varphi_i(x, u_{-i}) = \varphi((u_0, \dots, u_{i-1}, x, u_{i+1}, \dots, u_M)).$$

The following lemma states a helpful relationship of the directional derivatives of φ and the directional derivatives of the φ_i that we will use to prove local optimality of stationary points.

In order to calculate the directional derivatives we calculate partial derivatives, so we need to differentiate with respect to a single coordinate of $u = (u_0, u_1, \dots, u_M)$. This is the only time we need single coordinates, so in order to avoid confusion we introduce a specific notation. We denote single coordinates by u_{ij} , so that for each $i = 1, \dots, M$

$$u_i = \begin{pmatrix} u_{i_1} \\ \vdots \\ u_{i_{n-2}} \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_M \end{pmatrix} = \begin{pmatrix} u_{0_1} \\ \vdots \\ u_{0_{n-2}} \\ u_{1_1} \\ \vdots \\ u_{1_{n-2}} \\ u_{2_1} \\ \vdots \\ u_{M_{n-2}} \end{pmatrix}.$$

Lemma 5.48. *Let φ be continuously differentiable at u . Denote the directional derivative in direction y at u by $D_y(\varphi)(u)$ and use the according notation for the φ_i as well. Then*

$$D_y(\varphi)(u) = \sum_{i=0}^M \langle \nabla(\varphi_i(\cdot, u_{-i}))(u_i), y_i \rangle.$$

Proof. We have continuous differentiability of φ at u so we can rewrite the directional derivative as the sum of the partial derivatives of all coordinates

$$D_y(\varphi)(u) = \langle \nabla\varphi(u), y \rangle = \sum_{i=0}^M \sum_{j=1}^{n-2} \frac{\partial\varphi}{\partial u_{ij}}(u) \cdot y_{ij}, \quad (5.23)$$

Now, consider the partial derivatives with respect to the blocks $i \in \{1, \dots, M\}$ and

some coordinate $j \in \{1, \dots, n-2\}$ of the block u_i :

$$\begin{aligned}
\frac{\partial \varphi(u)}{\partial u_{ij}} &= \frac{\partial}{\partial u_{ij}} \left(\sum_{k=1}^M d(\psi(u_k), T_k) + \sum_{k=1}^M \|u_0 - u_k\|_2 \right) (u) \\
&= \frac{\partial d(\psi(u_i), T_i)}{\partial u_{ij}}(u) + \frac{\partial \|u_0 - u_i\|_2}{\partial u_{ij}}(u) \\
&= \frac{\partial d(\psi(u_i), T_i)}{\partial u_{ij}}(u_i) + \frac{\partial \|u_0 - u_i\|_2}{\partial u_{ij}}(u_i) \\
&= \frac{\partial}{\partial u_{ij}} \left(\sum_{k=1}^M d(\psi(u_k), T_k) + \sum_{k=1}^M \|u_0 - u_k\|_2 \right) (u_i) \\
&= \frac{\partial \varphi_i(\cdot, u_{-i})}{\partial u_{ij}}(u_i)
\end{aligned}$$

We want to verify this for $i = 0$ as well:

$$\begin{aligned}
\frac{\partial \varphi(u)}{\partial u_{0j}} &= \frac{\partial}{\partial u_{0j}} \left(\sum_{k=1}^M d(\psi(u_k), T_k) + \sum_{k=1}^M \|u_0 - u_k\|_2 \right) (u) \\
&= \sum_{i=1}^M \frac{\partial (\|u_0 - u_i\|_2)}{\partial u_{0j}}(u) = \sum_{i=1}^M \frac{\partial (\|u_0 - u_i\|_2)}{\partial u_{0j}}(u_0) \\
&= \frac{\partial}{\partial u_{0j}} \left(\sum_{k=1}^M d(\psi(u_k), T_k) + \sum_{k=1}^M \|u_0 - u_k\|_2 \right) (u_0) \\
&= \frac{\partial \varphi_0(\cdot, u_{-0})}{\partial u_{0j}}(u_0)
\end{aligned}$$

So it turns out that the partial derivatives for a coordinate u_{i_k} are independent of all other blocks u_j , $j \neq i$ and thus, the partial derivatives of φ with respect to some coordinate of block i are the same as the partial derivative for the respective coordinate of the $\varphi_i(\cdot, u_{-i})$. In particular, the continuity of all partial derivatives of the $\varphi_i(\cdot, u_{-i})$ implies that

$$\sum_{i=0}^M \sum_{j=1}^{n-2} \frac{\partial \varphi_i(\cdot, u_{-i})}{\partial u_{ij}}(u_i) \cdot y_{ij} = \sum_{i=0}^M \langle \nabla(\varphi_i(\cdot, u_{-i}))(u_i), y_i \rangle \quad (5.24)$$

Now, combining (5.23) and (5.24) we get

$$\begin{aligned} D_y(\varphi)(u) &= \sum_{i=0}^M \sum_{j=1}^{n-2} \frac{\partial \varphi}{\partial u_{ij}}(u) \cdot y_{ij} \\ &= \sum_{i=0}^M \sum_{j=1}^{n-2} \frac{\partial \varphi_i(\cdot, u_{-i})}{\partial u_{ij}}(u_i) \cdot y_{ij} = \sum_{i=0}^M \langle \nabla(\varphi_i(\cdot, u_{-i}))(u_i), y_i \rangle. \end{aligned}$$

□

With this formula for the directional derivatives, we can prove optimality for certain stationary points.

Theorem 5.49. *Let $u^* \in \text{int}(\mathbb{R}_+^{n'})$ be a stationary point of φ such that φ is continuously differentiable at u^* . Then u^* is a minimizer of φ .*

Proof. The structure of this proof follows [War63], Theorem 2.3.

Since u^* is a stationary point, we have that for $i = 0, 1, \dots, M$ and for all $y \in \mathbb{R}_+^{n-2}$

$$\varphi_i(u_i^*, u_{-i}^*) = \varphi(u^*) \leq \varphi(u_0^*, \dots, u_{i-1}^*, y, u_{i+1}^*, \dots, u_M^*) = \varphi_i(y, u_{-i}^*). \quad (5.25)$$

Let now v be an arbitrary point of $\mathbb{R}_+^{n'}$. φ is convex, hence for some point $u \in \mathbb{R}_+^{n'}$ it holds that

$$g_u : (0, 1) \rightarrow \mathbb{R}, \quad g_u(\delta) := \frac{\varphi(u^* + \delta(u - u^*)) - \varphi(u^*)}{\delta}$$

is a non-decreasing function of δ . Since φ is continuously differentiable at u^* , we can use Lemma 5.48 and get

$$\begin{aligned} \varphi(v) - \varphi(u^*) &\geq \lim_{\delta \searrow 0} \frac{\varphi(u^* + \delta(v - u^*)) - \varphi(u^*)}{\delta} \\ &= D_{v-u^*}(\varphi)(u^*) = \langle \nabla(\varphi)(u^*), v - u^* \rangle \\ &= \sum_{i=0}^M \langle \nabla(\varphi_i(\cdot, u_{-i}))(u_i), v_i - u_i^* \rangle. \end{aligned} \quad (5.26)$$

Now, for $i = 0, \dots, M$ plug $y = u_i^* + \delta(v_i - u_i^*)$ into (5.25). Then

$$\lim_{\delta \searrow 0} \frac{\varphi_i(u_i^* + \delta(v_i - u_i^*), u_{-i}^*) - \varphi_i(u_i^*, u_{-i}^*)}{\delta} \geq 0$$

and thus

$$\langle \nabla(\varphi_i(\cdot, u_{-i}))(u_i), v_i - u_i^* \rangle = \lim_{\delta \searrow 0} \frac{\varphi_i(u_i^* + \delta(v_i - u_i^*), u_{-i}^*) - \varphi_i(u_i^*, u_{-i}^*)}{\delta} \geq 0$$

Plugging this into (5.26) for $i = 0, \dots, M$ it follows that $\varphi(v) - \varphi(u^*) \geq 0$. □

Theorem 5.49 gives us convergence to a minimizer of φ . However, φ was solely constructed with the purpose of finding a minimizer of f in the orthant $\mathcal{O} \subset \mathcal{T}_n$ that contains an optimal solution to (P_{Med}) . The following corollary builds the bridge from our auxiliary Euclidean function φ to our objective function f in tree space.

Corollary 5.50. *Let $u^* \in \text{int}(\mathbb{R}_+^{n'})$ be a stationary point of φ such that φ is continuously differentiable at u^* . Then $\psi^{-1}(u_0^*)$ is a minimizer of f .*

Proof. Assume $U := \psi^{-1}(u_0^*)$ is not a minimizer of f . We know that f is a geodesically convex function and that a minimizer of f exists. Let X^* be a minimizer of f and $\gamma : [0, 1] \rightarrow \mathcal{T}_n$ parametrize the geodesic from U to X^* , $\Gamma(U, X^*)$. Due to the geodesic convexity of f (Definition 5.23), we get that for any $\lambda \in (0, 1)$

$$f(\gamma(\lambda)) \leq (1 - \lambda)f(\gamma(0)) + \lambda f(\gamma(1)) = (1 - \lambda)f(U) + \lambda f(X^*) < f(U) \quad (5.27)$$

since $f(X^*) < f(U)$. Due to Theorem 5.49, u^* is a minimizer of φ , so it is in particular a local minimizer. Thus, there exists an $\varepsilon > 0$ such that for all $w \in B_\varepsilon(u^*)$ we have $\varphi(w) \geq \varphi(u^*)$. Moreover it holds that $u^* \in \text{int}(\mathbb{R}_+^{n'})$, so there exists a $\bar{\lambda} > 0$, such that $\psi(\gamma(\lambda)) \in \text{int}(\mathbb{R}_+^{n-2})$ for all $\lambda \in [0, \bar{\lambda})$.

Now, define $v = (v_0, P_{T_1}(v_0), \dots, P_{T_M}(v_0))$ where $v_0 = \psi(\gamma(\lambda'))$, where $\lambda' \in [0, \bar{\lambda})$ can be chosen small enough such that $v \in B_\varepsilon(u^*)$, because the P_{T_i} are continuous. Then we use (5.27) for λ' to get

$$f(U) > f(\gamma(\lambda')) = f(\psi^{-1}(v_0)) = \varphi(v) \geq \varphi(u^*) = f(\psi^{-1}(u_0^*)) = f(U),$$

contradicting that U is no minimizer of f . □

Now, we are finally able to state a theorem for convergence of the BPA to an optimal solution for the problem (P_{Med}) . Therefor we still need all assumptions that ensure convergence to a stationary point of the BPA from Theorem 5.44 and moreover we need to make sure that the block-wise coordinate descent version of the BPA does not converge to a non-optimal stationary point in the boundary of the orthant.

Theorem 5.51. *Assume that there exists a unique optimal solution X^* to (P_{Med}) in $\text{int}(\mathcal{O})$ and that $\{T_1, \dots, T_M\}$ contains no balance point. Moreover, assume that for all $u_0 \in \mathbb{R}_+^{n-2}$ it holds that $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear. Denote $U_B := \min_{X \in \partial \mathcal{O}} f(X)$ and let $u^{(0)}$ be the starting point of the BPA. Then, if $\exists \ell' \in \mathbb{N}_0$ such that an iterate of the BPA $u^{(\ell')}$ satisfies $\varphi(u^{(\ell')}) < U_B$, then $u^{(\ell)} \rightarrow u^* \in \mathbb{R}_+^{n'}$ such that $\psi^{-1}(u_0^*) = X^*$*

Proof. First of all, since all assumptions of Theorem 5.44 hold, it follows that $(u^{(\ell)})$ has a subsequence $(u^{(\ell_m)})_{m \in \mathbb{N}}$ that converges to a stationary point \bar{u} . Since $\varphi(u^{(\ell)})$ is non-increasing in ℓ , $u^{(\ell')} < U_B$ holds for all $\ell \geq \ell'$, thus $\varphi(\bar{u}) < U_B$. This implies, that

$\varphi(\bar{u}) \notin \partial(\mathcal{O})$. Since \bar{u} is a stationary point it follows that $\psi^{-1}(\bar{u}_0)$ is a balance point due to Theorem 5.28. Hence, the prerequisites imply that $\psi^{-1}(\bar{u}_0) \notin \{T_1, \dots, T_M\}$. Since \bar{u} is a stationary point determined by the BPA, Proposition 5.47 implies that $v = (\bar{u}_0, R_{T_1}(\bar{u}_0), \dots, R_{T_M}(\bar{u}_0)) \in \text{int}(\mathbb{R}_+^{n'})$ is a stationary point as well. Moreover, we have established that $\psi^{-1}(\bar{u}_0) \notin \{T_1, \dots, T_M\}$, so Lemma 5.46 implies that φ is continuously differentiable at v . Finally, Corollary 5.50 implies that $\psi^{-1}(v_0) = \psi^{-1}(\bar{u}_0)$ is a minimizer of f .

It remains to be shown that not only the subsequence, but the whole sequence of iterates $(u^{(\ell)})_{\ell \in \mathbb{N}}$ converges. As X^* is the unique minimizer of f it follows that only points $u \in \mathbb{R}_+^{n'}$ with $u_0 = \psi(X^*)$ can be optimal for φ . As $u_0^{(\ell_m)} \rightarrow \psi(X^*)$ and the whole sequence $\varphi(u^{(\ell)})$ is non-increasing in ℓ it follows that $u_0^{(\ell)} \rightarrow \psi(X^*) = \bar{u}_0$ and as $P_{T_i}(\bar{u}_0) = \bar{u}_i$, Lemma 5.43 implies that $u^{(\ell)} \rightarrow \bar{u}$. \square

To conclude the section we want to state a theorem that offers a neighborhood for which the BPA is guaranteed to converge when all prerequisites of Theorem 5.51 are satisfied. To this end we need the notion of (lower) level sets:

Definition 5.52 ([Roc15]). *The (lower) level set of a function $g : \mathcal{X} \rightarrow \mathbb{R}$ with respect to level $\alpha \in \mathbb{R}$ is defined as*

$$\text{lev}_\alpha(g) := \{x \in \mathcal{X} : g(x) \leq \alpha\}.$$

As the level sets of convex functions are convex for all $\alpha \in \mathbb{R}$ ([Roc15]), we get a convex neighborhood of the unique optimal solution for which the BPA converges:

Corollary 5.53. *Assume that there exists a unique optimal solution X^* to (P_{Med}) in $\text{int}(\mathcal{O})$ and that $\{T_1, \dots, T_M\}$ contains no balance point. Moreover, assume that for all $u_0 \in \mathbb{R}_+^{n-2}$ it holds that $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ are not collinear. Denote $U_B := \min_{X \in \partial \mathcal{O}} f(X)$ and let $u^{(0)}$ be the starting point of the BPA. Then, if $u^{(0)} \in \text{lev}_\alpha(\varphi)$ with $\alpha < U_B$, then $u^{(\ell)} \rightarrow u^* \in \mathbb{R}_+^{n'}$ such that $\psi^{-1}(u_0^*) = X^*$*

Proof. $U_B > \alpha \geq \varphi(u^{(0)}) = f(\psi^{-1}(u_0^{(0)}))$, so we may apply Theorem 5.51 with $\ell' = 0$. \square

After having proven a convergence result for the BPA, we briefly come back to the GBPH and discuss its convergence. Besides any other assumptions for Theorem 5.51, we always require to know the orthant containing an optimal solution, in which we then apply the BPA.

The GBPH is designed in such a way, that we are guaranteed to apply the BPA in an orthant containing an optimal solution, as the preprocessing of the GBPH only rules out orthants that do not contain an optimal solution. Thus, when all assumptions of Theorem 5.51 hold, then the GBPH yields the optimal solution to the problem:

Theorem 5.54. *Let $\mathbf{T} = \{T_1, \dots, T_M\}$ be given. Suppose there exists a unique optimal solution $X^* \in \text{int}(\mathcal{O}) \subset \mathcal{T}_n$ and that \mathbf{T} and \mathcal{O} satisfy the prerequisites of Theorem 5.51. Then the GBPH yields the optimal solution.*

Proof. It is clear, that the preprocessing of the GBPH does not discard \mathcal{O} . Thus, \mathcal{O} is one of the candidate orthants for which the BPA gets applied. Theorem 5.51 then implies that the BPA yields the unique optimal solution in \mathcal{O} . Thus, this balance point is also found by the GBPH as it is the balance point with the minimal objective value. \square

Obviously, Theorem 5.51 and Theorem 5.54 require a lot of assumptions. Some of the assumptions rule out rather unlikely special cases, e.g., $P_{T_1}(u_0), \dots, P_{T_M}(u_0)$ will seldom be collinear when M is big and the dimension of tree space $n > 4$, as for $n > 4$ the boundaries of orthants are already two-dimensional. Also, the case that the BPA converges to a non-optimal stationary point at the boundary seems not to be to likely: When started in the “middle” of the orthant, i.e., equidistant from all bounding hyperplanes, then there is a good chance that the BPA iterates into a convex level set as in Corollary 5.53 and then converges to the optimal solution.

The crucial assumption that is most likely to be violated, is that the unique optimal solution is contained in the interior of an orthant, which is rarely the case, when the data is very spread and the number of facilities is high, as discussed in the Experiment Section 5.2.4. Nonetheless, the results of the heuristic for such instances were pretty good.

Of course, it is practically impossible to check if all assumptions of Theorem 5.54 hold. So when applying the GBPH, we do not know whether it yields the optimal solution or not. Nonetheless, whenever the BPA or the GBPH output a solution, that is contained in the interior of an orthant and is not a facility, then we know that it is an optimal solution to (P_{Med}) .

5.4 Restrictions of the BPA

In the previous sections we have tediously proven a convergence result for the BPA under specific assumptions. Moreover, we have seen in the experiments in Section 5.2.4 that the BPA seems to perform quite well, even in cases where it converges to the boundary. Still, there is no theoretical justification or guarantee that the BPA performs well for the boundary cases. In this section we sketch why the above-mentioned assumptions are necessary for the convergence theorem of the BPA and show how and why the BPA fails, when the assumptions do not hold.

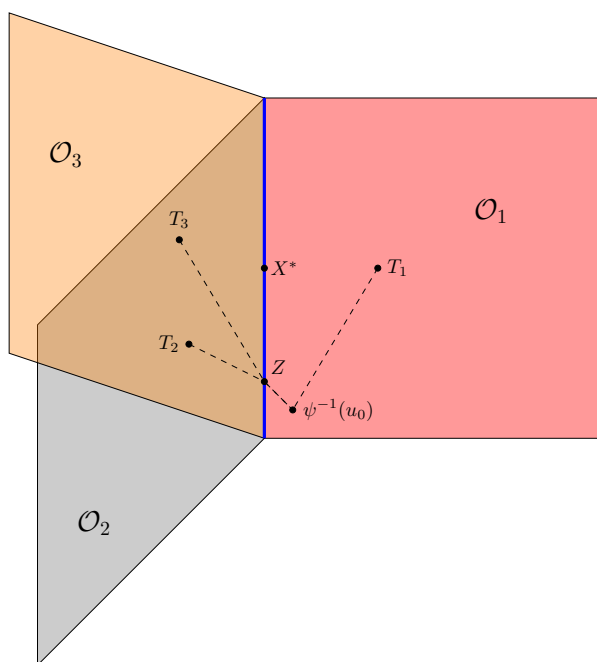


Figure 5.13: An instance where the BPA in \mathcal{O}_1 with starting point u_0 converges to the non-optimal balance point $Z \in \partial\mathcal{O}_1$.

