# The convergence rate of the Sandwiching algorithm for convex bounded multiobjective optimization

Ina Lammel and Karl-Heinz Küfer

Department of Optimization
Fraunhofer Institute for Industrial Mathematics (ITWM)
Fraunhofer Platz 1, 67663 Kaiserslautern, Germany
{ina.lammel, karl-heinz.kuefer}@itwm.fraunhofer.de

December 18, 2023

Sandwiching algorithms, also known as Benson-type algorithms, approximate the nondominated set of convex bounded multiobjective optimization problems by constructing and iteratively improving polyhedral inner and outer approximations. Using a set-valued metric, an estimate of the approximation quality is determined as the distance between the inner and outer approximation. The convergence of the algorithm is evaluated with respect to this approximation quality.

We show the convergence rate of a class of Sandwiching algorithms by extending results for a similar algorithm for the approximation of convex compact sets. In particular, we derive requirements for the nondominated set to have a twice continuously differentiable boundary. We show that two common quality indicators, the polyhedral gauge and the epsilon indicator, fulfill the necessary requirements and explicitly state the convergence rate for these two indicators. Under sufficient regularity assumptions, the convergence rate is optimal.

**Keywords** multiobjective optimization; convex optimization; approximation algorithm; convergence rate; quality indicator

**AMS Classification** 41A25; 52A27; 90C29; 90C59; 90-08

## 1. Introduction

Convex multiobjective optimization problems have several objectives that are to be optimized simultaneously. A common task is to compute the nondominated set. Since the number of nondominated solutions of a multiobjective optimization problem is in general unbounded, approximation algorithms are often used. One common approach is to construct an inner and outer polyhedral approximation of a convex bounded nondominated set using nondominated points and their gradients. Algorithms of this type are often called (Simplicial) Sandwiching algorithms (e.g. [5], [21], [32], [36]) or Benson-type algorithms (e.g. [2], [8], [11], [27]). By iteratively computing additional nondominated points, the inner and outer approximations describe the nondominated set more accurately.

The same algorithmic ideas are also applied to the approximation of convex functions (e.g. [33]) and convex sets (e.g. [17], [18]). Recently, the idea of Sandwiching algorithms has been applied to the approximation of multiple convex nondominated sets, e.g. for multiobjective mixed-integer convex optimization in [6] (bi-objective problems) and [23] (general number of objectives).

In this article, we will investigate the convergence behaviour of a Sandwiching algorithm. In particular, we show that the algorithm converges and how fast the approximation error decreases with the number of Sandwiching iterations. This property is also called convergence rate (see [33], [17], [21]). In contrast to the rate of convergence of sequences from numerical analysis, this convergence rate is a property of sets. We obtain the convergence rate result by extending the approach given in [17] for an algorithm for the inner and outer approximation of convex compact sets to our Sandwiching algorithm for multiobjective optimization and a class of different quality indicators.

For a variant of the Sandwiching algorithm, applied to the approximation of convex functions, a proof of the convergence rate has been derived in 1992 for the two-dimensional case [33]. This result has been transferred to the approximation of two-dimensional nondominated sets in [21].

For another Sandwiching variant used to approximate convex compact sets by polytopes, a proof of the convergence rate of an inner and outer approximation algorithm of convex compact sets has been published in 1996 by Kamenev for the general-dimensional case [17]. The article was published in Russian only[1]. Therefore, we state some straightforward extensions of proofs given in [17] in the appendix to make them accessible to the English-speaking community.

In this article, we will first introduce the Sandwiching algorithm for the approximation of the nondominated sets of convex bounded multiobjective optimization problems. Then we will state the algorithm for the inner and outer approximation of convex compact sets introduced in [17] and expand it so that it can be related to our Sandwiching algorithm for convex bounded multiobjective optimization.

Recently, in [3] a proof of the convergence rate of a polyhedral outer approximation algorithm of convex bounded nondominated sets was introduced, also based on Kamenev's

[1]We obtained the article in Russian and translated it. The article is cited in [28]. The citation list of the book mentions that the article has been translated to English. However, the English version cannot be found using the inter-library loan database digibib.

work.

In Section 4 we extend the convergence rate results to the expanded algorithm with a class of quality indicators. Then, we apply the results to the Sandwiching algorithm for convex bounded multiobjective optimization problems in Section 5. The main results of this article are given in Sections 5.2 and 5.3 where the convergence results for the Sandwiching algorithm with the quality indicators polyhedral gauge and epsilon indicator are derived.

Finally, we discuss the quality of the proved convergence rates. Under sufficient regularity assumptions, the convergence rate of the Sandwiching algorithm is optimal.

# 2. Preliminaries

## 2.1. Bounded convex multiobjective optimization

A convex multiobjective optimization problem is defined as

$$\min f(x) = (f_1(x), \dots, f_d(x)) \quad \text{subject to } x \in \mathcal{X}, \tag{1}$$

(Definition 2.1.3 of [30]), where $f(x)$ denotes the vector of $d$ convex objective functions $f_i : \mathbb{R}^k \to \mathbb{R}$, and the decision vectors $x \in \mathbb{R}^k$ are elements of the convex feasible set $\mathcal{X}$. We denote $\mathcal{Y} := \{f(x) \forall x \in \mathcal{X}\}$. We assume that the problem is solved with respect to a pointed, convex ordering cone $\mathcal{C}$ fulfilling $0 \in \mathcal{C}$ and $\mathcal{C} \supset \mathbb{R}^d_{\geq}$. Additionally, we assume that the problem is bounded in the sense that there is a point $\hat{p} \in \mathbb{R}^d$ such that $\text{cl}(\mathcal{Y} + \mathcal{C}) \subset \{\hat{p}\} + C$ where $\text{cl}(\cdot)$ denotes the closure [37].