Example 5.4.1. Consider the four splits on $\{0, 1, 2, 3, 4\}$

$$\begin{aligned} s_1 &= (\{1, 2\}|\{0, 3, 4\}), & s_2 &= (\{1, 3\}|\{0, 2, 4\}), \\ s_3 &= (\{2, 3\}|\{0, 1, 4\}), & s_4 &= (\{1, 2, 3\}|\{0, 4\}) \end{aligned}$$

with which we define the three orthants \mathcal{O}_i , for $i = 1, 2, 3$, that have the split sets

$$S_1 = \{s_1, s_4\}, \quad S_2 = \{s_2, s_4\}, \quad S_3 = \{s_3, s_4\},$$

respectively. The three orthants form a sort of 'open book' (see [HHL⁺13]) structured subset of \mathcal{T}_4 , see Figure 5.13 for an illustration. In each orthant \mathcal{O}_i we have one sample tree T_i , so $\mathbf{T} = \{T_1, T_2, T_3\}$ with

$$T_1 = ((s_1, s_4), (2, 3)), \quad T_2 = ((s_2, s_4), (2, 3)), \quad T_3 = ((s_3, s_4), (2, 3)).$$

The optimal solution to the problem is $T^* = ((s_4), (3))$, which is contained in the boundary of all of the three orthants. We want to apply the BPA in orthant \mathcal{O}_1 with the starting point $u_0^{(0)} = (0.5, 0.5)^t$. The geodesic from $\psi_{\mathcal{O}_1}^{-1}(u_0^{(0)})$ to T_2 and T_3 always contains the split s_4 , since it is contained in all splits. Its length changes with constant speed, so the exit point is $Z = ((s_4), (1))$ for both T_2 and T_3 , due

to the symmetry. This implies that $P_{T_2}(u_0^{(0)}) = (0, 1)^t = P_{T_3}(u_0^{(0)})$. Since $T_1 \in \mathcal{O}_1$ we have $P_{T_1}(u_0^{(0)}) = \psi(T_1) = (2, 3)^t$. Now, by definition of the BPA, $u_0^{(1)} = \arg \min_{y \in \mathbb{R}_+^2} \sum_{i=1}^3 \|y - P_{T_i}(u_0^{(0)})\|_2$. Since $(0, 1)^t$ has weight 2, it is the unique optimal solution to this problem, so $u_0^{(1)} = (0, 1)^t$. Thus, $\psi_{\mathcal{O}_1}^{-1}(u_0^{(1)}) = Z$.

But then we have that Z is the exit point for T_2 and T_3 once more and we still have $P_{T_1}(u_0^{(1)}) = \psi_{\mathcal{O}_1}(T_1)$. Hence we have the same representatives and the same optimization problem, that leads to $u_0^{(2)} = u_0^{(1)}$ being the optimal solution. That means that we have found a balance point and of course it follows that $u_i^{(2)} = P_{T_i}(u_0^{(2)}) = Z = P_{T_i}(u_0^{(1)}) = u_i^{(1)}$ holds, so we also arrive at a stationary point.

Unfortunately, this balance point $(1, 0)$ is not the optimal solution, which would be $(3, 0)$.

This example does not only show, that there is no guarantee for convergence to the optimal solution when it is contained in the boundary, it also shows that the BPA does depend on the starting point, when the optimal solution is contained in the boundary: If one would choose to start in \mathcal{O}_1 with the tree with weight vector $(1, 3)$, then one would get the optimal solution in the first step of the iteration, as the geodesics to T_2 and T_3 would exit \mathcal{O}_1 exactly at X^* . Hence, it is necessary to take caution when interpreting results of the BPA when it converged to the boundary of an orthant.

Another interesting idea to improve/generalize the BPA is to continue to search in lower-dimensional orthants, i.e., to apply the BPA in a non-maximal orthant. The hope is that, when the BPA converges to the boundary of a maximal orthant, one can improve the solution when continuing the procedure in this lower-dimensional orthant. Since we may still easily define exit points in these orthants and the minimizer of the Euclidean median problem would also be contained in the lower-dimensional orthant in each step, it is theoretically no problem to apply the method. In fact, we will do this in Chapter 6 in order to further improve solutions in lower-dimensional orthants. The next example, however, shows that we cannot guarantee an improvement and even less convergence with this approach as the next example shows.

Example 5.4.2. *We give an example for a non-optimal balance point in a non-maximal orthant in \mathcal{T}_4 . Consider the following five splits for $n = 4$:*

$$\begin{aligned} s_1 &= (\{1, 2\}|\{0, 3, 4\}), & s_2 &= (\{3, 4\}|\{0, 1, 2\}) \\ s_3 &= (\{1, 2, 3\}|\{0, 4\}), & s_4 &= (\{1, 3\}|\{0, 2, 4\}) & s_5 &= (\{2, 4\}|\{0, 1, 3\}) \end{aligned}$$

Assume we want to find the median for the points $\{T_1, T_2, T_3\}$, where

$$T_1 = ((s_1, s_2), (4, 2))$$

$$T_2 = ((s_1, s_3), (4, 2))$$

$$T_3 = (s_4, s_5), (2, 2).$$

We show that the tree $P = ((s_1), (4))$ is a balance point in the non-maximal orthant defined by split s_1 , but P is not optimal. The situation is illustrated in Figure 5.14. That P is not optimal can immediately be seen by comparing objectives with, for

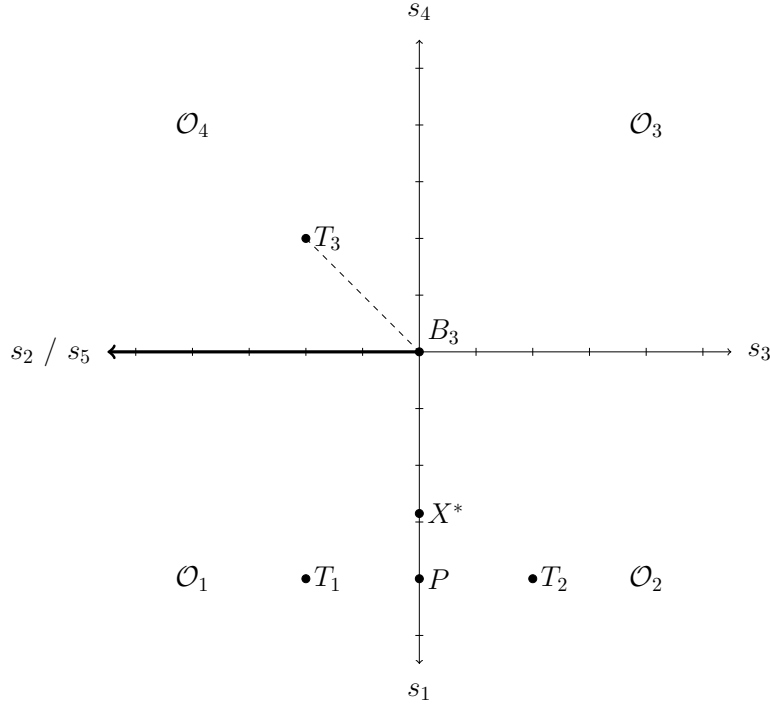


Figure 5.14: An example for a non-optimal balance point in a non-maximal orthant in \mathcal{T}_4 .

example, $X = ((s_1), (2))$:

$$f_{\text{Med}}(P) = 2 + 2 + (4 + \sqrt{8}) \approx 10.83 > 10.49 \approx \sqrt{8} + \sqrt{8} + (2 + \sqrt{8}) = f_{\text{Med}}(X)$$

So it remains to show that P is a balance point, when we apply the BPA in the non-maximal orthant \mathcal{O} that satisfies $\text{Split}(\mathcal{O}) = \{s_1\}$. Since s_2 and s_3 are compatible with s_1 it follows that the exit point for both T_1 and T_2 is P itself. Due to the incompatibility of s_1 with s_4 , the exit point for T_3 is B_3 . The median of $\{P, P, T_3\}$,

or to be more precise, the weighted median of $\{P, T_3\}$, where P has weight 2 and T_3 weight 1, is P itself. Thus, P is the median of its exit points, i.e., a balance point.

This example is designed to show the possibility of such non-optimal balance points but it does not really yield problems for the BPA in general: When starting the BPA in \mathcal{O}_1 or \mathcal{O}_2 , it actually yields the optimal solution, which is $X^* = ((s_1), (2.85))$.

The example does not only show that the existence of a non-optimal balance point in a non-maximal orthant, i.e., the boundary of an orthant, but it also gives some intuition as to why the BPA fails here. The reason are the double compatible splits. For a tree Y in a lower-dimensional orthant it may happen that a lot of sample trees T_i contain splits that are compatible with $\text{Split}(Y)$, which results in Y being the exit point, which we simply call B_i here, with respect to these trees T_i , i.e., $Y = B_i$. Then the optimality condition for the weighted median problem

$$\text{Test}_k := \left\| \sum_{\substack{i=1 \\ B_i \neq Y}}^M \frac{(Y - B_i)}{\|Y - B_i\|_2} \right\|_2 \leq |\{T_i \in \mathbf{T} : B_i = Y\}|,$$

(see [Kuh73]) can easily be met when the number of trees T_i for which $B_i = Y$ holds is high. This then results in Y being a balance point.

Another problematic case is, where the BPA actually converges to a non-optimal solution in the boundary of an orthant, even when the unique optimal solution to the problem is contained in the interior of that orthant:

Example 5.4.3. Figure 5.15 depicts a situation, where non-optimal stationary points may also exist in the boundary of the orthant that contains the unique optimal solution in its interior. In fact, every tree T that is contained in the lower-dimensional orthant $\mathcal{O}(\{s_2\})$ is a balance point, as the exit point will be $\psi(T)$ itself with a weight of 2, so $\psi(T)$ will solve the resulting Euclidean median problem.

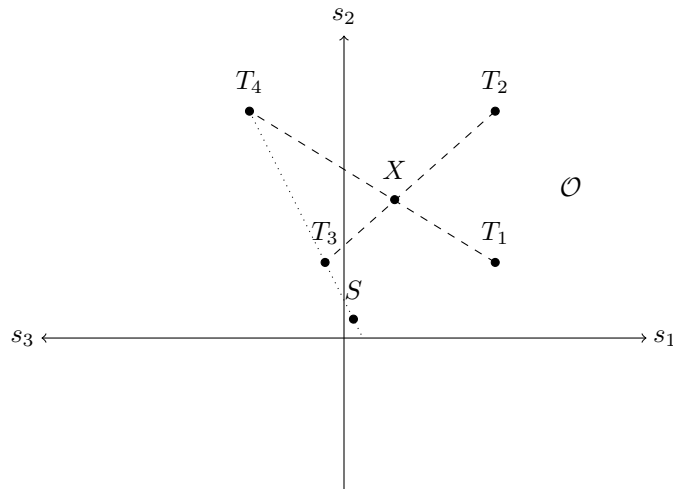


Figure 5.15: Non-optimal stationary points may also exist in the boundary of the orthant that contains the unique optimal solution in its interior.

In Figure 5.15, the unique optimal solution to the problem is marked by X , which is the intersection of the diagonals of the polygon, see [Pla05]. In the unlucky case that the BPA is started at $\psi(S)$, or any other point on the dotted line in \mathcal{O} , then the exit points for T_3 and T_4 will coincide and the BPA yields a non-optimal stationary point in $\partial(\mathcal{O})$.

Example 5.4.2 and Example 5.4.1 show where the BPA might fail and we also explain what causes this. Of course it would have been desirable to have a proof that the BPA also finds the best solution in the boundary of the orthant. Even though the examples show that this does not work, the BPA is still able to yield sensible information in these cases. First of all, when we converge to the boundary, we do not know if we have found an optimal solution; but we do know that the optimal solution is not contained in the interior of that orthant. If this behavior shows for several maximal orthants and the BPA always yields the same lower-dimensional orthant, then this is a strong hint that the optimal solution is contained within this orthant, which is a valuable information on its own. Secondly, having found this tendency one may use the BPA in the determined lower-dimensional orthant. Again, as the exit points may tend to be clutched at a single point as in Example 5.4.1 it is not clear if an improvement may be found, but it is also not ruled out. In Chapter 6 we show that the BPA significantly improves the solution in lower-dimensional orthants for the given real data set.

6 Real Data Example

6.1 Data Set

In this chapter we calculate the median of a real data set. The Apicomplexa data set was introduced in [KWK08] and consists of 268 phylogenetic trees with eight species. Seven of these species are in the *Apicomplexa* phylum, “*whose most infamous members [...] are the causative agents of Malaria*” ([KWK08]), one other species, a ‘free-living ciliate’ is included as an ‘outgroup’.

[WYH17] investigated this data set with a normalizing kernel algorithm in tree space that detected 16 of the 268 gene trees as outliers that we exclude from the data set as in [YZZ17]. Thus, we are left with 252 phylogenetic trees. The names of the eight species of which we would like to find the species tree may be found in the table in Table 6.1 together with their labels and abbreviations.

Label	Abbreviation	Name
1	Bb	Babesia bovis
2	Cp	Cryptosporidium parvum
3	Et	Eimeria tenella
4	Pf	Plasmodium falciparum
5	Pv	Plasmodium vivax
6	Ta	Theileria annulata
7	Tg	Toxoplasma gondii
8	Tt	Tetrahymena thermophila

Table 6.1: The abbreviations and labels of the *Apicomplexa* data set.

First we give a very rough depiction of the tree topologies in the data set, thereby omitting any information about the weights of the splits of the given trees, to keep the depiction as clear as possible. According to [KWK08], the data set consists of 32 tree topologies. When analyzing the data set we found a lot of splits with a length of 0, which probably comes from a specific setup and choice of species, with ‘Tt’ being a so-called *free living species*, which does not belong to the phylum *Apicomplexa*. As \mathcal{T}_n does not really contain trees with split lengths of 0, we removed these to get a better picture of the data set. On a side note, let T be a tree that contains

some artificially added splits of length 0 and let T' be the tree that is obtained when removing the splits of length 0 from T . Then $d(T, T') = 0$ (by the Euclidean distance formula), so $T = T'$. To avoid such ambiguities, \mathcal{T}_n is defined so that every split $s \in \text{Split}(T)$ also has a positive weight assigned to it.

It turns out that after removing these splits we end up with a total of 57 tree topologies. This may seem counter intuitive at first glance, but when removing these splits with 0 length this may result in two topologies for two trees that shared a topology before removing the split. This case occurs when the splits that had length 0 were different splits.

In Figure 6.1 we have depicted the four most common tree topologies. The tree topologies do not depict any indication on the edge weights, the lengths of the edges are simply chosen to let the trees look as neat as possible.

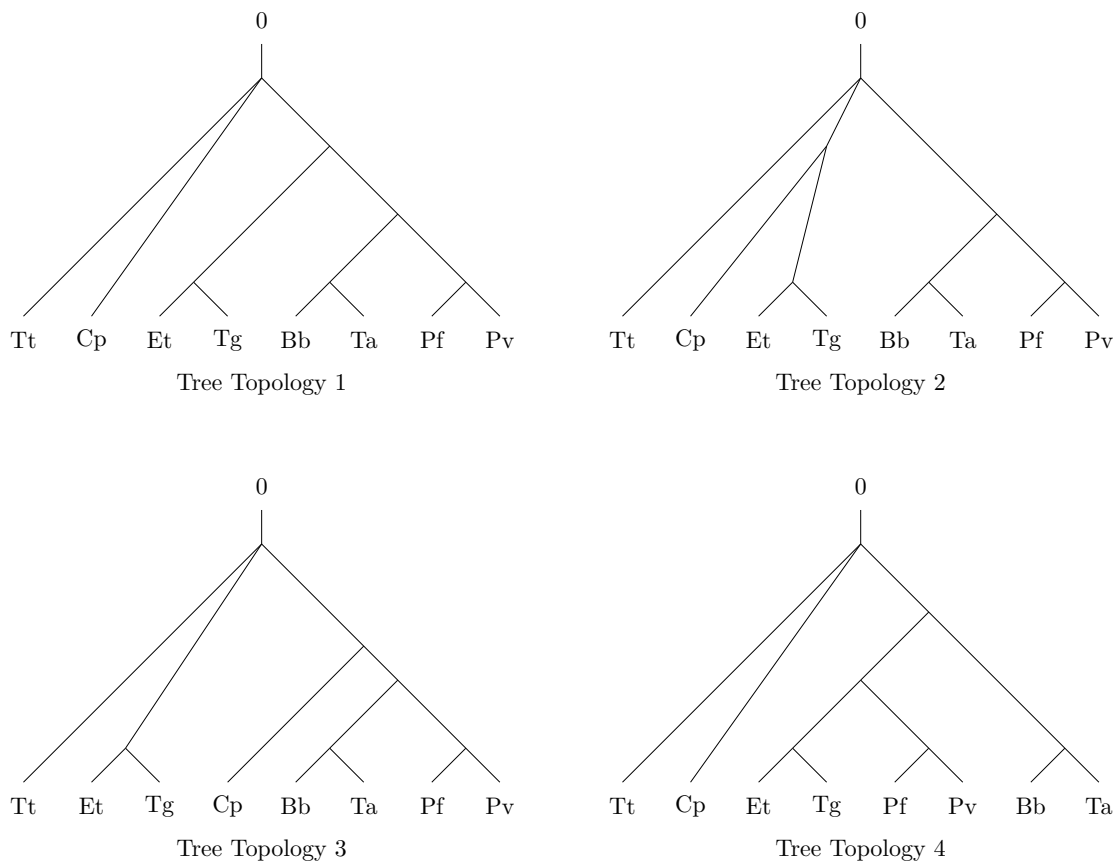


Figure 6.1: The four most common tree topologies of the sample trees from the Apicomplexa data set.

The whole data set may be found in Chapter 7, where trees are given in the Newick format, which includes the lengths of the splits.

6.2 Experiments

The experiments we conduct in this section pursue the objective of evaluating how good the BPA performs on a real data set. Not only does this show whether or not the algorithm performs well on a reasonably big data set in comparison to the rather small examples in Section 5.2.4 but also it allows for a first interpretation what the median incorporates from the structure of the data set, i.e., if the median is a helpful concept.

It is important to mention here, that we do not know the median of the presented Apicomplexa data set, as this has not yet been calculated to our knowledge. So, in order to be able to compare the results of our algorithm we implemented the *cyclic order version* of the *Proximal Point Algorithm* (PPA) from [Bac14a]. The PPA is a method that calculates medians in Hadamard spaces, thus it may also be applied in tree space and to our knowledge, this is the only other algorithm to determine medians in tree space.

In order to reasonably compare PPA and BPA we first investigate how to best choose starting points and parameters that may influence the outcomes of the algorithms. Then we compare their best runs and overall results for varying parameters. Note that we do not aim for finding the absolute best parameters in order to show how well the algorithms perform, we rather try to find parameters that enable us to carefully investigate and interpret the results.

To this end we apply the algorithms with several different settings and summarize what the best choice of settings is for each algorithm. Furthermore we conjecture how to best choose such parameters in a more general setting, when we are not bound to the present data set. Last but not least, the following experiments aim for finding an approximate median for the data set so we try to find the best method and setting to determine the best solution we can find.

6.2.1 PPA Experiments

We start with the Proximal Point Algorithm of [Bac14a], which we stated in Algorithm 6, Section 5.2.4.

In the following we carry out three different kinds of experiments for the PPA. We investigate the influence of the amounts of cycles, Table 6.2, the influence of different starting points, Tables 6.3 - 6.5, and the influence of the sequence (λ_k) in Table 6.6. The first and most obvious feature to investigate is how the amount of cycles influences the outcome of the algorithm. To this end we use a fixed starting point and

a fixed sequence (λ_k) and compare the objectives and the computation times after different amounts of cycles. The starting point we chose has the most common tree topology of the trees in the data set, which is Tree Topology 1 from Figure 6.1. The split set of this topology is

$$\mathcal{S}_1 = \{(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\}), \\ (\{1, 3, 4, 5, 6, 7\}, \{0, 2, 8\})\},$$

and the weight vector of the tree we use is

$$\bar{w} = (0.214165, 0.239492, 0.114332, 0.083201, 0.061012)^t.$$

We denote the starting point as $\bar{X} = (\mathcal{S}_1, \bar{w})$. The results of the experiments are summarized in Table 6.2. It can be seen quite nicely that due to the cyclic nature of the algorithm, the computation time is proportionate to the amounts of cycles.

# of cycles	time in s	objective value
10	4.14	76.401
100	38.68	72.297
1000	373.28	72.203
10000	3735.40	72.182

Table 6.2: Results of the PPA for the Apicomplexa data set. The sequence that was used for the PPA is $\lambda_k = \frac{10^{-3}}{k+1}$, the starting point is \bar{X} .

It may be seen that the algorithm makes significant progress in the beginning, whereas from 100 cycles to 1000 cycles as well as from 1000 to 10000 cycles the progress is rather small, especially in relation to the additional amount of computational effort.

It is important to note that the tree topology does not change at all. That is probably due to the *Tree Topology 1* actually being an optimal tree topology, i.e., there exists a median that has this topology, or at least it is very close to it. So the tree topology that we have after each cycle is the most common among the input trees, i.e., it is Tree Topology 1 from Figure 6.1.

Having tested a starting point that has a different topology we realized that the topology also stagnates after a few cycles. We use the sequence $\lambda_k = \frac{10^{-3}}{k+1}$ and the starting point $\bar{Y} = (\mathcal{S}_4, v)$ with the split set

$$\mathcal{S}_4 = \{(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{3, 4, 5, 7\}, \{0, 2, 3, 7, 8\}), \\ (\{1, 3, 4, 5, 6, 7\}, \{0, 2, 8\})\},$$

which is the split set of Tree Topology 3 from Figure 6.1 and the weight vector

$$\bar{v} = (0.393798, 0.273573, 0.206168, 0.104214, , 0.249161)^t.$$

It turns out that the topology of the trees also reaches Tree Topology 1 after only three cycles and does not change after that.

That the topology does not seem to change anymore after a certain amount of cycles is a very important observation for the BPA. When the topology is fixed from some point on it is very likely that the optimal solution is contained within this orthant. We may then use that topology to choose certain starting trees for the BPA. We will make use of this observation in Section 6.2.2.

In the next series of experiments the objective is to find out, how strongly Algorithm 6 depends on the given starting point. This question turns out to be somewhat intertwined with the choice of a reasonable scaling factor for the step size, as we will explain in the following. We perform experiments for a fixed set of five different starting points and show the influence on the results when choosing different step sizes. As for the starting points, we choose representatives from the data set of the four most common tree topologies, compare Figure 6.1, to get reasonably good, but still different starting points.

As earlier, for Tree Topology 1 and 4 we formally state the topologies of Tree Topology 2 and Tree Topology 3 as split sets:

$$\begin{aligned} \mathcal{S}_2 = & \{(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ & (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\}), \\ & (\{2, 3, 7\}, \{0, 1, 4, 5, 6, 8\})\}, \\ \mathcal{S}_3 = & \{(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ & (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\}), \\ & (\{1, 2, 4, 5, 6\}, \{0, 3, 7, 8\})\}, \end{aligned}$$

The weights that we use for the starting points stem from representatives of this topology in the data set:

$$\begin{aligned} w_1 &= (0.214165, 0.239492, 0.114332, 0.083201, 0.061012)^t \\ w_2 &= (0.427082, 0.23773, 0.074412, 0.10792, 0.045042)^t \\ w_3 &= (0.397007, 0.18038, 0.225526, 0.199062, 0.05536)^t \\ w_4 &= (0.393798, 0.273573, 0.206168, 0.104214, 0.249161)^t \end{aligned}$$

With this, we define our starting points as $X_i := (\mathcal{S}_i, w_i)$, $i = 1, \dots, 4$. We will also use $0 \in \mathcal{T}_n$ as an additional starting point and reference.

The first try for this experiment was conducted with the sequence $\lambda_k = \frac{1}{k+1}$. We received the same result for all starting points, as may be seen in Table 6.3 which is actually due to all starting points returning the exact same tree.

starting point	time in s	objective value
X_1	37.5	72.297
X_2	37.9	72.297
X_3	38.2	72.297
X_4	37.3	72.297
0	37.5	72.297

Table 6.3: Results of the PPA for the Apicomplexa data set. The sequence that was used is $\lambda_k = \frac{1}{k+1}$, the number of cycles was set to 100.

This phenomenon results from the structure of the data set and the definition of t_k^η : Since the weights of all trees in the data set are rather small, their distance is small as well, because it is bounded by $d(T_i, T_j) \leq \|t_i\|_2 + \|t_j\|_2$. The maximal distance between two trees in the data set is actually only 1.17 and the majority of pairwise distances is smaller than 1.

That makes it so that $t_0^1 = \min \left\{ 1, \frac{\lambda_0}{d(X, T_1)} \right\} = \min \left\{ 1, \frac{1}{d(X, T_1)} \right\} = 1$, as $\frac{1}{d(X, T_1)} \geq 1$ holds for all investigated starting points. Thus, the algorithm yields the same results for all starting points.

In order to rule this out, so that our different starting points actually have an influence on the algorithm, we scale down the λ_k by a factor $c > 0$. As a rather extreme try to get rid of this independent behavior of the starting points, we choose $c = 10^{-5}$ such that $t_1^0(X) = \min \left\{ 1, \frac{10^{-5}}{d(X, T_1)} \right\} < 1$.

Table 6.4 contains the results of this experiment and shows that the resulting step sizes are way too small:

starting point	time in s	objective value
X_1	3511	80.498
X_2	3801	84.984
X_3	3871	96.217
X_4	3798	105.483
0	3617	115.422

Table 6.4: Results of the PPA for the Apicomplexa data set. The sequence that was used is $\lambda_k = \frac{10^{-5}}{k+1}$, the number of cycles was set to 10000.

Theorem 5.18 states that the sequences of iterates do converge, no matter how small the scaling factor $c > 0$ becomes, but here we can see how slow the convergence may get, as we let the PPA run for 10000 cycles for each starting point and still have very bad objective values. Comparing these results to the run with 10000 cycles in Table 6.2 we can see that the choice of such a small step size is extremely maleficial. That shows, that it is not that easy to improve the result of PPA by only choosing different starting points, one also needs to be in the right range of scaling factors. After having conducted several different runs of the algorithm for varying starting points and step sizes, it seems that a scaling factor of $c = 10^{-3}$ works best empirically to compare different starting points. This is for two reasons, the first one being that for $c = 0.01, c = 0.1$ etc. we have $t_1^0(X) = 1$ for $X \in \{X_1, X_2, X_3, X_4, 0\}$ and they all yield the same result. The other reason is that for $c = 10^{-4}$ we get the same slow converging behavior as for $c = 10^{-5}$, just not as extreme. The results for our five starting points, a scaling factor of $c = 10^{-3}$ and a maximum number of cycles of 10000 are summarized in Table 6.5.

starting point	time in s	objective value
X_1	3666	72.180626
X_2	3684	72.180629
X_3	3674	72.180630
X_4	3673	72.180635
0	3675	72.180666

Table 6.5: Results of the PPA for the Apicomplexa data set. The sequence that was used for the PPA is $\lambda_k = \frac{10^{-3}}{k+1}$, the number of cycles was set to 10000.

The results clearly point out, that when taking a reasonable step size and a large number of cycles, the results are close to independent from the starting point. In

order to be able to depict the different outcomes we give six decimal places, so all resulting objective values are really close.

All in all, the influence of different starting points is not too big for this data set and the three experiments have shown that the step size has a much bigger impact on the outcome.

Thus we try to compare varying scaling factors c for the step size. We investigate the range from $c = 10^{-4}$ to $c = 10^3$, always multiplying c by 10. We execute the experiment with a fixed starting point, the origin, and a fixed number of 100 maximum cycles. One may see in Table 6.6 that it turns out that the best scaling

factor c	time in s	objective value
0.0001	36.02	98.920
0.001	36.83	72.212
0.01	36.83	72.182
0.1	37.42	72.203
1	37.12	72.297
10	37.74	76.401
100	36.73	125.444
1000	36.41	125.444

Table 6.6: Results of the PPA for the Apicomplexa data set. The number of cycles was set to 100, the starting point was the origin $X_s = 0$.

factor seems to be around $c = 0.01$ for the data set at hand. When the scaling factor gets bigger, then the first cycles should only jump from one sample tree to another as the bigger $\lambda_k = \frac{c}{k+1}$ implies that $t_k^\eta = \min \left\{ 1, \frac{\lambda_k}{d(X, T_\eta)} \right\}$ is bigger; hence we get $t_k^\eta = 1$ for a lot of k in the beginning before something starts to happen. This behavior culminates in the same objective value for $c = 100$ and $c = 1000$; the scaling factors are so big that $t_M^{100} = 1$ which results in the output of the sample tree T_M and its objective 125.444 for both values of c .

A last important point to notice is that the ‘optimal’ choice of the step size also strongly depends on the starting point. When one does not care for the structure of the data set and chooses an outlier as a starting point, one may not make any progress at all when the scaling factor is as small as $c = 0.01$. When we applied the PPA with this scaling factor and a maximal number of cycles of 1000 to the starting tree $X^b = (\mathcal{S}_1, (100, 100, 100, 100, 100))$ we received an objective of 51521 for the output tree

$$X_o^b = (\mathcal{S}_1, (91.581, 91.57, 91.563, 91.552, 91.553)),$$

so we are still light years away from an optimal solution after 1000 cycles. Hence, it is important to look at the data set before choosing a starting point and also choosing a fitting step size factor for the corresponding starting point.

We want to briefly summarize the results. The experiments have shown that for the data set at hand the most crucial factor for the performance of the algorithm was the step size. Without the correct step size, the algorithm makes very little progress, even when massively increasing the computational effort. The influence of different starting points may even be nullified by the step size, when it is not chosen suitably, as we have seen in Table 6.3.

It is, however, crucial to notice the dependence of the step size and the given data set. Our data set carries a lot of structure. The splits

$$\begin{aligned} &(\{4, 5\}|\{0, 1, 2, 3, 6, 7, 8\}), \\ &(\{1, 6\}|\{0, 2, 3, 4, 5, 7, 8\}), \\ &(\{3, 7\}|\{0, 1, 2, 4, 5, 6, 8\}) \end{aligned}$$

are contained in nearly all input trees and the split lengths from the trees are rather small, so the trees are not extremely far apart from each other. The maximal distance between two trees is actually only 1.17, as mentioned earlier. Thus it is not too surprising that a scaling factor $c = 10^{-2}$ is the best choice, as $\min\left\{1, \frac{c}{(k+1) \cdot d(X, T_\eta)}\right\}$ is more likely to be 1 for smaller k , when $d(X, T_\eta) \ll 1$ and c is not sufficiently small to balance this out.

So for our data set it is reasonable that $c < 1$ yield faster convergence than other c . For data sets, where the sample trees have big distances of, say 1000 in average, it is better to take a scaling factor $c > 1$, as one otherwise observes the same phenomenon as for our outlier example. This shows that there is no universal rule in choosing a good factor for the step size, the data set at hand needs to be investigated first.

6.2.2 BPA Experiments

In this subsection we conduct experiments for the BPA on the Apicomplexa data set. The first thing that needs to be clarified is that the BPA itself cannot solve this problem as it is a local improvement procedure that only operates on a fixed orthant. Hence, the BPA needs a starting point in a fixed orthant and since the BPA only yields trees in the orthant it was given, it even needs a starting point which is contained in an orthant with an optimal solution in order to perform well. This is also the reason why we started with the PPA experiments - the PPA yields an orthant with a very good solution if not an orthant containing the optimal solution. So the experiments we perform in this subsection are only sensible when the BPA is started after one has a result of the PPA, which we will assume in the following.

The best run of the PPA was found with starting point X_1 in the experiment summarized in Table 6.5. It returned the tree

$$X_P = (\mathcal{S}_1, (0.30847421, 0.18800566, 0.1120948, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})),$$

i.e., a tree with the most common Tree Topology 1, which has an objective of 72.180626. The most important thing to notice here is that the tree structure is unresolved. This immediately raises the question where the BPA should be started since we only developed theory for maximal orthants. Moreover, there may exist non-optimal balance points in non-maximal orthants, so it is not sensible to simply trust the outcome of the algorithm in non-maximal orthants, as we have already shown in Example 5.4.2.

The problems with the analysis of the BPA in a non-maximal orthant and the conjecture that an optimal solution of the median is to be found in a non-maximal orthant seem to suggest that the method fails. But in the following we will show that BPA performs quite well on such instances for several reasons. We show this via three different experiments. First, we simply apply the BPA to one tree for each of the three maximal orthants that *resolve* Tree Topology 1 that contains the median according to the PPA results, see Table 6.7. ‘Resolving’ means that we add compatible splits to the split set, such that we end up with a maximal split set, and thus in a maximal orthant. The second experiment investigates the influence of different starting points in a fixed maximal orthant, see Table 6.8. The third and most interesting experiment is to apply the BPA in non-maximal orthants, trying to improve solutions that the PPA yielded, see Table 6.9.

We start with the three topologies that resolve Tree Topology 1, which is the topology that the best output trees of PPA yield. Recall, that the split set of Tree Topology 1 is

$$\begin{aligned} \mathcal{S}_1 = \{ & (\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ & (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\}), \\ & (\{1, 3, 4, 5, 6, 7\}, \{0, 2, 8\}) \}. \end{aligned}$$

Now the three splits that are compatible with all five splits of \mathcal{S}_1 are

$$\begin{aligned} s_1^c &= (\{1, 2, 3, 4, 5, 6, 7\}, \{0, 8\}), \\ s_2^c &= (\{1, 3, 4, 5, 6, 7, 8\}, \{0, 2\}), \\ s_3^c &= (\{2, 8\}, \{0, 1, 3, 4, 5, 6, 7\}). \end{aligned}$$

With this we define three resolved tree topologies

$$\mathcal{R}_i = \mathcal{S}_1 \cup \{s_i^c\} \quad \text{for } i = 1, 2, 3.$$

We apply the BPA for one tree for each of these topologies and we use the weights

$$w_{eq} = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right)$$

for all three topologies. So our starting trees are given by $Y_i^s = (\mathcal{R}_i, w_{eq})$ for $i = 1, 2, 3$. The BPA is, just as the PPA, an iterative algorithm which converges to a balance point. In most cases it does not really reach it. Hence, we also give the BPA a maximal amount of iterations to control the computation time, just as for the PPA. One iteration consists of calculating one Euclidean median problem for the respective representatives of the current iterate. The results for our starting points in the three different orthants with a maximal number of iterations of 1000 are given in Table 6.7.

starting point	time in s	objective value
Y_s^1	9111	72.202
Y_s^2	8252	72.189
Y_s^3	8123	72.189

Table 6.7: Results of the BPA for the Apicomplexa data set with a cap on the number of iterations of 1000.

In the second experiment we choose only one orthant to see how sensitive the BPA is to different starting points. We choose the resolving tree topology \mathcal{R}_3 , as it yielded the best result in our first experiment. We compare the outcomes for the following five weight vectors:

$$\begin{aligned} w_1^s &= (0.3, 0.12, 0.1, 0.01, 0.01, 0.01) \\ w_2^s &= (0.1, 0.1, 0.1, 0.1, 0.1, 0.1) \\ w_3^s &= (0.1, 0.1, 0.1, 0.1, 0.1, 0.5) \\ w_4^s &= (1, 1, 1, 1, 1, 1) \\ w_5^s &= (100, 100, 100, 100, 100, 100) \end{aligned}$$

We chose the above weight vectors for the following reasons. w_1^s is a weight vector that is close to weights that PPA yielded as output, so it might be a good starting point. Another interesting investigation is to find the differences in the outcome of w_2^s and w_3^s : They are identical except for the weight for the last split, which is exactly the one that is not contained in \mathcal{S}_1 . We expect this weight to tend to zero for all weight vectors, so we investigate here, how fast this happens and what the difference is when choosing a bigger value for this split. Last but not least, we chose

two weight vectors w_4^s and w_5^s that has slightly bigger or exceptionally big weights to see how the BPA copes with starting points that are 'outliers', i.e., far away from any tree in the data set and from a median.

Table 6.8 summarizes the outcomes for the starting points $Y_i = (\mathcal{R}_3, w_i^s)$ for $i = 1, \dots, 5$. We can see that even though the optimal solution is not contained in the

starting point	time in s	objective value
Y_1	833	72.348
Y_2	851	72.218
Y_3	871	72.191
Y_4	829	72.199
Y_5	834	72.212

Table 6.8: Results of the BPA for the Apicomplexa data set with a cap on the number of iterations of 100.

interior of the orthant and we have no optimality analysis, it does not really matter too much with what weight one starts the algorithm in a given orthant. There are notable differences in the objective values, but when looking at the weights of the output trees we see that the resulting weight vectors are quite similar:

$$\begin{aligned}
 w_1^o &= (0.305, 0.169, 0.108, 7 \cdot 10^{-7}, 8 \cdot 10^{-7})^t \\
 w_2^o &= (0.299, 0.184, 0.111, 1 \cdot 10^{-6}, 1 \cdot 10^{-6})^t \\
 w_3^o &= (0.304, 0.186, 0.112, 1 \cdot 10^{-6}, 1 \cdot 10^{-6})^t \\
 w_4^o &= (0.312, 0.192, 0.117, 8 \cdot 10^{-7}, 8 \cdot 10^{-7})^t \\
 w_5^o &= (0.314, 0.193, 0.117, 7 \cdot 10^{-7}, 7 \cdot 10^{-7})^t
 \end{aligned}$$

For all starting weights the algorithm converged to the same boundary, which is the unresolved tree topology \mathcal{S}_1 , i.e., the split s_3^c gets thrown out. This also very well suits the conjecture that this tree topology contains a median, since the BPA always decreases the objective and also converges to this tree topology in our experiments so far.

That the BPA seems to be rather independent of the weight with which it is started is a helpful observation: It is no problem to start the BPA close to the boundary of the orthant (Y_1) or with weights that are far higher than the ones of the given sample trees (Y_5). Surprisingly, it seems that the BPA performs better, when it is not immediately started next to the boundary to which it converges. This is probably because once one gets close to this boundary, the Euclidean median problem has a lot of representative facilities in this boundary that are close to each other, so that the steps of the BPA are not that big early on and continue to get smaller.