A feasible solution $\hat{x} \in \mathcal{X}$ is called *efficient* if there is no other $x \in \mathcal{X}$ such that $f(x) \leq f(\hat{x})$. The set of all efficient points is denoted $\mathcal{X}_E$. If $\hat{x}$ is efficient, $f(\hat{x})$ is called *nondominated* .The set of all nondominated points is denoted $\mathcal{Y}_N$. A feasible solution $\hat{x} \in \mathcal{X}$ is called *weakly efficient* if there is no $x \in \mathcal{X}$ such that $f(x) < f(\hat{x})$, i.e. $f_i(x) < f_i(\hat{x})$ $\forall i = 1, \dots, d$. If $\hat{x}$ is weakly efficient, $f(\hat{x})$ is called *weakly nondominated* [10].

The weighted sum scalarization of problem is given by ((3.3) of [10])

$$\min_{x \in \mathcal{X}} \sum_{i=1}^{d} \lambda_i f_i(x) \tag{2}$$

with values $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d_{\geq}$. Well-known results include that for a convex optimization problem, every solution of the weighted sum problem is weakly nondominated and that every element of the nondominated set can be computed by solving weighted sum scalarization problems for $\lambda \in \mathcal{C}^* \backslash \{0\}$ where $\mathcal{C}^*$ denotes the dual cone (e.g. [10]).

A vector $k \in \mathbb{R}^d \backslash \{0\}$ is called a direction of the cone $\mathcal{C}$ if $\{c + \alpha k \in \mathbb{R}^d : c \in \mathcal{C}, \alpha > 0\} \subset \mathcal{C}$. The set of extreme directions of a convex cone is a minimal set of directions such that all directions of the cone lie in their convex hull. If the solution of the weighted sum scalarization problem (2) for all extreme directions of the dual cone $\mathcal{C}^*$ of $\mathcal{C}$ is bounded, the convex multiobjective optimization problem is bounded [37].

The point $y^N = (y^N_1, \dots, y^N_d)$ given by $y^N_k := \max_{x \in \mathcal{X}_E} f_k(x) = \max_{y \in \mathcal{Y}_N} y_k$is called the nadir point of the multiobjective optimization problem (Definition 2.22.2 of [10]).

## 2.2. The Sandwiching Algorithm and common quality indicators

### 2.2.1. The Algorithmic Idea of Sandwiching Algorithms

Sandwiching algorithms are used to approximate the nondominated sets of bounded convex multiobjective optimization problems.

To apply the convergence rate proof that will be presented in this article, the set that is approximated by the Sandwiching algorithm needs to be a convex compact set. Adding the domination cone to the feasible objective set does not have any effect on its nondominated set $(\mathcal{Y})_N = (\mathcal{Y} + \mathcal{C})_N$ (Lemma 1.1.5 of [22]), where $(\cdot)_N$ denotes the nondominated set operator. We use the property that $\mathcal{Y} + \mathcal{C}$ is a convex set (Lemma 1.19 of [36]). To obtain a bounded set, we remove those parts of the nondominated set exceeding the nadir point $y^N$ of the multiobjective optimization problem. No efficient method for determiming the nadir point for general multiobjective optimization problems is known [10]. Therefore, in practice the nadir point is approximated as $\tilde{y}^N$. Together, we obtain the set that will be approximated by the Sandwiching algorithm

$$\mathcal{P} := (\mathcal{Y} + \mathcal{C})_N \cap (\tilde{y}^N - \mathbb{R}^d_{\geq}). \tag{3}$$

Since we assume that the multiobjective optimization problem is bounded, removing the part of the nondominated set exceeding $\tilde{y}^N$ makes $\mathcal{P}$ a convex compact set. If the nadir estimation is exact, i.e. $\tilde{y}^N = y^N$, then $(\mathcal{P})_N = \mathcal{Y}_N$. If the nadir approximation is not dominated by the true nadir point, a part of the nondominated set may be cut off. Instead of $\tilde{y}^N$, a different upper bound on the part of the nondominated set that the decision maker is interested in can be used.

A similar approach of obtaining a convex compact set that is approximated by the Sandwiching algorithm is used in [2], where a concave minimization problem is solved to construct a halfspace that contains the nondominated set.

The Sandwiching algorithm as used in [22], [5], [36] is given in Algorithm 1. After an initial approximation, e.g. consisting of the extreme compromise solutions, has been computed, an inner and an outer approximation are constructed. The inner approximation is defined as the convex hull of the nondominated points, extended by the domination cone. The outer approximation is defined as the intersection of the half-spaces containing the nondominated set that support the nondominated points. Due to convexity, the nondominated set is contained in the outer approximation and contains the inner approximation.