Altogether the experiments suggest that starting points do not influence the BPA too much. The experiment shows that we get a reasonable result when simply taking a starting weight vector which is not too close to a boundary already. Moreover, it is nice to see that the method is robust to a seemingly bad choice of starting weights (Y_5), as the PPA does not perform too well in such a case.

As mentioned, the last experiment aims for improving the PPA solutions by applying the BPA on the output trees. The trees that we use for this are the output trees from Table 6.5. All outputs of the PPA from Table 6.5 have the Tree Topology 1 from Figure 6.1, whose splits are given by \mathcal{S}_1 . The output trees from Table 6.5 have the following weight vectors, which we now use as input weights

$$\begin{aligned} w_1^P &= (0.30847421, 0.18800566, 0.1120948, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})^t \\ w_2^P &= (0.30862393, 0.18805914, 0.11213941, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})^t \\ w_3^P &= (0.30863274, 0.18805648, 0.11216363, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})^t \\ w_4^P &= (0.30866245, 0.18809296, 0.11217955, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})^t \\ w_5^P &= (0.30823681, 0.18787339, 0.11199423, 1.7 \cdot 10^{-7}, 2.3 \cdot 10^{-7})^t \end{aligned}$$

Note that the weight vectors are really similar, which is due to the fact that the PPA eventually converges to a median. The similarities indicate that we are not too far away from a median.

Nevertheless, we now use these output trees from the PPA as starting points for the BPA in order to improve the solutions and look for changes. Since the BPA is guaranteed not to increase the objective we may try to improve any solution from the PPA like this, however, it might be that the BPA immediately returns the starting point, when a scenario as in Example 5.4.2 occurs.

The hope of is that we get a bigger decreases in the objective or a faster convergence to a median than by simply letting the PPA execute more cycles, as Table 6.2 suggests that the convergence gets really slow at some point.

Now, in order to finally present the results of the experiment, denote the start trees as $X_i^P = (\mathcal{S}_1, w_i^P)$ for $i = 1, \dots, 5$. Table 6.9 shows the results of the BPA.

Due to the nature of the BPA, the resulting trees have to have a split set that is a subset of the starting tree, i.e., a subset of the splits of the orthant in which it is started. For all five given trees, the length of the split $\bar{s} = (\{1, 3, 4, 5, 6, 7\}, \{0, 2, 8\})$ reached 0 at some point and it remained 0. So the resulting topology of the five output trees is

$$\mathcal{S}^r = \{(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), (\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), (\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}), (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\})\}.$$

starting point	time in s	objective value
X_1^P	452	72.180624
X_2^P	448	72.180628
X_3^P	461	72.180629
X_4^P	454	72.180634
X_5^P	452	72.180665

Table 6.9: Results of the BPA for the Apicomplexa data set. Improving the PPA solutions from Table 6.5 by applying BPA to the output trees.

The respective weight vectors of the output trees are

$$\begin{aligned}
w'_1 &= (0.30847421, 0.18800566, 0.11209484, 6.5 \cdot 10^{-10})^t, \\
w'_2 &= (0.30862391, 0.18805914, 0.11213941, 6.5 \cdot 10^{-10})^t, \\
w'_3 &= (0.30863273, 0.18805648, 0.11216362, 6.5 \cdot 10^{-10})^t, \\
w'_4 &= (0.30866244, 0.18809295, 0.11217954, 6.4 \cdot 10^{-10})^t, \\
w'_5 &= (0.30823681, 0.18787340, 0.11199424, 6.5 \cdot 10^{-10})^t.
\end{aligned}$$

Even though all output trees decrease the objective of the respective starting trees, the improvements are only marginal. That was to be expected, since it is quite likely that all starting points are already pretty close to an optimal solution as the PPA has already been performed with 10000 cycles with a good step length and reasonable starting trees. Nevertheless, the BPA actually decreased the objective for all five starting trees, even though it was applied in a lower-dimensional orthant where no guarantee of convergence to an optimal solution exists.

The more important observation is that the BPA strongly suggests that the optimal solution is to be found in an even lower-dimensional orthant: It removed the split $\bar{s} = (\{1, 3, 4, 5, 6, 7\}, \{0, 2, 8\})$ and it seems as though the length of the split $(\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\})$ also tends to 0. The weight vectors the PPA determined, i.e., the weight vectors of the starting trees already suggest, that these weights might converge to 0, but after 10000 cycles they still have weights of about 10^{-7} . In contrast to that the BPA has actually removed a split and the other weight seems to converge to 0 faster as for the PPA. In fact, the objective seems to further decrease when removing the split $\tilde{s} = (\{1, 4, 5, 6\}, \{0, 2, 3, 7, 8\})$. We take the best solution of the five BPA output trees, which is $\bar{X} = (S^r, w'_1)$, remove the split \tilde{s} and keep the other splits with the same weights as in w'_1 . The resulting tree is $X' = (S^r \setminus (\tilde{s}), (0.30847421, 0.18800566, 0.11209484))$. Then we actually have

$$f(\bar{X}) = 72.1806242978 > 72.1806242901 = f(X').$$

These are extremely small improvements, which would usually be practically irrelevant when only searching for a minimum. But in this case the structure of the minimizer changes: The comparison suggests that the optimal solution is contained in the orthant with only the splits

$$\begin{aligned} &(\{4, 5\}, \{0, 1, 2, 3, 6, 7, 8\}), \\ &(\{1, 6\}, \{0, 2, 3, 4, 5, 7, 8\}), \\ &(\{3, 7\}, \{0, 1, 2, 4, 5, 6, 8\}). \end{aligned}$$

The BPA has reached values of about $6.5 \cdot 10^{-10}$ for the length of \tilde{s} and stopped since the employed exactness threshold of 10^{-15} has been reached, i.e., two successive iterates of the BPA had a distance of less than 10^{-15} . Even though the BPA has not deleted the split \tilde{s} it still shows the right tendency to converge to solutions in the lower-dimensional orthants and it has done so faster than the PPA.

To wrap up the results of the section we want to stress that the main insight is that the BPA handles the real data set better than expected. Not only did it give reasonably good results, when started in maximal orthants, but it also showed a nice behavior when applied to non-maximal orthants, which allowed for a better interpretation of the results than with the PPA outputs. Also the computation times for the BPA are still manageable here. As a final note we mention that quite some computational effort of the BPA is put into the optimality test for the Euclidean median problem as well as into the Weiszfeld-iterations. It is hence an interesting future line of research to practically explore how to improve the BPA algorithmically and to see which algorithms for the Euclidean median problem may outperform the Weiszfeld algorithm to be able to gain a significant speed up for the BPA.

6.3 Summary of Results and Comparison to other Algorithms

In this section we extract the essence of the BPA and PPA experiments on the real data set. First it needs to be said that the BPA strongly relies on a good starting solution that is produced by the PPA. When the BPA is started in some random orthant, there would not be a lot of reasonable results, so the BPA is dependent on the PPA. Nevertheless, the BPA has proven to be a very useful tool in combination with the PPA, even if the optimality analysis of balance points may not be applied in the lower-dimensional orthants. This is for two reasons. Firstly the BPA guarantees not to increase the objective, so we cannot ‘lose’ anything by applying it to a solution of the PPA. Practice has shown that we even get a decrease in the objective. The

second reason is that the BPA provides a better insight as to where the median is probably located: in an even lower-dimensional orthant. The PPA was not able to converge to the boundary of the orthant with Tree Topology 1 as quickly as the BPA, see Table 6.9. This convergence to the boundary enables a better interpretation of the results and it is quite likely that a median is contained in that lower-dimensional orthant the BPA hints at, which is important to know.

In contrast to many other algorithms in tree space that use the local Euclidean structure to determine the Fréchet mean, the BPA may also be applied at the boundary of the orthants, i.e., in lower-dimensional orthants. Most of the other algorithms make use of the differentiability and other features that do not hold at the boundary of the orthants. This is also an issue for our optimality analysis. Nevertheless, the algorithmic scheme of the BPA may be applied to lower-dimensional orthants and we can hence search for better solutions in lower-dimensional orthants that seem to contain an optimal solution. This works as the iterative scheme does not need differentiation and only requires the calculation of the geodesics and Euclidean medians. That is a big upside in comparison to other algorithms, for which the iterative scheme may not be applied to lower-dimensional orthants, as e.g., the descent methods in [MOP15].

Besides the comparison of BPA and PPA we also want to interpret our results and compare them with other insights from the literature. Therefore we first need to settle on what the actual tree topology of a median looks like. Since the lowest objective value we were able to find was the tree

$$X' = (\mathcal{S}^r \setminus (\tilde{s}), (0.30847421, 0.18800566, 0.11209484))$$

with only three splits, and the PPA and BPA have very low weights on the splits that are not in $\mathcal{S}^r \setminus (\tilde{s})$ we conjecture that this topology is the most likely to be the topology of a median.

One of the probably easiest obtainable hypotheses for the species tree are consensus trees. [KWK08] were the first to work with this Apicomplexa data set and have used several methods to obtain the set of gene trees. With the set of gene trees at hand they calculate the extended majority rule consensus tree with the method *consensus* of the widely used PHYLIP package. The extended majority rule consensus calculates a consensus tree using the following criteria: Any split that appears in more than 50% of the trees is included. The program then considers the other splits in order of their frequency with which they have appeared, adding to the consensus tree any of them which are compatible with it until the tree is fully resolved. The resulting tree topology is on the left hand side in Figure 6.2, on the right hand side is the presumed tree topology of a median.

One immediately notices that the *Median Tree Topology* contains less splits than the *Consensus Tree Topology*. As we know, split sets uniquely describe the tree

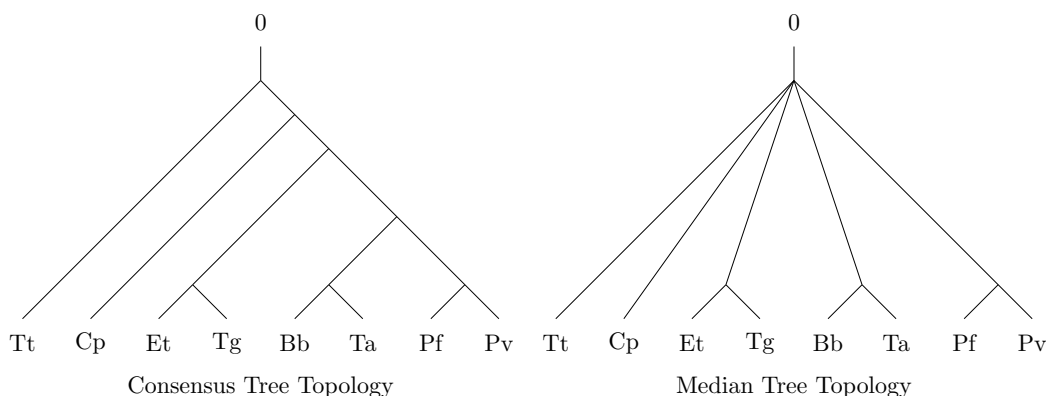


Figure 6.2: The tree topology of the extended majority rule consensus tree (left) and the topology of the approximate solution to the median problem determined by PPA and BPA (right).

topologies, and we have that the consensus tree contains three additional splits, that the Median Tree Topology does not contain. These splits are

$$(\{1, 4, 5, 6\}|\{0, 2, 3, 7, 8\}), \quad (\{1, 3, 4, 5, 6, 7\}|\{0, 2, 8\}), \quad (\{1, 2, 3, 4, 5, 6, 7\}|\{0, 8\}).$$

So, three of the six splits of the Consensus Tree Topology are also contained in the Median Tree Topology and the Median Tree Topology does not contain a split that is not contained in the Consensus Tree Topology. This hints that the Median Tree Topology seems to extract only the splits with the highest consent from a data set, even though it may not be too many.

We do not try to give a practical biological interpretation of the difference between the two trees, as this is best to be left to experts, though we note that from a biological perspective it is not the most desirable result to have topologies that have few splits. But on the other hand, it does not really seem reasonable to enforce resolvedness of trees, when the model does not allow for it. This is also why it is difficult to fairly compare the majority rule consensus tree to median trees or the Fréchet mean tree, as the median and the mean tree might very likely be unresolved as there is no artificially added constraint in the method forcing them to be unresolved.

Moreover we emphasize that it is plausible that the split $(\{1, 2, 3, 4, 5, 6, 7\}|\{0, 8\})$ that distinguishes the species *Cryptosporidium parvum* ('Cp', label 2) from the origin is not contained in the Median Tree Topology and should also not be contained in the species tree topology; our conjecture is that this split is only included in the Consensus Tree Topology as the extended majority rule enforces resolved tree topologies. We noted in Section 6.1, that we removed splits of length 0 from the trees in the data set, as these do not carry any information in the tree space and may

be omitted. After that transformation many trees did not distinguish the species *Cryptosporidium parvum* ('Cp') from the origin anymore and it is hence not very surprising that the Median Tree Topology does not depict this distinction as well. From our perspective, that also arises the question how sensible it is to calculate consensus trees on a set of trees where edges of length 0 are present, since this may heavily influence the result. As edge lengths are not taken into account in the method, these 0 weight edges have a big influence on the outcome of the consensus tree method, whereas they should actually not be considered at all.

Another paper that dealt with the Apicomplexa data set is [YZZ17], where they investigated the data set in the space of ultrametric trees. The space of ultrametric trees is another metric space in which phylogenetic tree problems may be modeled, as already mentioned in Section 2.3. [YZZ17] calculated two-dimensional *Stiefel tropical linear spaces* which are the analogon of principal components in this space using a randomized search heuristic. Since they calculated these special principal components which relate more to variance minimization and hence the Fréchet mean than the median we may not really compare the results. Nevertheless, after applying their PCA method and projecting the tree topologies they conclude that their topologies are 'largely congruent with the accepted phylogeny' as [YZZ17] describe it. The 'accepted phylogeny' groups 'Pf' and 'Pv' together, as well as the four species 'Bb', 'Ta', 'Et', 'Tg', whereas 'Tt' is an isolated species. The Median Tree Topology does not meet all of these three requirements as it does not group 'Bb', 'Ta', 'Et', 'Tg' together. Nevertheless, it is at least possible to resolve the tree to meet the requirements of the accepted phylogeny, as the split that groups 'Bb', 'Ta', 'Et', 'Tg', together is compatible with the Median Tree Topology. This is not the case for the Consensus Tree Topology, where the split 'Bb', 'Ta', 'Pf', 'Pv' is actually conflicting with 'Bb', 'Ta', 'Et', 'Tg'.

Altogether, the conducted experiments on the real data set have shown promising results. We see that the median may be a helpful concept in the following sense: Even if it may yield unresolved trees, the resulting tree seems to incorporate the most important shared features of the sample trees. The three splits that have been included in nearly all sample trees have also been contained in our approximate solutions of a median. In this way the median seems to yield similar tree topologies as majority rule consensus trees, for which it is not forced that the consensus tree is unresolved. From that point of view the median is a concept that is close to majority consensus trees but additionally includes edge weights in its analysis.

7 Conclusion

The thesis consists of three major contributions. The first and probably most important contribution is to merge two different fields of research by introducing optimization methods, Facility Location methods in particular, to the phylogenetic tree related research. Facility Location is a research field which is often restricted to business applications and therefore most of the research restricts to the Euclidean plane. With our application of Facility Location in tree space we have found a non-business application for the field. Moreover we motivate future research by showing that Facility Location can be applied to metric spaces such as the tree space thereby yielding promising results.

The second contribution is based on connecting the tree space problems and known location problems. We point out a connection of special cases of tree space location problems to Euclidean *fixed gate point location problems*. Since we are able to find solutions for these fixed gate point location problems in many cases this also implies that we can solve the special cases of the location problems in tree space from which the problems originated from.

The third and theoretically biggest contribution is the development and in-depth analysis of the BPA. The BPA is designed in such a way that it exploits the local Euclidean structure which enables the use of known methods of Facility Location. The idea of improving a given solution by local Euclidean methods seems to only provide a rough heuristic at first glance. Nevertheless, when experimenting with the method, see Section 5.2, we realized that the GBPH, that is based on the BPA yields surprisingly good solutions in average. That motivated the convergence analysis of the method, resulting in the proof of convergence in Section 5.3. The clear drawback of the BPA is that there is no guarantee for it to converge when the set of medians is contained in a lower-dimensional orthant and practice shows that it is often the case that medians are contained in lower-dimensional orthants. Nonetheless, the experiments in \mathcal{T}_4 and \mathcal{T}_5 as well as the real data example have shown that the algorithm is still very useful. First of all, it yields valuable information about the orthants we apply it in; when the BPA converges to the boundary, it is an indicator that the optimal solution is probably contained in the lower-dimensional orthant the BPA converges to. This way it is possible to identify lower-dimensional orthants that are likely to contain an optimal solution and the BPA may be applied in these lower-dimensional orthants as well. Even though there is no guarantee for convergence

to an optimal solution in lower-dimensional orthants, the experiments in Chapter 6 show that the BPA works well in lower-dimensional orthants for the data set at hand and was able to improve any given solution.

In the introduction we motivated the research undertaken in this thesis by the species tree problem. We pointed out that there are various methods and different approaches on how to produce a guess or a hypothesis for the species tree when given a set of sampled gene trees. The thesis focuses on calculating the median of such a set of trees. Therewith we add another approach to the list of calculating hypotheses for the species tree, but since none of the approaches so far is generally accepted by biologists to always infer the correct species tree, it does no harm to offer a new candidate. In Chapter 6 we applied the BPA and received a tree that is very close to the median of the data set and we have seen that the result is not far off the species tree that is agreed upon, or at least the most likely species tree according to biologists. This may both be seen as an advantage as well as a disadvantage. It is certainly a good validation of the concept and the model assumptions when the median yields reasonable or expectable results. On the other hand, when the results are too close to known or expected results, then our method does not offer a real new hypothesis.

In general, it is hard to evaluate methods for species tree problems as there is little to no possibility to actually prove or validate results for these problems due to the reverse engineering nature of the problem. Moreover, we do not aim for overinterpreting the results here, as interpretation of phylogenetic trees should be left to experts.

In conclusion, the thesis has pointed out the potential of introducing methods from Facility Location to tree space. On one hand, the achieved results show that we have a good solution approach for the median problem, but that there is further research left to explore how to precisely solve cases with optimal solutions in the boundary. Theoretically one might, for example, try to find optimality conditions for boundary cases using directional derivatives, similar to the conditions that [MOP15] develop for the mean problem. Such conditions could be used to improve the BPA by calculating directional derivatives at balance points that are located in boundaries. Then the directional derivatives might be used to tackle two central problems: Firstly, they naturally yield an optimality criterion for points on the boundary, so it is possible to evaluate the solutions found by the BPA and secondly the directional derivatives may hint at a specific lower-dimensional orthant in which one could continue the search for the optimal tree.

Appendix

CAT(0) spaces

We have referred to this section for a brief and concise introduction of geodesic metric and CAT(0) spaces. We start with the definition of geodesic metric spaces.

Definition 7.1 ([Bac14b]). *We say that a metric space (X, d) is a geodesic space if for every two points $x, y \in X$ there exists a path $\gamma : [0, 1] \rightarrow X$ such that*

$$d(\gamma(s), \gamma(t)) = |s - t|d(\gamma(0), \gamma(1)),$$

where $\gamma(0) = x$, $\gamma(1) = y$. The path γ is called geodesic from x to y .

As uniqueness of geodesics has a lot of important implications we define *uniquely geodesic spaces*:

Definition 7.2 ([Bac14b]). *A geodesic metric space (X, d) is uniquely geodesic if for every two points $x, y \in X$ there exists a unique geodesic from x to y .*

The concept of curvature may be introduced in a much more general than in our summary here. We restrict to only defining the required CAT(0) instead of general CAT(k) spaces in order to avoid unnecessary technicalities. Curvature is measured by comparing distances of particularly parametrized points on opposing sides of triangles in the metric space with the distance between particularly parametrized points on opposing sides of a triangle in Euclidean space:

Definition 7.3 ([Bac14b]). *Consider a geodesic metric space (X, d) and let $\gamma_1, \gamma_2, \gamma_3 : [0, 1] \rightarrow X$ be geodesics. $\gamma_1, \gamma_2, \gamma_3$ form a geodesic triangle with vertices x, y, z if*

$$\begin{aligned}\gamma_1 &: [0, 1] \rightarrow X, \gamma_1(0) = x, \gamma_1(1) = y, \\ \gamma_2 &: [0, 1] \rightarrow X, \gamma_2(0) = x, \gamma_2(1) = z, \\ \gamma_3 &: [0, 1] \rightarrow X, \gamma_3(0) = y, \gamma_3(1) = z.\end{aligned}$$

A comparison triangle $\Delta(x', y', z')$ for the geodesic triangle with vertices x, y, z is given by $x', y', z' \in \mathbb{R}^2$ such that

$$\|x' - y'\|_2 = d(x, y) \quad \|y' - z'\|_2 = d(y, z) \quad \|z' - x'\|_2 = d(z, x).$$

With this we may define CAT(0) spaces.

Definition 7.4 ([Bac14b]). *Let (X, d) a geodesic metric space. (X, d) is CAT(0) if for a geodesic triangle with respect to geodesics $\gamma_1, \gamma_2, \gamma_3$ with vertices x, y, z and a comparison triangle $\Delta(x', y', z')$ it holds that:*

$$d(\gamma_1(\lambda), \gamma_2(\mu)) \leq \|(1 - \lambda) \cdot x' + \lambda \cdot y' - ((1 - \mu) \cdot x' + \mu \cdot z')\|_2,$$

for $\lambda, \mu \in [0, 1]$.

A key feature of CAT(0) spaces is the following:

Theorem 7.5 ([Bac14b]). *Let (X, d) be a CAT(0) space. Then (X, d) is uniquely geodesic.*

We have mentioned a few times that the tree space \mathcal{T}_n is a *Hadamard space*.

Definition 7.6 ([Bac14b]). *A Hadamard space is a complete CAT(0) space.*

Since Theorem 2.16 states that \mathcal{T}_n is a CAT(0) space, it is clear that \mathcal{T}_n is also a Hadamard space, since the Euclidean orthants of which it consists are complete.

Apicomplexa Data Set

The data set contains 252 trees that are given in the widely used Newick format. The names and abbreviations of the investigated species may be found in the table below (compare Section 6.1). The abbreviations are used in the Newick-trees on the following pages.

Abbreviation	Name
Bb	Babesia bovis
Cp	Cryptosporidium parvum
Et	Eimeria tenella
Pf	Plasmodium falciparum
Pv	Plasmodium vivax
Ta	Theileria annulata
Tg	Toxoplasma gondii
Tt	Tetrahymena thermophila

In order to depict the trees in a single line per tree, we use the landscape format.

(((Pv:0.10887,Pf:2e-06):0.450566,(((Tg:0.140425,Et:0.237933):0.164784,Cp:0.357374):2e-06,(Ta:0.414277,Bb:0.075828):0.271068):0.01032):0.Tt:0.862226);
 (((Pv:0.056005,Pf:0.160324):0.14564,(((Tg:0.15779,Et:0.190757):0.244146,Cp:0.446469):0.081707,(Ta:0.210332,Bb:0.154176):0.248695):0.240365):0.Tt:1.801646);
 (((Pv:0.026935,Pf:0.000513):0.218681,(Ta:0.106259,Bb:0.132035):0.084837):0.024391,(Cp:0.219855,(Tg:0.094113,Et:0.296386):0.042994):0.023744):0.Tt:0.445389);
 (((((Ta:0.075982,Bb:0.093578):0.075704,(Tg:0.138806,Et:0.275626):0.13922):0.0304,(Pv:0.026348,Pf:0.029386):0.213323):0.082747,Cp:0.178225):0.Tt:0.375947);
 (((Pv:0.120634,Pf:0.151612):0.404303,Cp:0.708297):0.086897,(((Tg:0.080211,Et:0.10618):0.173588,(Ta:0.384918,Bb:0.245501):0.178588):0.035334):0.Tt:0.71643);
 (((Tg:0.126597,Et:0.253507):0.114332,((Ta:0.253121,Bb:0.191217):0.239492,(Pv:0.034181,Pf:0.025932):0.214165):0.083201):0.061012,Cp:0.220292):0.Tt:0.854742);
 (((Pv:0.027887,Pf:0.047181):0.230827,((Ta:0.30681,Bb:0.203542):0.389604,(Tg:0.093847,Et:0.242006):0.040231):0.044772):0.194744,Cp:0.23426):0.Tt:2.517163);
 (((Ta:0.229544,Bb:0.178428):0.23773,(Pv:0.035869,Pf:0.051972):0.427082):0.10792,(Tg:0.1254,Et:0.275284):0.074412,Cp:0.372214):0.045042):0.Tt:0.618957);
 (((Ta:0.289736,Bb:0.214659):0.338835,(Pv:0.091951,Pf:0.162502):0.641198,(Tg:0.235021,Et:0.306122):0.272838):0.081131):0.190672,Cp:0.556049):0.Tt:0.516209);
 (((Tg:0.062333,Et:0.134755):0.061877,((Pv:0.035521,Pf:0.032494):0.274249,(Ta:0.126251,Bb:0.129893):0.149285):0.050452):0.058634,Cp:0.192667):0.Tt:0.952282);
 (((Pv:0.143379,Pf:0.062074):0.397007,(Ta:0.354146,Bb:0.323986):0.18038):0.199062,Cp:1.141397):0.055536,(Tg:0.266961,Et:0.488291):0.225526):0.Tt:1.293297);
 (((Ta:0.130224,Bb:0.058549):0.205358,((Cp:0.256853,(Tg:0.072916,Et:0.141928):0.118337):0.093659,(Pv:0.020733,Pf:1e-06):0.271465):0.185451):0.Tt:0.487914);
 (((Pv:0.071919,Pf:0.017275):0.349886,Et:0.253803):2e-06,Tg:1.803907):0.184068,(Ta:0.228266,Bb:0.209849):0.296846):0.208639,Cp:0.141124):0.Tt:2.310614);
 (((Pv:0.024354,Pf:0.029489):0.253206,(Tg:0.058873,Et:0.679713):0.028338,((Ta:0.078708,Bb:0.0642):0.047523,Cp:0.504046):0.057038):0.076151):0.Tt:0.146047);
 (((Bb:0.262953,Ta:0.531894):0.114717,(Pv:0.103989,Pf:0.199476):0.387423,(Et:0.443798,Cp:0.498809):0.041155):0.077986):0.114014,Tg:0.204379):0.Tt:0.537288);
 (((Pv:0.057689,Pf:0.024604):0.191867,(((Tg:0.089389,Et:0.115823):0.020286,Cp:0.214586):0.062534,(Ta:0.201277,Bb:0.085637):0.183523):0.07731):0.Tt:0.545186);
 (((Pv:0.029941,Pf:0.060283):0.323463,(Ta:0.157127,Bb:0.199001):0.160836):0.052287,(Tg:0.260188,Et:0.557902):0.113638):0.10113,Cp:0.301123):0.Tt:1.146346);
 (((Ta:0.316675,Bb:0.613288):0.187568,(Pv:0.075637,Pf:0.113557):0.446743):0.110804,Cp:0.362001):0.156538,Tg:0.111395):0.055157,Et:0.286228):0.Tt:0.850859);
 (((Pv:0.102661,Pf:0.050225):0.281671,(Ta:0.437517,Bb:0.19399):0.163806):0.066657,Cp:0.314472):0.023138,(Tg:0.123183,Et:0.472552):0.100983):0.Tt:0.436302);
 (((Pv:1e-06,Pf:0.07116):0.453561,(Tg:0.141626,Et:0.129416):0.117029):0.090585,Cp:0.198221):0.09351,(Ta:0.232889,Bb:0.103701):0.25547):0.Tt:0.858059);
 (((Pv:0.098269,Pf:0.015508):0.435007,(Cp:1.036682,(Tg:0.259151,Et:0.57619):0.18312):0.098869):2e-06,(Bb:0.65267,Ta:2.468047):0.425855):0.Tt:0.728167);
 (((Pv:0.010015,Pf:1e-06):0.166664,(Ta:0.009351,Bb:0.034284):0.123746,((Tg:0.028078,Et:0.071772):0.02386,Cp:0.066315):0.025669):0.054454):0.Tt:0.455164);
 (((Ta:0.22651,Bb:0.297486):0.302993,(Tg:0.117459,Et:0.13912):0.172189):0.124553,(Pv:0.05034,Pf:0.094004):0.450814):0.145739,Cp:0.571306):0.Tt:2.037631);
 (((Pv:0.005716,Pf:1e-06):0.196719,Cp:0.206916):0.060486,Tg:0.038558):0.045761,(Ta:0.14549,Bb:0.074017):0.199205):0.039022,Et:0.111277):0.Tt:0.777314);
 (((Ta:0.110569,Bb:0.199276):0.442759,(((Pv:0.033473,Pf:0.010846):0.246454,Et:0.077112):0.042621,Tg:0.00446):0.228576,Cp:0.497754):0.141986):0.Tt:0.227037);
 (((Ta:0.177107,Bb:0.196665):0.120183,Et:3.672233):0.090845,(Pv:0.057793,Pf:0.078442):0.433603):0.168704,Cp:1.149202):0.082889,Tg:0.218375):0.Tt:1.978348);
 (((Tg:0.093373,Et:0.015633):0.248999,(Ta:0.286541,Bb:0.062318):0.250999,(Pv:0.007612,Pf:0.028076):0.053101):0.077317):0.151217,Cp:0.166764):0.Tt:0.8895);
 (((Pv:0.019201,Pf:0.010223):0.193053,((Ta:0.085376,Bb:0.140658):0.23953,(Tg:0.155887,Et:0.110339):0.047465):0.037368):0.033803,Cp:0.23047):0.Tt:0.428727);
 (((Pv:0.103724,Pf:0.095121):0.335659,(Ta:0.314546,Bb:0.444168):0.342029):0.214853,(Et:0.772831,Tg:0.200034):0.15912):0.089282,Cp:0.642557):0.Tt:1.071328);

(((Pv:0.084271,Pf:0.132715):0.424653,((Tg:0.201948,Et:0.618559):0.276919,Cp:0.272491):0.031741,(Ta:0.122816,Bb:0.250528):0.214013):0.044442):0,Tt:0.427801);
 (((Pv:0.033887,Pf:0.081555):0.339449,Cp:0.887891):0.070236,((Tg:0.225443,Et:0.351029):0.351336,(Ta:0.416027,Bb:0.440432):0.182826):0.041576):0,Tt:0.558724);
 (((Cp:0.24004,(Pv:0.020754,Pf:1e-06):0.39512,(Tg:0.131956,Et:0.29794):0.120723):0.083219):0.208648,(Ta:0.359369,Bb:0.217365):0.341256):0,Tt:1.178107);
 (((Tg:0.209917,Et:0.190213):0.111198,(Pv:0.055513,Pf:0.026169):0.403448):0.01912,(Ta:0.101417,Bb:0.144886):0.070988):7e-06,(Tt:100,Cp:0.242048):0);
 (((Ta:0.023164,Bb:0.189022):0.335089,Et:0.719725):0.023327,(Tg:0.13494,(Pv:0.030372,Pf:0.007223):0.172228,Cp:0.333878):0.015743):0.031997):0,Tt:0.594317);
 (((Pv:0.107806,Pf:0.013946):0.818367,(Ta:0.333552,Bb:0.224926):0.200826):0.17136,Cp:1.035538):0.070396,(Tg:0.31438,Et:0.449908):0.116709):0,Tt:1.625665);
 (((Pv:0.097107,Pf:0.189989):0.243724,Cp:0.798548):0.092971,Tg:0.739278):0.074347,(Ta:0.471097,Bb:0.20685):0.384194):0.16569,Et:0.577956):0,Tt:0.425173);
 (((Tg:0.100286,Et:0.418264):0.129748,(Cp:0.376786,(Ta:0.071883,Bb:0.109416):0.175457,(Pv:0.03405,Pf:0.018055):0.041602):0.035805):0,Tt:0.449665);
 (((Pv:0.076033,Pf:0.067538):0.422397,(Ta:0.192732,Bb:0.512075):0.241155):0.088576,Cp:0.701443):0.207338,(Tg:0.254832,Et:0.333921):0.139981):0,Tt:1.210863);
 (((Cp:0.127325,Et:0.011066):0.010119,(Ta:0.034024,Bb:0.039681):0.162299,(Pv:1e-06,Pf:1e-06):0.005524,Tg:1e-06):0,Tt:0.341747);
 (((Pv:0.030483,Pf:0.036181):0.207183,Cp:0.493508):0.079288,(Ta:0.247198,Bb:0.117784):0.061681,(Tg:0.121308,Et:1.150041):0.098238):0.078153):0,Tt:0.344205);
 (((Pv:0.039293,((Pv:1e-06,Pf:0.014713):0.17399,(Tg:0.092919,Et:0.018132):0.076829):0.021797,Cp:0.110883):0.02251):0.02515,Bb:0.017937):0,Tt:0.577954);
 (((Pv:0.075757,Pf:0.075648):0.296971,(Ta:0.228841,Bb:0.189079):0.17155):0.072976,Cp:0.267629):0.043884,(Tg:0.118194,Et:0.808587):0.095075):0,Tt:0.49839);
 (((Pv:0.040979,Pf:0.045976):0.321774,(Ta:0.348364,Cp:0.43917):0.080744,Bb:0.167176):0.101715):0.051525,(Tg:0.187644,Et:0.278719):0.155054):0,Tt:1.508091);
 (((Ta:0.122551,Bb:0.085227):0.14011,(Pv:0.037549,Pf:1e-06):0.13866):0.02209,(Tg:0.083543,Et:0.122234):0.113773):0.143664,(Tt:1.354018,Cp:0.585714):0);
 (((Pv:0.13207,Pf:0.035349):0.567678,(Ta:0.350632,Bb:0.277729):0.201051):0.117238,((Tg:0.229411,Et:0.3327):0.10255,Cp:0.481034):0.08928):0,Tt:0.833025);
 (((Pv:0.079239,Pf:0.053326):0.393798,(Tg:0.193964,Et:0.376032):0.206168):0.104214,(Ta:0.181211,Bb:0.17817):0.273573):0.249161,Cp:0.392058):0,Tt:0.749229);
 (((Cp:0.810105,((Pv:0.117861,Pf:0.199672):0.527552,(Tg:0.121492,Et:0.374794):0.303068):0.122503):0.136855,(Ta:0.513278,Bb:0.401286):0.597848):0,Tt:1.298045);
 (((Pv:0.027206,Pf:0.090028):0.192216,(Ta:0.181049,Bb:0.1913):0.186787,((Tg:0.075788,Et:0.512384):0.064422,Cp:0.448129):0.072593):0.101547):0,Tt:1.268759);
 (((Pv:0.04561,Pf:1e-06):0.454618,(Ta:0.194599,Bb:0.10407):0.112895):0.108382,Cp:0.437101):0.111285,Tg:0.22805):0.098386,Et:0.102808):0,Tt:0.86979);
 (((Ta:0.082536,Bb:0.10913):0.051604,(Pv:0.006102,Pf:0.021995):0.327566,(Tg:0.128406,Et:0.334342):0.105416):0.033295):0.249778,Cp:0.533634):0,Tt:0.46505);
 (((Pv:0.026151,Pf:1e-06):0.398538,(Ta:0.309685,Bb:0.273126):0.180828):0.232203,(Tg:0.197663,Et:0.407468):0.047268,Cp:0.62794):0.097463):0,Tt:0.464264);
 (((Ta:0.334952,Bb:0.183088):0.272797,(Cp:0.163667,((Tg:0.030108,Et:0.155303):0.113057,(Pv:0.01407,Pf:0.024493):0.281365):0.024711):0.075712):0,Tt:0.3689);
 (((Tg:0.062308,Et:0.20854):0.064577,(Ta:0.257321,Bb:0.218839):0.267834,(Pv:0.003983,Pf:0.005807):0.157449):0.016606):0.054712,Cp:0.201413):0,Tt:0.465543);
 (((Pv:0.090088,Pf:0.057415):0.832388,(Ta:0.183816,Bb:0.284586):0.270421):0.158776,Cp:0.387508):0.110329,(Tg:0.156672,Et:0.332716):0.076427):0,Tt:1.079966);
 (((Ta:0.099436,Bb:0.067755):0.180592,((Et:0.243159,(Tg:0.095845,(Pv:0.01398,Pf:1e-06):0.084875):0.036423,Cp:0.183537):9e-06):0,Tt:0.618376);
 (((Pv:0.072933,Pf:0.047587):0.337611,((Tg:0.262725,Et:0.255091):0.145223,Cp:0.482801):0.039978):0.072961,(Ta:0.457483,Bb:0.493163):0.419907):0,Tt:0.365418);
 (((Pv:0.038849,Pf:0.031771):0.535343,(Ta:0.210673,Bb:0.320975):0.331565):0.372286,((Tg:0.056064,Et:1.117886):0.412233,Cp:0.529964):0.122772):0,Tt:0.423792);
 (((Pv:0.171255,Pf:0.081663):0.503822,(Ta:0.206798,Bb:0.312864):0.294529):0.103684,Cp:0.579723):0.068862,(Tg:0.33211,Et:0.228025):0.280022):0,Tt:1.341608);