Then, the approximation quality is determined. The approximation quality is the maximal minimal distance between the inner and the outer approximation with respect to some quality indicator. It is an upper bound on the approximation error. Common quality indicators are the epsilon indicator and the polyhedral gauge. The facet of the inner approximation where the largest distance to the outer approximation was attained is used as the starting point for computing a new nondominated point $z^{n+1}$ by using the facet normal as the parameter for the weighted sum scalarization problem. Then, the tangential hyperplane of the resulting weakly nondominated point is parallel to this facet of the former inner approximation.

Then, the inner approximation is updated by $I^{n+1} = \text{conv}\{z^{n+1}, I^n\}$, the outer approximation is updated by $O^{n+1} := HS(w^{n+1}, b^{n+1}) \cap O^n$.

**Algorithm 1** Scheme of the Sandwiching algorithm

**Require:** Target quality $\epsilon$, quality indicator $\delta$

Set $n := 0$. Compute initial approximations $I^0$, $O^0$, nadir point approximation $\tilde{y}^N$, initial approximation quality $\delta := \delta(I^0, O^0)$ and candidate weight $\pi^0$.

**while** $\delta > \epsilon$ **do**

  Set $n := n + 1$.

  Solve $z := \min\{\sum_{i=1}^{d} \pi_i^{n-1} f_i(x),\ x \in \mathcal{X}\}$ to obtain a new nondominated point.

  Update $I^n := \text{conv}\{z, I^{n-1}\}$, $O^n := H(\pi^{n-1}, \pi^{n-1}z) \cap O^n$.

  Compute the approximation quality $\delta := \delta(I^n, O^n)$ and compute the candidate weight $\pi^n$.

**return** $I^n$, $O^n$, $\delta$

### 2.2.2. Common Quality Measures: Epsilon Indicator and Polyhedral Gauge

One quality criterion commonly used in Sandwiching algorithms (e.g. [5], [32]) but also other algorithms approximating nondominated sets or studies in multiobjective optimization, e.g. in [7], [25] or [38] is the *epsilon indicator* or *$\epsilon$-indicator*.

**Definition 2.1.** The epsilon indicator $\delta^\epsilon(I, O)$ of a Sandwiching approximation $I, O$ is the smallest number $\epsilon \geq 0$ such that for every $z \in O$ there exists a point in the inner approximation $z' \in I$ such that $z' \leq z + \epsilon \cdot e$ where $e = (1, \dots, 1) \in \mathbb{R}^d$ (Definitions 1 and 2 of [7]).

To determine the epsilon indicator between an inner and outer Sandwiching approximation, it suffices to determine the epsilon indicator between vertices of the outer approximation and the inner approximation (Proposition 4.1 of [5]).

While the epsilon indicator is an absolute measure of the approximation quality, there are also relative quality indicators. One example that is used to measure the approximation quality of the Sandwiching approximation, for example in [21] and [36], is based on the *polyhedral gauge*. We translate the bounded nondominated set such that the nadir point lies in the origin. For $I^n$, we define the reflected inner approximation as the reflection set of $I^n$ (see Definition 2 of [34])

$$I_R^n := R(I^n) := \bigcup_{z \in I^n} \left\{ w \in \mathbb{R}^d : |w_i| = |z_i| \ \forall i = 1, \dots, d \right\}.$$

**Definition 2.2.** For a vertex $s$ of the outer approximation, the polyhedral gauge

$$\gamma_{I_R^n}(s) := \min\{\lambda \geq 0 : \ z \in \lambda I_R^n\}$$

(Definition 2.1.1 of [21]) describes the factor by which $I_R^n$ has to be scaled to reach $s$. As a quality criterion we use the polyhedral gauge subtracted by one so that the quality value is zero for exact approximation.

Since all nondominated points and $s$ lie in the same orthant $\mathbb{R}^d_{\leq}$, we can omit reflecting $I^n$ and work with it directly.

Both the epsilon indicator and the polyhedral gauge distance between the inner and outer approximation of a Sandwiching algorithm can be computed by solving a linear program. For an efficient way to compute these qualities, see [24].

### 2.3. The convergence rate

The convergence rate is an asymptotic property of an approximation algorithm. It describes how fast the approximation error decreases with the number of iterations. While this definition is used e.g. in [33], [17], [21] and [3], there also exist different definitions of convergence rates in literature which might not be equivalent.

**Definition 2.3.** Let $I^n$ be the inner, $O^n$ the outer approximation constructed by the Sandwiching algorithm in iteration $n$ of a convex compact set or of the nondominated set of a bounded convex multiobjective optimization problem.

Then the Sandwiching algorithm has the convergence rate $r$ if $r$ is the largest real number such that for every $\epsilon > 0$ there exists a $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ it holds

$$d(I^n, O^n) \leq C(\epsilon)n^{-r}$$

for a constant $C$ depending on $\epsilon$ and a metric $d$ on finite-dimensional nonempty compact sets, e.g. the Hausdorff metric.