(((((Pv:0.188347,Pf:0.103133):0.513962,(Tg:0.289772,Et:0.184061):0.36169):0.159874,(Ta:0.449806,Bb:0.232716):0.664215):0.248787,Cp:0.503034):0,Tt:1.175713);
 (((((Pv:0.122067,Pf:0.135724):0.516851,Cp:0.47079):0.033005,(Tg:0.17057,Et:0.42302):0.260653):0.163928,(Ta:0.409614,Bb:0.351634):0.188432):0,Tt:0.662377);
 (((Ta:0.092092,Bb:0.256378):0.292881,((Pv:0.050294,Pf:1e-06):0.154462,((Tg:0.098053,Et:0.228512):0.148922,Cp:0.292319):0.063667):0.047441):0,Tt:0.975872);
 (((((Pv:0.057083,Pf:0.044035):0.401228,(Tg:0.213936,Et:0.179567):0.097344):0.049117,(Ta:0.421116,Bb:0.206297):0.429119):0.15645,Cp:0.401132):0,Tt:0.991515);
 (((Tg:0.393437,Et:0.713591):0.30157,(Pv:0.0227,Pf:0.010397):0.354528,(Ta:0.310795,Bb:0.336109):0.40044):0.06211):0.303201,Cp:0.511456):0,Tt:0.966784);
 (((Pv:0.024388,Pf:0.040776):0.58886,((Ta:0.151785,Bb:0.246877):0.152289,((Tg:0.102579,Et:0.085293):0.094577,Cp:0.253464):0.087778):0.08895):0,Tt:0.585681);
 (((Pv:0.055914,Pf:0.017874):0.3548,Cp:0.1941):0.001953,((Tg:0.083677,Et:0.078139):0.112275,(Ta:0.068807,Bb:0.171601):0.197962):0.066889):0,Tt:0.777141);
 (((Pv:2e-06,Pf:0.045013):0.305148,(Ta:0.141382,Bb:0.164168):0.15801):0.091167,(Cp:0.267875,(Tg:0.23312,Et:0.206524):0.082025):0.070984):0,Tt:0.804872);
 (((((Pv:0.094751,Pf:0.037885):0.688008,Cp:0.26843):0.067134,(Ta:0.158495,Bb:0.215299):0.209154):0.076787,(Et:1.690312,Tg:0.327221):0.121249):0,Tt:0.916948);
 (((((Pv:0.058128,Pf:0.158343):0.412799,(Ta:0.172454,Bb:0.229555):0.260441):0.073267,(Tg:0.248918,Et:0.22193):0.22132):0.078487,Cp:0.397652):0,Tt:0.763158);
 (((Ta:0.216454,Bb:0.115798):0.120958,((Pv:0.074862,Pf:5e-06):0.336647,Et:0.54201):0.015326,Tg:0.070367):0.07221):0.116706,Cp:0.204519):0,Tt:0.461722);
 (((((Pv:0.145029,Pf:0.084892):0.463566,(Tg:0.197287,Et:0.551099):0.229538):0.121327,(Bb:0.451764,Ta:0.242818):0.324666):0.068063,Cp:0.520014):0,Tt:0.727243);
 (((((Pv:0.064687,Pf:0.059843):0.269237,(Ta:0.393662,Bb:0.251011):0.145491):0.124312,(Tg:0.099622,Et:0.151224):0.158512):0.11428,Cp:0.451495):0,Tt:0.454833);
 (((Ta:0.094779,Bb:0.088734):0.262699,((Et:0.633415,(Tg:0.081656,(Pv:1e-06,Pf:0.018957):0.385033):0.041553):0.033429,Cp:0.250753):0.17232):0,Tt:1.043903);
 (((Ta:0.504341,Bb:0.36533):0.517322,((Pv:0.08154,Pf:0.058974):0.418936,(Tg:0.122479,Et:0.296151):0.134766):0.024914):0.065675,Cp:0.463083):0,Tt:0.561176);
 (((((Pv:0.065527,Pf:0.06493):0.659522,(Tg:0.376184,(Ta:0.45315,Bb:0.301234):0.448371):0.08509):0.051465,Et:100):1.9e-05,Cp:0.377255):0,Tt:0.75075);
 (((((Pv:0.032732,Pf:0.052941):0.346249,(Tg:0.179101,Et:0.193167):0.090075,Cp:0.392333):0.125741):0.148464,Bb:0.265259):0.151894,Ta:3e-05):0,Tt:100);
 (((Tg:0.068189,Et:0.122243):0.084725,((Ta:0.013031,Bb:0.108711):0.268678,(Pv:0.034992,Pf:0.00671):0.197248):0.02497):0.086137,Cp:0.171096):0,Tt:0.842932);
 (((Ta:0.403159,Bb:0.306383):0.100922,(Pv:0.085732,Pf:0.071659):0.587923):0.093982,((Tg:0.139456,Et:0.173257):0.092505,Cp:0.270305):0.034764):0,Tt:0.799172);
 (((Et:0.123958,((Pv:0.151034,Pf:0.065188):0.603861,(Ta:0.141379,Bb:0.170974):0.207656):0.125404,Cp:0.327885):0.106881):0.034951,Tg:0.126313):0,Tt:0.521152);
 (((Ta:0.196256,Bb:0.14133):0.22654,((Pv:0.028753,Pf:0.051562):0.401636,Cp:0.361235):0.090567):0.028463,(Tg:0.209903,Et:0.288728):0.056517):0,Tt:0.661471);
 (((Cp:0.234889,(Tg:0.094116,Et:0.059617):0.025024):0.024723,(Pv:1e-06,Pf:0.022996):0.172669):0.031178,Bb:0.10565):0.038109,Ta:0.039696):0,Tt:0.449861);
 (((Et:0.619819,Tg:0.126783):0.078853,Cp:0.226477):0.039887,((Ta:0.11721.Bb:0.187008):0.157711,(Pv:0.00959,Pf:0.053893):0.20598):0.057925):0,Tt:0.656886);
 (((Pv:0.047053,Pf:0.011508):0.106611,(Ta:0.098134,Bb:0.094647):0.072071):0.016308,(Tg:0.076522,Et:0.157437):0.041045,Cp:0.211622):0.071213):0,Tt:0.925304);
 (((((Pv:0.122227,Pf:0.082386):0.416611,Cp:0.654551):0.170153,(Tg:0.228245,Et:0.321498):0.209795):0.111976,(Bb:0.530784,Ta:0.294374):0.251838):0,Tt:0.871347);
 (((Bb:0.429005,(((Pv:0.108863,Pf:0.128598):0.505157,Ta:1.482901):0.134992,Cp:0.626612):0.088222):0.097776,(Et:3.318163,Tg:0.059961):0.560329):0,Tt:0.94264);
 (((Ta:0.385046,Bb:0.316188):0.162538,((Pv:0.101224,Pf:0.01281):0.182513,((Tg:0.120818,Et:0.198557):0.178601,Cp:0.369724):0.059682):0.141753):0,Tt:0.512023);
 (((Pv:2e-06,Pf:0.189709):0.38917,(Ta:0.498595,Bb:0.142255):0.345929):0.106072,(Tg:0.05539,Et:0.337233):0.113904):0.155706,Cp:0.39151):0,Tt:0.872216);
 (((Bb:0.572474,Ta:0.832036):0.437998,((Pv:0.169958,Pf:0.061668):0.324076,(Tg:0.206792,Et:1.395496):0.348454):0.121263):0.117276,Cp:1.170549):0,Tt:0.760768);

((((((Ta:0.189788,Bb:0.28572):0.086986,(Pv:0.019658,Pf:0.03546):0.453036):0.068677,Cp:0.27206):0.027723,Et:0.199873):0.067763,Tg:0.05551):0,Tt:0.628706);
 ((((((Tg:0.152502,Et:0.194626):0.139693,((Pv:0.052492,Pf:0.098341):0.273014,Cp:0.281784):0.085978):0.100609,(Ta:0.243146,Bb:0.330589):0.085023):0,Tt:0.819596);
 ((((((Tg:0.333256,Et:0.511402):0.36574,((Pv:0.282069,Pf:0.053001):1.036802,(Bb:0.692715,Ta:0.568367):0.295342):0.103477):0.097565,Cp:0.497051):0,Tt:0.721915);
 (((Ta:0.091509,(((Tg:0.083858,Et:0.077261):0.052042,(Pv:1e-06,Pf:0.01784):0.227369):0.069255,Bb:0.045524):3e-06):0.409477,(Tt:1.454904,Cp:0.402926):0);
 ((((((Pv:0.133149,Pf:0.023902):0.487191,Ta:0.308845):0.170997,Bb:0.233326):0.054943,(Cp:0.621676,(Tg:0.129366,Et:0.20793):0.165619):0.187476):0,Tt:1.139548);
 ((((((Pv:0.01812,Pf:1e-06):0.230674,(Ta:0.138556,Bb:0.116446):0.076181):0.078623,Tg:0.053094):0.197592,Et:0.303894):0.299784,(Tt:1.109634,Cp:0.980022):0);
 ((((((Tg:0.109656,Et:0.322061):0.051633,((Pv:0.045274,Pf:0.038154):0.279851,(Ta:0.226974,Bb:0.286792):0.110957):0.113616):0.04973,Cp:0.428264):0,Tt:2.139567);
 (((Ta:0.048785,Bb:0.05956):0.076559,Cp:0.159647):0.030007,(Tg:0.033141,Et:1.23391):0.053272):0.034139,(Tt:0.41208,(Pv:0.011069,Pf:0.024892):0.154099):0);
 (((Pv:0.109628,Pf:0.037157):0.335379,(Ta:0.344664,Bb:0.208965):0.133245):0.097017,((Tg:0.261511,Et:0.341045):0.116074,Cp:0.502199):0.050517):0,Tt:1.102335);
 (((Pv:0.006988,Pf:0.043261):0.257316,((Tg:0.103511,Et:0.314694):0.101324,Cp:0.38349):0.057811):0.037905,(Ta:0.149812,Bb:0.160695):0.160063):0,Tt:0.497034);
 ((((((Pv:0.069334,Pf:0.03039):0.383255,(Ta:0.154217,Bb:0.159862):0.264056):0.149444,(Tg:0.153229,Et:0.422726):0.210688):0.16716,Cp:0.419561):0,Tt:2.32486);
 (((Tg:0.088284,Et:0.056649):0.067445,((Pv:0.033086,Pf:0.033397):0.185768,(Ta:0.076976,Bb:0.166882):0.084468):0.07799):0.080748,Cp:0.314307):0,Tt:0.786089);
 (((Pv:0.042721,Pf:0.064563):0.277658,((Tg:0.059592,Et:0.312222):0.047718,(Cp:0.160929,(Ta:0.173747,Bb:0.216856):0.233485):0.025151):0.086029):0,Tt:2.009643);
 (((Tg:0.0765,Et:0.285133):0.050282,((Ta:0.130018,Bb:0.057335):0.091237,(Pv:0.020173,Pf:0.015944):0.083635):0.033723):0.060193,Cp:0.662854):0,Tt:0.263704);
 (((Pv:1e-06,Pf:0.045453):0.335119,((Tg:0.126825,Et:0.145297):0.165753,Cp:0.18193):0.114148):0.109738,Ta:0.247694):0.018424,Bb:0.019576):0,Tt:1.149055);
 (((Pv:0.141268,Pf:0.042759):0.280117,Cp:0.595532):0.041351,(Ta:0.276892,Bb:0.356187):0.113566):0.164472,(Et:0.692898,Tg:0.110591):0.056314):0,Tt:0.675282);
 (((Pv:0.019223,Pf:0.014225):0.335235,(Ta:0.22528,Bb:0.392838):0.126374):0.086513,Cp:0.481766):0.042112,(Tg:0.040708,Et:0.995116):0.088374):0,Tt:0.572028);
 (((Pv:0.041458,Pf:0.055047):0.199083,((Ta:0.12958,Bb:0.154307):0.072015,Cp:0.486829):0.03524):0.033995,(Tg:0.051286,Et:0.084608):0.095863):0,Tt:0.367371);
 (((Pv:0.070799,Pf:0.024267):0.288361,Cp:0.180939):0.059594,((Tg:0.220177,Et:0.395368):0.075734,(Ta:0.223079,Bb:0.24371):0.088541):0.032242):0,Tt:0.180192);
 (((Pv:0.07895,Pf:0.166157):0.752931,((Tg:0.217885,Et:0.329563):0.466244,Cp:0.585927):0.088237,(Ta:0.281058,Bb:0.23747):0.298331):0.101782):0,Tt:1.1839);
 (((Pv:1e-06,(Cp:0.203453,((Ta:0.242972,Bb:0.05222):0.236072,(Tg:0.07596,Et:0.140655):0.093877):0.021933):0.160766):0.062182,Pf:2e-06):0,Tt:1.256113);
 (((Pv:0.090953,Pf:0.084462):0.462598,(Ta:0.25728,Bb:0.376757):0.244008):0.104811,((Tg:0.143457,Et:0.250855):0.084886,Cp:0.435804):0.067641):0,Tt:0.621876);
 (((Pv:0.01628,Pf:0.003619):0.163181,(Bb:0.1314,Ta:0.617406):0.092742):0.047354,(Tg:0.035791,Et:0.27456):0.053695):0.033372,Cp:0.231798):0,Tt:0.341648);
 (((Ta:0.199832,Bb:0.093068):0.166427,((Pv:0.028316,Pf:0.068045):0.301409,(Tg:0.150013,Et:0.365662):0.208132):0.082419):0.244606,Cp:0.66759):0,Tt:0.540342);
 (((Tg:0.036869,Et:0.120034):0.0435,((Pv:0.00877,Pf:0.007442):0.173756,(Ta:0.102482,Bb:0.058761):0.074327):0.021395):0.032709,Cp:0.448792):0,Tt:0.199319);
 (((Bb:0.065505,(((Tg:0.041898,Et:0.096852):0.070129,Cp:0.414201):0.112775,(Pv:0.011191,Pf:0.017638):0.109623):0.094105):0.056537,Ta:4e-06):0,Tt:0.880265);
 (((Pv:0.08379,Pf:0.0449):0.317906,(Ta:0.317743,Bb:0.411808):0.165374):0.022827,(Tg:0.293804,Et:0.157186):0.082817,Cp:0.390462):0.065939):0,Tt:0.582373);
 (((Pv:0.142562,Pf:0.133097):0.307503,(Ta:0.22695,Bb:0.379558):0.321812):0.044724,(Cp:0.359281,(Tg:0.216407,Et:0.253266):0.097339):0.103976):0,Tt:0.81064);
 (((Tg:0.015304,Et:0.036016):0.04984,(Pv:0.016248,Pf:0.010897):0.105491,(Ta:0.056082,Bb:0.062537):0.084765):0.032111):0.038446,Cp:0.075971):0,Tt:0.328806);

(((Bb:0.472656,Ta:0.097872):0.364175,(Pv:0.0653,Pf:2e-06):0.299431,(Tg:0.097669,Et:1.128665):0.405724):0.391281,Cp:0.069481):0,Tt:3.049199);
 (((Ta:0.208299,Bb:0.134582):0.091164,(Tg:0.136156,Et:0.235878):0.172881,(Pv:0.027684,Pf:0.018154):0.165398):0.151329):0.207864,Cp:0.566257):0,Tt:0.632752);
 (((Pv:0.062638,Pf:0.042576):0.234204,(Ta:0.319935,Bb:0.218876):0.333913,(Tg:0.186771,Et:0.165163):0.145084,Cp:0.475804):0.114885):0.041094):0,Tt:0.87694);
 (((((Pv:0.076749,Pf:1e-06):0.342181,Et:0.333469):0.072459,(Ta:0.091884,Bb:0.110656):0.212249):0.047684,Tg:0.236311):0.064649,Cp:0.19863):0,Tt:0.410201);
 (((((Ta:0.049125,Bb:0.130977):0.062087,(Pv:0.011647,Pf:0.013886):0.058401):0.008329,(Tg:0.025612,Et:0.067186):0.096769):0.081038,Cp:0.094615):0,Tt:0.203611);
 (((((Pv:0.030482,Pf:0.049468):0.43103,(Ta:0.158747,Bb:0.165615):0.341209):0.083094,(Et:1.527127,Tg:0.148296):0.34484,Cp:0.283285):0.050104):0,Tt:0.459656);
 (((((Pv:0.094285,Pf:0.066009):0.437339,(Ta:0.276969,Bb:0.195935):0.258556):0.116224,(Tg:0.182039,Et:0.568657):0.253824):0.117594,Cp:0.761451):0,Tt:0.762923);
 (((((Pv:0.050363,Pf:0.046094):0.269186,(Ta:0.16424,Bb:0.100419):0.069201,(Tg:0.104736,Et:0.119147):0.083752,Cp:0.231095):0.061645):0.174187):0,Tt:1.802828);
 (((((Ta:0.281507,Bb:0.252634):0.170364,(Pv:0.080775,Pf:0.060035):0.340736,(Tg:0.14356,Et:0.133447):0.230764):0.037276):0.174406,Cp:0.865659):0,Tt:1.432959);
 (((((Pv:0.041905,Pf:0.115103):0.552228,(Ta:0.421187,Bb:0.615866):0.558047):0.127516,(Tg:0.139259,Et:0.294743):0.150881):0.092453,Cp:0.315926):0,Tt:1.007648);
 (((((Pv:0.058082,Pf:0.043275):0.368346,(Tg:0.165769,Et:0.325961):0.159972):0.032847,(Ta:0.328471,Bb:0.336825):0.385733):0.048798,Cp:0.424109):0,Tt:0.676253);
 (((((Ta:0.064758,Bb:0.05265):0.030375,(Pv:1e-06,Pf:0.009267):0.08776):0.025585,(Tg:0.004224,Et:0.984038):0.088089):0.018368,Cp:0.04564):0,Tt:0.154437);
 (((((Tg:0.205581,Et:0.269539):0.293377,(Ta:0.252215,Bb:0.137342):0.163964,(Pv:0.017548,Pf:0.008173):0.167491):0.082283):0.390099,Cp:0.542046):0,Tt:1.344154);
 (((((Pv:0.063868,Pf:0.155055):0.781273,(Ta:0.430978,Bb:0.391952):0.279685):0.123004,Cp:0.422307):0.092825,(Tg:0.272577,Et:0.416029):0.252681):0,Tt:0.70208);
 (((((Pv:0.006629,Pf:0.106732):0.218339,(Tg:0.097245,Et:0.317649):0.232232):0.177084,(Ta:0.406252,Bb:0.343151):0.314531,Cp:0.157567):0.102097):0,Tt:0.379835);
 (((((Ta:0.244104,Bb:0.234462):0.262453,Cp:0.536877):0.043634,(Pv:0.027904,Pf:0.053671):0.246021):0.090783,Et:0.40995):0.098929,Tg:0.286151):0,Tt:0.603364);
 (((((Ta:0.146957,Bb:0.13041):0.046473,(Tg:0.055755,Et:0.111955):0.16158,(Pv:0.014694,Pf:1e-06):0.020524):0.059622):0.138569,Cp:0.33732):0,Tt:0.211587);
 (((((Tg:0.136537,Et:0.177795):0.285553,(Ta:0.235315,Bb:0.268459):0.27155),(Pv:0.09462,Pf:0.020101):0.610571):0.174645,Cp:0.363085):0,Tt:0.850848);
 (((((Pv:0.006472,Pf:0.034041):0.196683,(Ta:0.336929,Bb:0.234819):0.406915,(Tg:0.054333,Et:0.224953):0.264117):0.056654):0.057641,Cp:0.382604):0,Tt:0.585701);
 (((((Et:0.056994,((Ta:0.030093,Bb:0.062733):0.113175,(Pv:1e-06,Pf:1e-06):0.105347):0.075563,Tg:0.018737):0.048574):0.003386,Cp:0.114706):0,Tt:0.720698);
 (((((Ta:0.521685,Cp:0.393343):0.170642,(Pv:1e-06,Pf:0.026715):0.284928,Bb:0.381273):0.075214):0.077138,(Tg:0.165935,Et:0.149842):0.137233):0,Tt:0.258722);
 (((((Pv:1e-06,Pf:0.027603):0.152768,(Ta:0.038261,Bb:0.056834):0.045099):0.078949,Cp:0.74897,(Tg:0.040869,Et:0.039863):0.0503):0.052692):0,Tt:0.303034);
 (((((Ta:0.110992,Bb:0.179273):0.330509,(Pv:1e-06,Pf:0.012347):0.280985,((Tg:0.108209,Et:0.206652):0.230494,Cp:0.355891):0.107202):0.189591):0,Tt:0.704696);
 (((((Tg:0.032269,Et:0.277463):0.051633,(Ta:0.047975,Bb:0.070335):0.052806):0.032312,((Pv:0.046436,Pf:0.049476):0.013142,Cp:0.184068):0.033536):0,Tt:0.297923);
 (((((Pv:0.012012,Pf:0.020347):0.138045,(Tg:0.196577,Et:0.724735):0.246703):0.120049,Bb:0.143093):0.158774,(Ta:0.136511,Cp:0.801322):0.072891):0,Tt:0.666477);
 (((((Pv:0.013721,Pf:0.090572):0.122029,Cp:0.617218):0.081152,(Ta:0.138098,Bb:0.05992):0.112447):0.044299,(Et:2.069831,Tg:0.041336):0.308886):0,Tt:1.665626);
 (((((Ta:0.500304,Bb:0.642712):0.359311,Tg:0.203374):0.100115,(Pv:0.087833,Pf:1e-06):0.255219):0.148805,Et:1.849579):0.279644,Cp:0.15345):0,Tt:1.622042);
 (((((Pv:0.026892,Pf:0.033117):0.538451,(Tg:0.1463,Et:0.248166):0.116108,(Ta:0.261742,Bb:0.102226):0.136457):0.107642):0.070543,Cp:0.224109):0,Tt:0.545899);
 (((((Pv:0.095504,Pf:0.191708):0.503603,(Tg:0.128751,Et:0.223703):0.194576):0.032409,Cp:0.536605):0.094274,(Ta:0.525824,Bb:0.412927):0.148326):0,Tt:0.779163);