## 3. Polyhedral approximation of convex compact sets

### 3.1. Introducing the Kamenev algorithm for the approximation of convex compact sets

In his 1996 article [17] (in Russian), Kamenev introduces an algorithm for the simultaneous inner and outer approximation of convex compact sets. To distinguish between those algorithms approximating convex compact sets and those algorithms that approximate convex bounded nondominated sets, we use different notation. For the algorithms approximating nondominated sets, the inner approximation is denoted by $I$, the outer approximation is denoted by $O$. When approximating convex compact sets, we denote the polytope forming the inner approximation by $P$ and the outer approximation by $Q$.

Before stating the algorithm, we introduce some notation and definitions.

We denote by $B_\rho(x)$ a ball of radius $\rho$ with centre $x$. The volume of a unit ball in $\mathbb{R}^d$ is denoted by $\pi_d$.

**Definition 3.1** ( [17]). Let $\mathscr{C}$ be the set of convex compact subsets of $\mathbb{R}^d$ with nonempty interior.

For $C \in \mathscr{C}$ we denote by $\partial C$ the boundary of $C$. Its asphericity $\alpha(C)$ is defined as the minimal ratio of the radii of concentric outer and inner spheres. The surface area of $C$ is denoted by $\sigma(C)$.

For $C \in \mathscr{C}$ and $u \in \mathbb{R}^d$ let us introduce

- the *support function* of $C$: $g_C(u) := \max\{\langle u, x\rangle : x \in C\}$,
- the *supporting half-space*: $H_C(u) := \{x \in \mathbb{R}^d : \langle u, x\rangle \leq g_C(u)\}$,
- the *supporting hyperplane*: $\partial H_C(u) = \{x \in \mathbb{R}^d : \langle u, x\rangle = g_C(u)\}$ and
- the set of *tangent points*: $T_C(u) := C \cap \partial H_C(u) = \{x \in C : \langle u, x\rangle = g_C(u)\}$,

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product on $\mathbb{R}^d$.

**Definition 3.2.** Let $A \subset \mathbb{R}^d$ be a compact subset. The set $A$ has twice continuously differentiable boundary if for every boundary point $a \in \partial A$ there exists an open neighbourhood $U \subset \mathbb{R}^d$ and a twice continuously differentiable function $\psi : U \to \mathbb{R}$ that fulfils $A \cap U = \{x \in U : \psi(x) \leq 0\}$ and $\nabla\psi(x) \neq 0\ \forall x \in U$ (adaptation of [12] p. 179).

Let $\mathscr{C}^2$ be the set of convex compact subsets of $\mathbb{R}^d$ with non-empty interior, with twice continuously differentiable boundary and positive principal curvatures ( [17]).

For $C \in \mathscr{C}^2$ the radius of curvature of a point $x \in \partial C$ is given by the maximal radius of balls $B_r(y)$ for some $y \in C$ that fulfil $B_r(y) \subset C$ and $x \in \partial B_r(y)$. We denote by $\rho_{\min}(C)$ and $\rho_{\max}(C)$ the minimum and maximum radii of curvature of the surface of $C$.

**Definition 3.3** (p. 16f of [9])**.** A polyhedron is defined as the intersection of a finite number of closed half-spaces. A bounded polyhedron is also called a *polytope*.

**Definition 3.4** ( [17])**.** Let $\mathscr{P}$ be the set of convex polytopes. For $P \in \mathscr{P}$ we denote by $n^v(P)$ the number of its vertices, and by $n^f(P)$ the number of its facets.

For $C \in \mathscr{C}$ let us introduce the set of inscribed polytopes $\mathscr{P}^i(C) \subset C$ whose vertices belong to the boundary $\partial C$ of $C$, and the set of circumscribed polytopes $\mathscr{P}^c(C) \supset C$ whose facets touch $\partial C$.

For $P \in \mathscr{P}$ we denote by $U(P) \subset S^{d-1} := \{x : ||x|| = 1\}$ the set of vectors of unit outer normals to the facets of $P$.

**Definition 3.5** ( [17])**.** A $(d-1)$-dimensional face of a polytope $P \in \mathscr{P}$, $P \subset \mathbb{R}^d$ is a called a *facet* of $P$.

In the case $d = 3$, for example, triangular faces are facets, while one-dimensional faces are segments and zero-dimensional faces are points.

Next, we introduce the algorithm given in [17] for the inner and outer approximation of convex compact sets with nonempty interior.

We construct an inner and outer approximation of $C \in \mathscr{C}$. Let an initial inner and outer approximation $P^0 \in \mathscr{P}^i(C)$, $Q^0 \in \mathscr{P}^c(C)$ be given. Algorithm 2 describes the $(n+1)$-th iteration of the algorithm. First, a facet of $P^n$ with normal $u^*$ is determined where the support function distance between the inner and outer approximation is maximal. If the distance does not vanish, a point $p^* \in \partial C$ is determined that has normal $u^*$. Then, the inner approximation is augmented by the new point $p^*$ by forming the convex hull $\operatorname{conv}\{\{p^*\} \cup P^n\}$. The supporting hyperplane of $C$ in $p^*$ cuts the former outer approximation to yield the next outer approximation.