(((Et:0.923418,Tg:0.121727):0.170059,((Ta:0.192306,Bb:0.305267):0.119378,(Pv:0.057957,Pf:0.054691):0.145979):0.091344,Cp:0.238653):0.113553):0,Tt:0.20433);
 (((((Bb:0.281472,(Tg:0.103163,Et:0.735598):0.061204):0.049975,Ta:0.358041):0.080616,(Pv:0.03847,Pf:0.086986):0.161273):0.165872,Cp:0.507013):0,Tt:0.511048);
 (((((Tg:0.047949,Et:0.355585):0.130984,(Ta:0.165168,Bb:0.164034):0.168228,(Pv:0.017489,Pf:0.015663):0.151574):0.029329):0.053745,Cp:0.137923):0,Tt:0.813844);
 (((((Tg:0.307702,Et:0.259945):0.561317,(Ta:0.248527,Bb:0.190832):0.240303):0.095971,(Pv:0.170127,Pf:0.080202):0.53303):0.117105,Cp:0.493744):0,Tt:0.925912);
 (((((Cp:0.648714,(Ta:0.482278,Bb:0.588864):0.184936):0.13435,Et:0.418992):0.054054,(Pv:0.074676,Pf:0.128075):0.813981):0.186708,Tg:0.239786):0,Tt:0.476983);
 (((((Pv:0.002677,Pf:0.009026):0.132217,(Ta:0.138285,Bb:0.097366):0.23171,Cp:0.100831):0.045208):0.041967,Et:0.147301):0.022334,Tg:0.033773):0,Tt:0.254846);
 (((((Ta:0.245796,Bb:0.107992):0.202615,(Pv:0.004434,Pf:0.033573):0.154707):0.054568,(Tg:0.111578,Et:0.117899):0.195179):0.204803,Cp:0.277434):0,Tt:0.68931);
 (((((Pv:0.079249,Pf:0.180276):0.406631,(Tg:0.670322,(Bb:0.506524,(Et:1.701363,(Cp:0.708572,Ta:0.61364):0.006983):0.017486):0.039651):0.010873):0,Tt:0.791135);
 (((((Pv:0.011105,Pf:0.153546):0.462507,(Tg:0.250072,Et:0.391667):0.157923):0.090092,(Ta:0.344185,Bb:0.548889):0.194215):0.071597,Cp:0.623445):0,Tt:0.809404);
 (((((Tg:0.099951,Et:0.12924):0.089285,(Pv:0.09866,Pf:0.063475):0.76235):0.046576,(Ta:0.236923,Bb:0.150557):0.294103):0.057743,Cp:0.217021):0,Tt:0.669728);
 (((((Pv:0.019077,Pf:0.019179):0.447834,(Ta:0.388681,Bb:0.566172):0.519712):0.193516,Cp:0.631286):0.058286,(Tg:0.261965,Et:0.516855):0.223026):0,Tt:1.068591);
 (((((Pv:0.050255,Pf:0.034067):0.120084,Et:0.231782):0.029429,Tg:0.121589):0.026029,(Ta:0.377274,Bb:0.128092):0.219879):0.05653,Cp:0.645506):0,Tt:0.40263);
 (((((Pv:0.066767,Pf:0.039806):0.405251,(Ta:0.229571,Bb:0.186183):0.108212):0.059488,Cp:0.609926):0.115331,(Et:2.185991,Tg:0.013382):0.319851):0,Tt:0.641237);
 (((((Tg:0.318528,(Ta:0.064975,Bb:0.271356):0.427606,Et:0.606):0.100906):0.037416,(Pv:0.140555,Pf:0.031463):0.71304):0.143601,Cp:0.248217):0,Tt:0.643248);
 (((((Pv:0.082872,Pf:0.093831):0.300687,(Tg:0.145155,Et:0.291191):0.096396,(Ta:0.266435,Bb:0.433868):0.287499,Cp:0.377513):0.024345):0.126417):0,Tt:0.538444);
 (((((Pv:0.064644,Pf:0.035663):0.436173,(Ta:0.13577,Bb:0.249036):0.334538):0.114969,Cp:0.485517):0.167842,Et:0.414266):0.112764,Tg:0.110658):0,Tt:1.151416);
 (((((Pv:0.043322,Pf:0.012641):0.341827,(Ta:0.353707,Bb:0.331794):0.435548):0.057377,Cp:0.573557):0.024675,(Tg:0.003295,Et:1.539333):0.503514):0,Tt:0.80924);
 (((((Pv:0.052584,Pf:2e-06):0.178781,(Ta:0.255759,Bb:0.088245):0.153965):0.050155,Tg:0.12942):0.03893,Et:0.155452):0.077604,Cp:0.184632):0,Tt:0.601358);
 (((((Pv:0.037398,Pf:0.039446):0.470074,(Tg:0.264132,Et:0.203163):0.215046,Cp:0.433771):0.108523):0.161419,(Ta:0.328424,Bb:0.13488):3e-06):0,Tt:2.1742823);
 (((((Ta:0.248206,Bb:0.183437):0.405575,(Et:1.044253,Tg:0.076856):0.106391):0.048752,Cp:0.217647):0.045371,(Pv:0.01965,Pf:0.059066):0.227382):0,Tt:0.824072);
 (((((Pv:0.130276,Pf:0.039981):0.708905,(Ta:0.127692,Bb:0.298842):0.456597):0.191855,(Tg:0.181282,Et:0.240719):0.226374,Cp:0.403028):0.055349):0,Tt:0.36441);
 (((((Bb:0.375685,Ta:0.400344):0.368258,(Pv:0.590674,Pf:0.155432):0.177417):0.147492,(Tt:1.607121,(Cp:0.852796,(Tg:0.492636,Et:0.526651):0.429496):0.091731):0);
 (((((Pv:1e-06,Pf:0.042555):0.352833,(Ta:0.70545,Bb:0.455808):0.182796,Cp:0.610317):0.09856,(Et:0.84213,Tg:0.161478):0.160317):0,Tt:2.129905);
 (((((Pv:0.026942,Pf:0.00473):0.108802,(Ta:0.148428,Bb:0.114832):0.104581):0.0333,(Tg:0.052458,Et:0.55167):0.043822):0.029744,Cp:0.074521):0,Tt:0.371907);
 (((((Tg:0.086994,Et:0.226052):0.118598,(Ta:0.339536,Bb:0.264673):0.1213,(Pv:0.040298,Pf:0.055154):0.215383):0.035768):0.085475,Cp:0.645952):0,Tt:0.401792);
 (((((Cp:0.244638,(Pv:0.02215,Pf:0.019447):0.417868,(Ta:0.178044,Bb:0.292203):0.131685):0.056736):0.051813,(Tg:0.077498,Et:0.097625):0.077725):0,Tt:1.713278);
 (((((Pv:0.095506,Pf:0.086441):0.25891,(Tg:0.14107,Et:0.391509):0.100942):0.052016,(Ta:0.441645,Bb:0.208621):0.232025):0.083062,Cp:0.712303):0,Tt:0.85371);
 (((((Pv:0.026531,Pf:0.026535):0.378625,(Ta:0.203199,Bb:0.28239):0.130865,Cp:0.324888):0.05853):0.056656,(Tg:0.098017,Et:1.063685):0.139117):0,Tt:0.384456);
 (((((Tg:0.085157,Et:0.070765):0.075781,(Pv:0.096071,Pf:0.0434):0.249784,Cp:0.280587):0.04956):0.085959,(Ta:0.133107,Bb:0.136663):0.126322):0,Tt:0.689654);

(((Tg:0.108131,((Ta:0.193662,Bb:0.170375):0.159285,Cp:1.283718):0.097627,(Pv:0.033791,Pf:0.053801):0.20625):0.218525):0.134298,Et:0.068497):0,Tt:1.405152);
 (((Tg:0.125855,((Ta:0.056698,Bb:0.216489):0.129726,Et:0.077138):0.082405):0.075347,(Pv:0.098005,Pf:0.031063):0.173797):0.09485,Cp:0.265318):0,Tt:0.873006);
 (((Tg:0.232278,Et:0.268787):0.350357,((Pv:0.028696,Pf:0.038626):0.369709,(Bb:0.202816,Ta:0.81418):0.251649):0.131116):0.220705,Cp:0.648215):0,Tt:0.8635);
 (((Ta:0.250466,Bb:0.21238):0.10411,((Pv:0.05268,Pf:0.088186):0.142687,Cp:1.474152):0.12739):0.101473,(Tg:0.11335,Et:0.146226):0.050074):0,Tt:0.823175);
 (((Pv:0.11928,Pf:0.013989):0.243555,((Tg:0.199947,Et:0.210614):0.18668,Cp:0.538258):0.086712):0.140573,(Ta:0.393282,Bb:0.330509):0.374041):0,Tt:2.248605);
 (((Pv:0.014818,Pf:0.060512):0.747844,(Ta:0.104481,Bb:0.339871):0.723746):0.121277,(Tg:0.179348,Et:0.386733):0.124386,Cp:0.651363):0.120887):0,Tt:0.538091);
 (((Pv:0.037473,Pf:0.024668):0.289582,(Ta:0.150788,Bb:0.148397):0.085832):0.055134,Cp:0.145385):0.052983,(Tg:0.085046,Et:0.153018):0.059788):0,Tt:0.543387);
 (((Tg:0.100597,Et:0.164939):0.218479,((Pv:0.058769,Pf:1e-06):0.355088,(Ta:0.345041,Bb:0.19478):0.107529):0.13083):0.125117,Cp:0.408602):0,Tt:0.551082);
 (((Pv:0.070574,Pf:0.075907):0.316023,((Tg:0.143246,Et:0.230269):0.097976,Cp:0.844922):0.060861):0.111274,(Ta:0.619803,Bb:0.348093):0.305768):0,Tt:2.630155);
 (((Tg:0.179653,Et:0.263034):0.120021,((Pv:0.058591,Pf:0.101227):0.458108,(Ta:0.371144,Bb:0.30566):0.107646):0.088906):0.071729,Cp:0.306929):0,Tt:0.841021);
 (((Ta:0.147651,Bb:0.307464):0.195463,(Pv:0.016738,Pf:0.008586):0.117144):0.07812,(Tg:0.077156,Et:0.391139):0.191808):0.315257,Cp:0.369603):0,Tt:0.845569);
 (((Pv:0.158451,Pf:0.069462):0.737491,(Ta:0.255952,Bb:0.329139):0.336794):0.353323,(Tg:0.312924,Et:0.288956):0.124803):0.147372,Cp:0.934833):0,Tt:0.678826);
 (((Tg:0.04975,Et:0.072014):0.048708,Cp:0.125642):0.041638,(Pv:0.110758,Pf:0.032813):0.269617):0.090825,(Ta:0.282382,Bb:0.12425):0.188824):0,Tt:0.543024);
 (((Pv:0.040472,Pf:0.036257):0.403726,Cp:0.256993):0.065136,(Tg:0.136621,Et:0.239876):0.25566):0.066421,(Ta:0.150345,Bb:0.429372):0.224353):0,Tt:0.686211);
 (((Pv:1e-06,Pf:0.077558):0.747611,Cp:0.473827):0.162034,((Tg:0.150936,Et:0.257787):0.174053,(Ta:0.109869,Bb:0.154313):0.277926):0.031751):0,Tt:0.574144);
 (((Ta:0.309156,Bb:0.101327):0.182887,(Pv:0.057973,Pf:0.146349):0.41026,(Tg:0.076004,Et:0.302734):0.296915):0.071852):0.092431,Cp:1.452374):0,Tt:1.069045);
 (((Pv:1e-06,Pf:0.066798):0.311069,((Et:1.385652,Tg:0.269752):0.263089,(Ta:0.326083,Bb:0.313525):0.150413):0.079306):0.12527,Cp:0.431337):0,Tt:1.064576);
 (((Pv:0.082795,Pf:0.056911):0.370581,((Tg:0.109483,Et:0.211891):0.144736,(Ta:0.18895,Bb:0.065176):0.103701):0.061096):0.09506,Cp:0.147579):0,Tt:0.43915);
 (((Bb:0.017281,(Cp:0.224488,((Tg:1e-06,Et:0.0105):0.046633,(Pv:0.010389,Pf:1e-06):0.041294):0.032227):1e-06):0.017066,Ta:0.014124):0,Tt:0.320176);
 (((Ta:0.095194,Bb:0.064541):0.136615,(Cp:0.124793,((Tg:0.052665,Et:0.131095):0.020358,(Pv:0.004163,Pf:0.007421):0.105815):0.028311):0.030876):0,Tt:0.504225);
 (((Pv:0.051829,Pf:0.03484):0.348603,((Tg:0.149289,Et:0.228514):0.097929,Cp:0.223364):0.053993):0.034166,(Ta:0.13728,Bb:0.10834):0.084748):0,Tt:0.359297);
 (((Pv:0.092275,Pf:1e-06):0.650402,(Cp:0.661627,(Ta:0.178046,Bb:0.289289):0.14006):0.110624):0.025394,(Tg:0.119603,Et:0.357309):0.076095):0,Tt:0.488701);
 (((Pv:0.010795,Pf:0.044099):0.208233,Cp:0.238169):0.036971,Tg:0.14065):0.010297,Et:0.83288):0.105399,(Tt:0.719514,(Ta:0.070272,Bb:0.085036):0.12597):0);
 (((Tg:0.065694,Et:0.143929):0.13653,(Ta:0.216285,Bb:0.201736):0.241257,(Pv:0.019937,Pf:0.013349):0.163234):0.06968):0.049074,Cp:0.146787):0,Tt:0.903136);
 (((Pv:0.079079,Pf:0.043464):0.557873,((Tg:0.334361,Et:0.324406):0.252719,Cp:0.58599):0.03426,(Ta:0.335613,Bb:0.318726):0.237104):0.068826):0,Tt:0.729902);
 (((Pv:0.07398,Pf:0.017675):0.568989,(Ta:0.247521,Bb:0.150695):0.2008):0.198278,((Tg:0.13523,Cp:0.372661):0.031231,Et:0.25126):0.06094):0,Tt:0.491069);
 (((Pv:0.061696,Pf:0.077586):0.399651,Cp:0.274951):0.074373,(Tg:0.168738,Et:0.098412):0.120076):0.21332,(Ta:0.730702,Bb:0.699928):0.589273):0,Tt:0.727918);
 (((Pv:0.095622,Pf:0.12798):0.12798):0.668676,Tg:0.544049):0.063096,(Et:0.896171,(Cp:1.397071,(Ta:0.464104,Bb:0.405108):0.166683):0.184555):0.084598):0,Tt:0.795221);
 (((Pv:0.029525,Pf:0.025063):0.723036,(Ta:0.254021,Bb:0.424591):0.439651):0.121835,(Tg:0.183345,Et:0.481538):0.299181):0.22537,Cp:0.741589):0,Tt:2.447591);