**Algorithm 2** Step $n+1$ of the inner and outer approximation of convex compact sets [17]

Let $P^n \in \mathscr{P}^i(C)$, $Q^n \in \mathscr{P}^c(C)$ be constructed.
Find $u^* \in \arg\max\{g_{Q^n}(u) - g_{P^n}(u) : u \in U(P^n)\}$
**if** $g_{Q^n}(u^*) - g_{P^n}(u^*) \neq 0$ **then**
    Find $p^* \in T_C(u^*)$
    Set $P^{n+1} = \text{conv}\{\{p^*\} \cup P^n\}$.
    Set $Q^{n+1} = Q^n \cap H_C(u^*)$.
**else**
    Stop.

Let us denote by $L_C(n)$ the number of problems of calculating the distance between inner and outer approximation and the problems of finding a point of contact of $C$ with its supporting hyperplane, solved for finding approximating polytope $P^n$ and $Q^n$. Then, (see [17])

$$n^v(P^n) \leq n^v(P^0) + n, \quad n^f(Q^n) \leq n^f(Q^0) + n, \quad L_C(n) = L_C(0) + n. \tag{4}$$

### 3.2. Reformulating the Kamenev algorithm for a class of quality criteria

The Kamenev algorithm [17] (Algorithm 2) selects the facet of the inner approximation starting from which a new point will be added by computing

$$u^* = \arg\max\{g_{Q^n}(u) - g_{P^n}(u) : u \in U(P^n)\},$$

the facet normal $u^*$ is selected where the support function distance between the inner and outer approximation is maximal. The term $g_{Q^n}(u) - g_{P^n}(u)$ for $u \in U(P^n)$ is therefore used as a quality indicator of the current approximation.

We modify Algorithm 2 to allow different quality indicators and introduce a stopping criterion. In the following, we will call this algorithm *Sandwiching with variable selection*. We will call the Algorithm 2 version given in [17] *Sandwiching with support function selection*.

Again, we assume that some initial inner and outer approximations are given. The following Algorithm 3 describes the $(n+1)$-th iteration of the algorithm.

**Algorithm 3** Sandwiching algorithm with variable selection

Select a facet of the inner approximation by its normal $u^*$ that attains the largest distance (by some (pseudo)metric) to the outer approximation.
**if** the calculated distance does not satisfy the stopping criterion **then**
    Find $p^* \in T_C(u^*)$ ($p^* \in \partial C$, $u^*$ is unit outer normal of $C$ in $p^*$).
    Add $p^*$ to the inner approximation: set $P^{n+1} = \text{conv}\{p^*, P^n\}$.
    Add the halfspace defined by the tangent of $C$ with normal $u^*$ to the outer approximation: $Q^{n+1} = Q^n \cap H_C(u^*)$.

# 4. Proving the convergence rate of an inner and outer approximation algorithm for convex compact sets

In the following, we derive a convergence proof and a bound on the decrease of the approximation error in the course of Algorithm 3. We achieve this by extending the proof of the convergence rate of Algorithm 2 given in [17] to the more general Algorithm 3. To provide a self-contained proof, we state some statements from previous work of the same author, mostly from [16].

## 4.1. The quality criterion of the Kamenev algorithm and equivalent metrics

The quality criterion $\delta_P$ of Algorithm 2 is defined using the support function. In this section we will derive an upper and lower bound on the value $\delta_P(P, C)$ using the Hausdorff distance $\delta^H(P, C)$. We introduce the class of metrics that are strongly equivalent to $\delta^H$ and extend the bounds to those metrics.

**Definition 4.1** (From [17])**.** For $C, C' \in \mathscr{C}$ and $P \in \mathscr{P}$ let us define

$$\delta_P(C, C') = \max\{|g_C(u) - g_{C'}(u)| : u \in U(P)\}. \tag{5}$$

**Lemma 4.2.** *The function $\delta_P$ is a pseudometric on the set of convex compact sets $\mathscr{C}$.*

*Proof.* The values of $\delta_P$ are obviously non-negative and $\delta_P(C, C) = 0 \ \forall C \in \mathscr{C}$. The function is symmetric since for $A, B \in \mathscr{C}$, $\delta_P(A, B) = \max\{|-1| \cdot |g_B(u) - g_A(u)| : u \in U(P)\} = \delta_P(B, A)$.

The triangle inequality follows from the subadditivity of the maximum function and the triangle inequality of the absolute value: Let $A, B, C \in \mathscr{C}$. Then

$$\begin{aligned}\delta_P(A, C) + \delta_P(B, C) \geq\ & \max\{|g_A(u) - g_C(u)| + |g_B(u) - g_C(u)| : u \in U(P)\} \\ \geq\ & \max\{|g_A(u) - g_B(u)| : u \in U(P)\} = \ \delta_P(A, B).\end{aligned}$$

□

Note that $\delta_P$ is not a metric. It can hold $\delta_P(A, B) = 0$ for $A \neq B$ as long as their support function value is the same in the finite number of directions $U(P)$.

**Definition 4.3.** The Hausdorff metric between two nonempty compact sets $C_1, C_2 \subset \mathbb{R}^d$ can be defined using a norm $||\cdot||$ of $\mathbb{R}^d$ as

$$\delta^H_{||\cdot||}(C_1, C_2) = \max\{\sup\{||x - C_2|| : x \in C_1\}, \sup\{||x - C_1|| : x \in C_2\}\},$$

where $||x - C|| := \inf\{||x - y|| : \ y \in C\}$. For the Hausdorff metric with Euclidean norm $||\cdot||_2$ we use the abbreviation $\delta^H(C_1, C_2) := \delta^H_{||\cdot||_2}(C_1, C_2)$.