(((Tg:0.16263,Et:0.173633):0.190291,((Ta:0.199879,Bb:0.196462):0.292257,(Pv:0.144507,Pf:0.23913):0.680253,Cp:1.408215):0.148374):0.098257):0,Tt:0.870171);
 (((((Pv:2e-06,Pf:0.074909):0.297728,Cp:0.403094):0.106332,(Tg:0.123756,Et:0.3509):7e-06,(Ta:0.673932,Bb:0.251):0.450222):0,Tt:1.300926);
 (((Tg:0.120041,Et:0.341283):0.143228,(Pv:1e-06,Pf:0.059902):0.297371,(Ta:0.06777,Bb:0.130926):0.217814):0.173557):0.166191,Cp:0.59561):0,Tt:1.222209);
 (((Tg:0.122543,Et:0.386826):0.124693,(Pv:0.077039,Pf:0.021322):0.285739,(Ta:0.239594,Bb:0.06754):0.137541):0.064383):0.137686,Cp:0.125854):0,Tt:0.768279);
 (((((Pv:0.094926,Pf:0.073582):0.342022,(Ta:0.46917,Bb:0.317795):0.19045):0.057939,(Tg:0.206129,Et:0.25223):0.131939):0.172339,Cp:0.43255):0,Tt:0.680984);
 ((((((Pv:0.016053,Pf:0.042898):0.24872,Tg:0.436908):0.050412,(Ta:0.256875,Bb:0.277968):0.429251):2e-06,Et:99.999999):0.090369,Cp:0.356268):0,Tt:1.03971);
 (((((Pv:0.031852,Pf:0.032649):0.470617,(Cp:0.604961,(Ta:0.876878,Bb:0.451809):0.645144):0.094568):0.09186,(Tg:0.138728,Et:0.188721):0.200411):0,Tt:0.723777);
 (((Tg:0.257574,Et:0.241835):0.174399,(Pv:0.006438,Pf:0.024508):0.372837,(Ta:0.112308,Bb:0.116121):0.253137):0.151784):0.342545,Cp:0.553048):0,Tt:0.504387);
 (((Cp:0.190933,(Ta:0.111124,Bb:0.30076):0.191257,(Pv:1e-06,Pf:1e-06):0.076176):0.059908):0.034482,(Tg:0.078564,Et:0.127022):0.070177):0,Tt:0.194637);
 (((Pv:0.141434,Pf:0.105649):0.321112,((Tg:0.071948,Et:0.2577):0.243183,Cp:0.55354):0.067352,(Ta:0.157641,Bb:0.171372):0.178875):0.188975):0,Tt:0.827232);
 (((Cp:0.367598,(Pv:0.022818,Pf:0.052324):0.442115,(Ta:0.310329,Bb:0.2703):0.154757):2e-06):0.068688,(Tg:0.184939,Et:0.097394):0.365677):0,Tt:0.601069);
 (((Tg:0.149812,Et:0.21152):0.10661,(Pv:0.029527,Pf:0.018209):0.20526,(Ta:0.143087,Bb:0.133829):0.080209):0.094766):0.078146,Cp:0.392808):0,Tt:0.997403);
 (((Tg:0.155885,Et:0.272917):0.099119,(Pv:0.081135,Pf:0.057799):0.341692,(Ta:0.252808,Bb:0.301099):0.173702):0.046049):0.102258,Cp:0.369952):0,Tt:0.651653);
 (((Pv:1e-06,Pf:0.012021):0.127007,(Cp:0.146235,(Tg:0.030634,(Ta:0.073141,Bb:0.082086):0.119292):0.014318,Et:0.104127):0.04725):0.106484):0,Tt:1.521213);
 (((((Tg:1e-06,Et:0.135148):0.193023,(Ta:0.195998,Bb:0.351308):0.254785):0.088981,Cp:0.207819):0.057136,(Pv:1e-06,Pf:0.035624):0.238836):0,Tt:0.370999);
 (((Pv:0.261459,Pf:0.258763):0.280038,(Cp:0.387024,(Tg:0.276185,Et:0.29027):0.320519):0.201439):0.142713,(Bb:0.306099,Ta:0.558704):0.20858):0,Tt:1.092396);
 (((((Pv:0.051411,Pf:0.061531):0.443163,(Tg:0.439959,Et:0.280309):0.10674):0.149178,Ta:0.563974):0.060177,(Cp:0.747623,Bb:0.435503):0.099686):0,Tt:0.911897);
 (((((Pv:0.069374,Pf:0.067164):0.303748,(Ta:0.070164,Bb:0.237826):0.227903):0.09636,Cp:0.101332):0.054687,(Tg:0.373046,Et:1.418243):0.523055):0,Tt:0.506022);
 (((Ta:0.127235,Bb:0.083123):0.130864,(Pv:0.076785,Pf:0.044271):0.218265):0.037921,(Tg:0.091387,Et:0.210055):0.09418):0.189662,(Tt:0.6041,Cp:0.348072):0);
 (((((Pv:0.030203,Pf:0.027843):0.726707,(Ta:0.192228,Bb:0.285334):0.312712):0.063648,(Tg:0.088904,Et:0.414483):0.322909):0.105011,Cp:0.470225):0,Tt:1.101456);
 (((Et:99.999999,(Ta:0.066051,Bb:0.11471):0.05728):3.1e-05,((Pv:3e-06,Pf:0.023854):0.121211,Cp:0.126021):0.076473):0.079058,Tg:0.050007):0,Tt:0.516324);
 (((Ta:0.076954,Bb:0.082923):0.131292,(Tg:0.053512,Et:0.063473):0.091215,(Pv:1e-06,Pf:0.01448):0.077435):0.052273):0.060623,Cp:0.100404):0,Tt:0.830818);
 (((Pv:0.021658,Pf:0.041601):0.381262,Cp:0.420875):0.070897,(Tg:0.119323,Et:0.209065):0.072014,(Ta:0.193489,Bb:0.227131):0.260417):0.038597):0,Tt:1.237378);
 (((((Tg:0.078563,Et:0.333408):0.326933,(Ta:0.173466,Bb:0.357598):0.410468):0.134255,(Pv:0.106376,Pf:0.04425):0.321962):0.101597,Cp:0.558084):0,Tt:0.982213);
 (((Bb:0.012204,Ta:0.423542):0.096594,(Tg:1.114763,Et:0.047377):0.006382,Cp:0.650909):0.108216):0.040797,(Pv:0.075167,Pf:1e-06):0.162135):0,Tt:0.2948);
 (((((Pv:0.154097,Pf:0.112049):0.309811,(Tg:0.34784,Et:0.37118):0.132237):0.049206,Cp:0.811955):0.221527,(Ta:0.233179,Bb:0.543825):0.086912):0,Tt:1.181579);
 (((((Pv:0.075824,Pf:0.084423):0.494409,(Ta:0.363313,Bb:0.335417):0.132515):0.096444,Cp:0.647453):0.086779,(Tg:0.300422,Et:0.42948):0.227257):0,Tt:0.491854);
 (((((Pv:0.018744,Pf:0.049533):0.273267,(Ta:0.291284,Bb:0.23796):0.306769):0.094499,(Tg:0.071126,Et:0.142207):0.192168):0.146154,Cp:0.466314):0,Tt:2.147951);
 (((Pv:0.042127,Pf:0.028084):0.304067,(Bb:0.22759,Ta:0.954716):0.153334):0.088189,(Cp:0.384261,(Tg:0.064078,Et:0.087138):0.085554):0.024438):0,Tt:0.578663);

(((((Pv:0.08939,Pf:0.035963):0.299913,(Ta:0.199775,Bb:0.19524):0.217463):0.108807,Cp:0.331656):0.034971,(Tg:0.226774,Et:0.405346):0.296787):0,Tt:0.628091);
 (((((Pv:0.008599,Pf:0.007068):0.151523,(Ta:0.180141,Bb:0.150228):0.089184):0.029297,(Tg:0.056454,Et:1.032657):0.168964):0.053444,Cp:0.242544):0,Tt:0.241587);
 (((Tg:0.104772,Et:0.11104):0.137187,(Pv:0.03396,Pf:0.018173):0.245983,(Ta:0.18118,Bb:0.135294):0.107956):0.079395,Cp:0.331723):0,Tt:1.992714);
 (((Et:0.41829,((Pv:0.031876,Pf:1e-06):0.253909,(Ta:0.044121,Bb:0.312364):0.13388):0.037538,Tg:0.107613):0.035219):0.133796,Cp:0.2053368):0,Tt:0.765788);
 (((Ta:0.524239,Bb:0.330469):0.358397,(Pv:0.174469,Pf:0.354136):0.54315):0.117806,((Cp:0.684463,Et:0.405948):0.097723,Tg:0.344686):0.275478):0,Tt:0.832437);
 (((Pv:0.119539,Pf:0.158954):0.680636,(Ta:0.679867,Bb:0.494603):0.541027):0.173536,(Et:1.055297,Tg:0.233226):0.26692,Cp:1.146918):0.165594):0,Tt:0.978399);
 (((Ta:0.072856,Bb:0.065646):0.09507,(Pv:1e-06,Pf:0.018711):0.106543):0.041062,(Tg:0.107736,Et:0.267948):0.278273):0.04603,Cp:0.194103):0,Tt:1.285246);
 (((((Pv:0.095738,Pf:0.123039):0.214234,(Ta:0.254538,Bb:0.298065):0.229085):0.101808,Et:0.880888):0.057514,Tg:0.399353):0.072869,Cp:0.288729):0,Tt:0.870463);
 (((Pv:0.050713,Pf:0.096938):0.500754,(((Tg:0.140265,Et:0.19504):0.114759,(Ta:0.153161,Bb:0.097176):0.27264):0.138762,Cp:0.420354):0.050285):0,Tt:0.45006);
 (((Pv:1e-06,Pf:0.007889):0.08671,(((Ta:0.03197,Bb:0.07948):0.140726,Cp:0.066223):0.017349,Tg:1e-06):0.007585):0.012129,Et:0.024126):0,Tt:0.020492);
 (((((Pv:0.108212,Pf:0.034728):0.680816,Cp:0.565365):0.048219,Et:100):1e-06,((Ta:0.381365,Bb:0.196111):0.545761,Tg:0.359448):0.095824):0,Tt:0.541188);
 (((((Pv:0.124516,Pf:0.072591):0.658292,(Ta:0.262733,Bb:0.354608):0.302822):0.121032,(Et:0.6209,Tg:0.381123):0.304884):0.137245,Cp:0.690004):0,Tt:0.520432);
 (((Ta:0.138519,Bb:0.158637):0.168739,(Pv:0.033536,Pf:0.027712):0.174158):0.041068,(Tg:0.168529,Et:0.210581):0.115811,Cp:0.338082):0.069868):0,Tt:0.624287);
 (((Pv:0.066269,Pf:0.143214):0.357196,((Cp:0.964411,(Ta:0.186299,Bb:0.163371):0.091223):0.06362,(Tg:0.154213,Et:0.288045):0.189308):0.097623):0,Tt:0.750155);
 (((Ta:0.043678,Bb:0.073143):0.122364,(Pv:0.006499,Pf:0.006251):0.076212):0.048243,(Tg:0.027897,Et:0.038301):0.068871,Cp:0.088401):0.031287):0,Tt:0.959667);
 (((Tg:0.342942,Et:0.195346):0.13679,(Ta:0.543854,Bb:0.329835):0.151936,(Pv:0.071893,Pf:0.157536):0.404897):0.153894):0.394134,Cp:0.728881):0,Tt:1.026702);
 (((((Pv:2e-06,Pf:0.009982):0.094886,(Tg:0.016528,Et:0.043378):0.022967):0.059128,(Ta:0.057786,Bb:0.044046):0.158352):0.080663,Cp:0.09457):0,Tt:0.517277);
 (((Pv:0.025207,Pf:0.05035):0.255072,Cp:0.630452):9e-05,(((Ta:0.42381,Bb:0.206967):0.193525,Et:0.527499):0.055724,Tg:0.088031):0.116533):0,Tt:100);
 (((Pv:0.045282,Pf:0.028196):0.403583,((Tg:0.06844,Et:0.45975):0.144099,Cp:0.284151):0.055286):0.063826,(Ta:0.240829,Bb:0.203575):0.16795):0,Tt:1.095438);
 (((Tg:0.456933,Et:0.249355):0.23083,Cp:0.677242):0.175847,(Pv:0.062741,Pf:0.114902):0.558076):0.093479,(Ta:0.346573,Bb:0.455427):0.382366):0,Tt:1.443289);

Bibliography

- [AI72] Edward N Adams III. Consensus techniques and the comparison of taxonomic trees. *Systematic Biology*, 21(4):390–397, 1972.
- [AP94] Yash P Aneja and M Parlar. Algorithms for Weber facility location in the presence of forbidden regions and/or barriers to travel. *Transportation Science*, 28(1):70–76, 1994.
- [Bac14a] Miroslav Bacák. Computing medians and means in Hadamard spaces. *SIAM Journal on Optimization*, 24(3):1542–1566, 2014.
- [Bac14b] Miroslav Bacák. *Convex analysis and optimization in Hadamard spaces*, volume 22. Walter de Gruyter GmbH & Co KG, 2014.
- [BDS91] Martin Barrett, Michael J Donoghue, and Elliott Sober. Against consensus. *Systematic Zoology*, 40(4):486–493, 1991.
- [BHV01] Louis J Billera, Susan P Holmes, and Karen Vogtmann. Geometry of the space of phylogenetic trees. *Advances in Applied Mathematics*, 27(4):733–767, 2001.
- [BKW03] Jack Brimberg, H Taghizadeh Kakhki, and George O Wesolowsky. Location among regions with varying norms. *Annals of Operations Research*, 122(1-4):87–102, 2003.
- [BKW05] Jack Brimberg, H Taghizadeh Kakhki, and George O Wesolowsky. Locating a single facility in the plane in the presence of a bounded region and different norms. *Journal of the Operations Research Society of Japan*, 48(2):135–147, 2005.
- [BLO⁺13] Dennis Barden, Huiling Le, Megan Owen, et al. Central limit theorems for Fréchet means in the space of phylogenetic trees. *Electron. J. Probab.*, 18(25):1–25, 2013.
- [BO17] Daniel G Brown and Megan Owen. Mean and Variance of Phylogenetic Trees. *arXiv preprint arXiv:1708.00294*, 2017.
- [Bot18] Marco Botte. Fixed gate point location problems. 2018.

- [Bry03] David Bryant. A classification of consensus methods for phylogenetics. *DIMACS series in discrete mathematics and theoretical computer science*, 61:163–184, 2003.
- [Coh17] Kai-Hendrik Cohrs. Schranken für Medianprobleme im phylogenetischen Tree Space, 2017.
- [DR06] James H Degnan and Noah A Rosenberg. Discordance of species trees with their most likely gene trees. *PLoS genetics*, 2(5):e68, 2006.
- [EH72] Jack Elzinga and Donald W Hearn. Geometrical solutions for some minimax location problems. *Transportation Science*, 6(4):379–394, 1972.
- [Fel85] Joseph Felsenstein. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39(4):783–791, 1985.
- [FGK03] Kaspar Fischer, Bernd Gärtner, and Martin Kutz. Fast smallest-enclosing-ball computation in high dimensions. In *European Symposium on Algorithms*, pages 630–641. Springer, 2003.
- [FVGA12] L Franco, Francisco Velasco, and L Gonzalez-Abril. Gate points in continuous location between regions with different l-p norms. *European Journal of Operational Research*, 218(3):648–655, 2012.
- [Gro87] Mikhael Gromov. Hyperbolic groups. In *Essays in group theory*, pages 75–263. Springer, 1987.
- [Hae] Ernst Haeckel. Tree of Life.
- [HHL⁺13] Thomas Hotz, Stephan Huckemann, Huiling Le, James S Marron, Jonathan C Mattingly, Ezra Miller, James Nolen, Megan Owen, Vic Patrangenaru, Sean Skwerer, et al. Sticky central limit theorems on open books. *The Annals of Applied Probability*, 23(6):2238–2258, 2013.
- [HSSC08] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using NetworkX. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.
- [KC81] I Norman Katz and Leon Cooper. Facility location in the presence of forbidden regions, I: Formulation and the case of Euclidean distance with one forbidden circle. *European Journal of Operational Research*, 6(2):166–173, 1981.

- [KMY03] Piyush Kumar, Joseph SB Mitchell, and E Alper Yildirim. Approximate minimum enclosing balls in high dimensions using core-sets. *Journal of Experimental Algorithmics (JEA)*, 8:1–1, 2003.
- [Kuh73] Harold W Kuhn. A note on Fermat’s problem. *Mathematical programming*, 4(1):98–107, 1973.
- [KWK08] Chih-Horng Kuo, John P Wares, and Jessica C Kissinger. The Apicomplexan whole-genome phylogeny: An analysis of incongruence among gene trees. *Molecular biology and evolution*, 25(12):2689–2698, 2008.
- [LMY18] Bo Lin, Anthea Monod, and Ruriko Yoshida. Tropical Foundations for Probability & Statistics on Phylogenetic Tree Space. *arXiv preprint arXiv:1805.12400*, 2018.
- [LPT95] Martine Labbé, Dominique Peeters, and Jacques-François Thisse. Location on networks. *Handbooks in operations research and management science*, 8:551–624, 1995.
- [Mad97] Wayne P Maddison. Gene trees in species trees. *Systematic biology*, 46(3):523–536, 1997.
- [Meg83] Nimrod Megiddo. The weighted Euclidean 1-center problem. *Mathematics of Operations Research*, 8(4):498–504, 1983.
- [MM81] Timothy Margush and Fred R McMorris. Consensus-trees. *Bulletin of Mathematical Biology*, 43(2):239–244, 1981.
- [MMN83] Fred R McMorris, David B Meronk, and Dean A Neumann. A view of some consensus methods for trees. In *Numerical Taxonomy*, pages 122–126. Springer, 1983.
- [MOP12] Ezra Miller, Megan Owen, and J Scott Provan. Averaging metric phylogenetic trees. *arXiv preprint arXiv:1211.7046*, 2012.
- [MOP15] Ezra Miller, Megan Owen, and J Scott Provan. Polyhedral computational geometry for averaging metric phylogenetic trees. *Advances in Applied Mathematics*, 68:51–91, 2015.
- [MP91] Joseph SB Mitchell and Christos H Papadimitriou. The weighted region problem: Finding shortest paths through a weighted planar subdivision. *Journal of the ACM (JACM)*, 38(1):18–73, 1991.
- [Nel93] G Nelson. Why crusade against consensus? A reply to Barrett, Donoghue, and Sober. *Systematic Biology*, 42(2):215–216, 1993.

- [NTWY17] Tom MW Nye, Xiaoxian Tang, Grady Weyenberg, and Ruriko Yoshida. Principal component analysis and the locus of the Fréchet mean in the space of phylogenetic trees. *Biometrika*, 104(4):901–922, 2017.
- [Nye11] Tom MW Nye. Principal components analysis in the space of phylogenetic trees. *The Annals of Statistics*, pages 2716–2739, 2011.
- [Nye15] Tom MW Nye. Convergence of random walks to Brownian motion in phylogenetic tree-space. *arXiv preprint arXiv:1508.02906*, 2015.
- [OP11] Megan Owen and J Scott Provan. A fast algorithm for computing geodesic distances in tree space. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 8(1):2–13, 2011.
- [Owe08] Megan Owen. *Distance computation in the space of phylogenetic trees*. PhD thesis, Cornell University, 2008.
- [Owe11] Megan Owen. Computing geodesic distances in tree space. *SIAM Journal on Discrete Mathematics*, 25(4):1506–1529, 2011.
- [Par94] M Parlar. Single facility location problem with region-dependent distance metrics. *International journal of systems science*, 25(3):513–525, 1994.
- [PC97] Roderic DM Page and Michael A Charleston. From gene to organismal phylogeny: Reconciled trees and the gene tree/species tree problem. *Molecular phylogenetics and evolution*, 7(2):231–240, 1997.
- [Pla05] Frank Plastria. 4-point Fermat location problems revisited. New proofs and extensions of old results, 2005.
- [Roc15] Ralph T Rockafellar. *Convex analysis*. Princeton university press, 2015.
- [Stu02] Karl-Theodor Sturm. Nonlinear Markov operators, discrete heat flow, and harmonic maps between singular spaces. *Potential Analysis*, 16(4):305–340, 2002.
- [Stu03] Karl-Theodor Sturm. Probability measures on metric spaces of nonpositive curvature. *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces: Lecture Notes from a Quarter Program on Heat Kernels, Random Walks, and Analysis on Manifolds and Graphs: April 16-July 13, 2002, Emile Borel Centre of the Henri Poincaré Institute, Paris, France*, 338:357, 2003.
- [Vog07] Karen Vogtmann. Geodesics in the space of trees, 2007.

- [VZ00] Yehuda Vardi and Cun-Hui Zhang. The multivariate L1-median and associated data depth. *Proceedings of the National Academy of Sciences*, 97(4):1423–1426, 2000.
- [War63] Jack Warga. Minimizing certain convex functions. *Journal of the Society for Industrial and Applied Mathematics*, 11(3):588–593, 1963.
- [War83] Arthur R Warburton. Quasiconcave vector maximization: Connectedness of the sets of Pareto-optimal and weak Pareto-optimal alternatives. *Journal of optimization theory and applications*, 40(4):537–557, 1983.
- [Web22] Alfred Weber. *Über den Standort der Industrien. Teil 1+ 2*. JCB Mohr, 1922.
- [Wei37] Endre Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematical Journal, First Series*, 43:355–386, 1937.
- [WYH17] Grady Weyenberg, Ruriko Yoshida, and Daniel Howe. Normalizing kernels in the Billera-Holmes-Vogtmann treespace. *IEEE/ACM transactions on computational biology and bioinformatics*, 14(6):1359–1365, 2017.
- [YZZ17] Ruriko Yoshida, Leon Zhang, and Xu Zhang. Tropical principal component analysis and its application to phylogenetics. *arXiv preprint arXiv:1710.02682*, 2017.