From [16], we get an upper and lower bound on $\delta_P$ using the Hausdorff metric.

**Lemma 4.4** (Theorem 1 of [16]). *For $C \in \mathscr{C}$ and $P \in \mathscr{P}^i(C)$ it holds*

$$\delta^H(P, C)/\alpha(P) \leq \delta_P(P, C) \leq \delta^H(P, C) \tag{6}$$

*where $\delta^H$ denotes the Hausdorff metric and $\alpha(P)$ the asphericity of $P$, i.e. the minimal ratio of the radii of concentric outer and inner spheres.*

To be able to apply the convergence rate proof derived in [17] to our Sandwiching algorithm, we extended their algorithm to allow different criteria that select the facet that is used to compute a new approximation point (see Algorithm 3). In particular, we will extend the convergence rate proof of [17] to a class of selection criteria that are strongly equivalent to the Hausdorff metric. After defining strong equivalence, we extend Equation 6 of Lemma 4.4 to metrics that are strongly equivalent to the Hausdorff metric. We will call the Sandwiching algorithm with variable selection (Algorithm 3) for selection criteria that are strongly equivalent to $\delta_H$ also *Sandwiching algorithm with strongly equivalent selection.*

**Definition 4.5.** Let $X$ denote a non-empty set and $\delta_A$, $\delta_B$ two metrics on $X$. The metrics $\delta_A$ and $\delta_B$ are *strongly equivalent* if and only if there exist positive constants $c_1, c_2 > 0$ such that for every $x, y \in X$ it holds

$$c_1\delta_A(x, y) \leq \delta_B(x, y) \leq c_2\delta_A(x, y).$$

*Remark* 1.

- When the two metrics $\delta_A$, $\delta_B$ are induced by norms $\|\cdot\|_A$ and $\|\cdot\|_B$ respectively, then strong equivalence of the metrics is equivalent to the equivalence of the norms, i.e. $c_1\|x\|_A \leq \|x\|_B \leq c_2\|x\|_A \; \forall x \in X$.
- The Hausdorff metric $\delta^H_{||\cdot||}$ defined using norm $||\cdot||$ of $\mathbb{R}^d$ is strongly equivalent to $\delta^H_{||\cdot||_2} = \delta^H$ since the norms $||\cdot||_2$ and $||\cdot||$ are equivalent in the finite dimensional vector space $\mathbb{R}^d$.
- From Lemma 4.4 we do not obtain that $\delta_P$ is strongly equivalent to $\delta^H$. Additionally to $\delta_P$ not being a metric, $\alpha(P)$ is not constant.

Let $\delta$ be a metric on the set $\mathscr{C}$ that is strongly equivalent to the Hausdorff metric. Then there exist constants $c_1, c_2 > 0$ such that for $C \in \mathscr{C}$ and $P \in \mathscr{P}^i(C)$ it holds

$$c_1\delta(P, C) \leq \delta^H(P, C) \leq c_2\delta(P, C). \tag{7}$$

Equation 6 describes the relation between $\delta^H(P, C)$ and $\delta_P(P, C)$. Together with the strong equivalence of $\delta$ and $\delta^H$ in Equation 7, we obtain the following relation between $\delta(P, C)$ and $\delta_P(P, C)$:

$$\frac{c_1}{\alpha(P)}\delta(P, C) \leq \frac{1}{\alpha(P)}\delta^H(P, C) \leq \delta_P(P, C) \leq \delta^H(P, C) \leq c_2\delta(P, C). \tag{8}$$

In a part of the convergence rate proof, we will need another lower bound on the support function quality which is an extension of Lemma 1 of [17]. The proof is stated in the Appendix.

**Lemma 4.6.** *Let $C \in \mathscr{C}^2$, $P \in \mathscr{P}^i(C)$, $\delta^H(P,C) < \rho_{\min}(C)$ and $C' \in \mathscr{C}$, $P \in \mathscr{P}^i(C')$. Let $\delta$ be a metric that is strongly equivalent to the Hausdorff metric with constants $c_1, c_2 > 0$. Then*

$$c_2\delta(P,C') \geq \delta^H(P,C') \geq \delta_P(P,C') \geq \left(1 - \frac{\delta^H(P,C)}{\rho_{\min}(C)}\right)\delta^H(P,C').$$

## 4.2. Convergence rate of Sandwiching with strongly equivalent selection criteria

The general idea of the convergence rate proof is to estimate the volume added to the inner and removed from the outer approximation within one iteration of Algorithm 3. This is done in the following way.

In Lemma 4.8, a result estimating the volume of a pyramid defined by a sphere, an external point and a hyperplane is used to estimate the change in volume of the inner and outer approximation in one iteration of the Sandwiching algorithm in terms of the distance between the inner and outer approximation. This result is developed in Lemma 4.9 to a bound on the distance between inner and outer approximation in terms of the iteration number which can also be used to prove the convergence of the algorithm in Theorem 4.10. After estimating the volume difference between the inner and outer approximation from one iteration to the other in Lemma 4.11, the convergence rate proof for convex compact sets is completed in Theorem 4.12. An improved convergence rate of convex compact sets with twice continuously differentiable boundary is proved in Theorem 4.15.

In the original proof in [17], the quality of the approximation, i.e. the distance between the inner and outer approximation, is measured using two quality indicators, the Hausdorff metric (see Definition 4.3) and the symmetric difference volume metric.

**Definition 4.7** ( [17]). For $C_1,\ C_2 \in \mathscr{C}$ the *symmetric difference volume metric* is defined as $\delta^S(C_1, C_2) := \mu(C_1 \Delta C_2)$ with $\mu$ the Lebesgue measure and $C_1 \Delta C_2 = (C_1 \backslash C_2) \cup (C_2 \backslash C_1)$.

We first extend a result from [17] to metrics that are strongly equivalent to the Hausdorff metric. It uses a result estimating the volume of a pyramid defined by a sphere, an external point and a hyperplane to estimate the change in volume of the inner and outer approximation in one iteration of the Sandwiching algorithm.

**Lemma 4.8.** *Let $(P^n, Q^n)$ belong to a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) using a metric $\delta$ that is strongly equivalent to the Hausdorff metric with constants $c_1, c_2 > 0$. Let $C \in \mathscr{C}$, $u^* \in U(P^n)$ and $p^* \in T_C(u^*)$ the direction and point, chosen at an iteration of the algorithm, and $\zeta = \max\{\zeta_1, \zeta_2\}$ where*

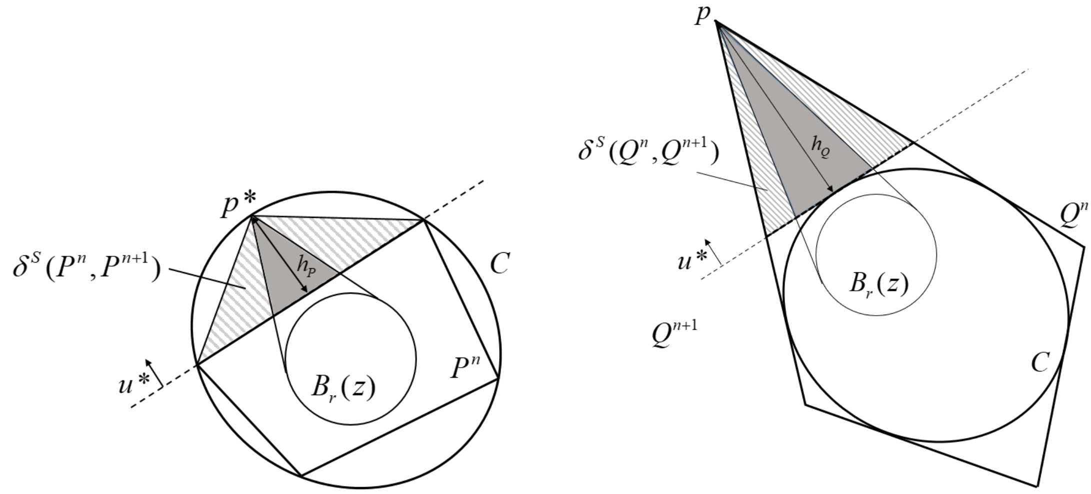
$$\zeta_1 := \inf\left\{\frac{||p^* - z||}{r} :\ z \in C,\ B_r(z) \subset P^n\right\},$$
$$\zeta_2 := \inf\left\{\frac{||p - z||}{r} :\ z \in C,\ B_r(z) \subset C,\ p \in T_{Q^n}(u^*)\right\}.$$

*Let $\delta$ be a metric that is strongly equivalent to the Hausdorff metric. Then*

$$\begin{aligned}
&\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \\
&\quad \geq \frac{\pi_{d-1}}{2^{d-1}d}(\zeta^2 - 1)^{(1-d)/2} \left(\alpha(P^n)\right)^{-d} \left(\delta^H(P^n, Q^n)\right)^d \\
&\quad \geq \frac{\pi_{d-1}}{2^{d-1}d}(\zeta^2 - 1)^{(1-d)/2} \left(\frac{c_1}{\alpha(P^n)}\right)^d \left(\delta(P^n, Q^n)\right)^d.
\end{aligned}$$

*Additionally, it holds*

An illustration of $\zeta_1$ and $\zeta_2$ is given in Figure 1.




(a) $||p^* - z||/r$ is an upper bound on $\zeta_1$ (b) $||p - z||/r$ is an upper bound on $\zeta_2$

Figure 1: Illustration of definitions of Lemma 4.8 and Lemma 2 of [17]

*Proof.* From Lemma 2 of [17] we obtain for $(P^n, Q^n)$ elements of a sequence of pairs of polyhedra, generated by the Sandwiching algorithm with support function selection (Algorithm 2),

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \geq \frac{\pi_{d-1}}{2^{d-1}d}(\zeta^2 - 1)^{(1-d)/2}[\delta_{P^n}(P^n, Q^n)]^d.$$

Note that in the original version of Lemma 2 in [17], $\zeta_2$ is defined with $B_r(z) \subset P^n$. We believe this to be a mistake and will use $B_r(z) \subset C$ in the following. Since this result is an important step in the convergence proof and it is only available in Russian, we state the proof in the Appendix. Note that the proof given in [17] for the Sandwiching algorithm with support function selection also holds for variable selection since the proof does not depend on the selection process. In particular, $g_{Q^{n+1}}(u^*) = \langle u^*, p^* \rangle = g_{P^{n+1}}(u^*)$ holds for every selection criterion.

Equation 8 contains the relation

$$\frac{c_1}{\alpha(P^n)}\delta(P^n, Q^n) \leq \frac{1}{\alpha(P^n)}\delta^H(P^n, Q^n) \leq \delta_{P^n}(P^n, Q^n).$$

We want to replace $\alpha(P^n)$ by a term that is independent of the current iteration. The asphericity of $P^n$ is defined as the minimal ratio of the radii of concentric outer and inner spheres, we denote $\alpha(P^n) = r_{\text{outer}}(P^n)/r_{\text{inner}}(P^n)$. Using the relations $r_{\text{outer}}(C) \geq r_{\text{outer}}(P^n)$ and $r_{\text{inner}}(P^n) \geq r_{\text{inner}}(P^0)$ we can bound the asphericity of $P^n$ from above by $\alpha(P^n) \leq r_{\text{outer}}(C)/r_{\text{inner}}(P^0)$. The statement then follows with Equation 8. □

Using Lemma 4 of [17], a result on monotonically decreasing sequences stated in the Appendix as Lemma A.1, a bound on the distance between inner and outer approximation in terms of the iteration number is derived.

**Lemma 4.9.** *Let $\{(P^n, Q^n)\}_{n=0,1,...}$ be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}$. Let $\delta$ be the metric equivalent to the Hausdorff metric with constants $c_1, c_2 > 0$ that is used as a selection criterion in the Sandwiching algorithm. Let $z \in C$ and $r, R > 0$ such that $B_r(z) \subset P^0 \subset Q^0 \subset B_R(z)$. Then for $n \geq d+1$ it holds*

$$\delta^S(P^n, Q^n) \leq (\lambda_a n^{1/(d-1)})^{-1}; \delta^H(P^n, Q^n) \leq (\lambda_b n^{1/(d-1)})^{-1}; \delta(P^n, Q^n) \leq (c_1 \lambda_b n^{1/(d-1)})^{-1}.$$

*where the quantities $\lambda_a$ and $\lambda_b$ depend only on $d$, $r$ and $R$ and $c_1$ only depends on the metric $\delta$.*

*Proof.* The proof is similar to that of Lemma 5 in [17] and given in the Appendix. □

Using Lemma 4.9, the convergence of the Sandwiching algorithm with strongly equivalent selection can be proven.

**Theorem 4.10.** *Let $\{(P^n, Q^n)\}_{n=0,1,...}$ be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) with a metric $\delta$ that is strongly equivalent to the Hausdorff metric for $C \in \mathscr{C}$. Then, for $\tilde{\delta} = \delta^S, \delta^H$ or $\delta$ it holds $\lim_{n\to\infty} \tilde{\delta}(P^n, C) = \lim_{n\to\infty} \tilde{\delta}(Q^n, C) = 0$.*

*Proof.* Since $\tilde{\delta}(P^n, Q^n) \geq \max\{\tilde{\delta}(P^n, C), \tilde{\delta}(C, Q^n\}$, for $\tilde{\delta} = \delta^S, \delta^H, \delta$, the assertions of the theorem follow directly from Lemma 4.9. □

### Convergence rate of the Sandwiching algorithm for convex compact sets

After showing that the Sandwiching algorithm converges, we will now investigate the convergence rate, i.e. how fast the approximation error decreases in the course of the algorithm. We derive two different convergence rates depending on the regularity of the convex compact set. First, we show the result for a general convex compact set.

To do this, we estimate the volume difference between the inner and outer approximation from one iteration to the other in terms of the Hausdorff distance between the inner and outer approximation.

**Lemma 4.11.** *Let $\{(P^n, Q^n)\}_{n=0,1,...}$ be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}$, $\alpha(C) \neq 1$. Let $\delta$ be the metric used as a selection criterion in the Sandwiching algorithm* *that is strongly equivalent to the Hausdorff metric with constants* $c_1, c_2 > 0$*. Then for any* $\epsilon$ *with* $1 > \epsilon > 0$*, there exists a number* $n_0$ *so that for all* $n \geq n_0$ *it holds*

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \geq \xi_1(\epsilon)[\delta^H(P^n, Q^n)]^d,$$

$$\textit{with}\ \ \xi_1(\epsilon) = (1-\epsilon)\frac{\pi_{d-1}}{2^{d-1}d}\left(\alpha(C)^2 - 1\right)^{(1-d)/2}\alpha(C)^{-2d}\left(\frac{c_1}{c_2}\right)^d.$$

*Proof.* The proof is similar to that of Lemma 6 in [17] and given in the Appendix. □

We can now complete the proof of the convergence rate for convex compact sets.

**Theorem 4.12.** *Let* $\{(P^n, Q^n)\}_{n=0,1,\ldots}$ *be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for* $C \in \mathscr{C}$ *with* $\alpha(C) \neq 1$*. Let* $\delta$ *be the metric strongly equivalent to the Hausdorff metric with constants* $c_1, c_2 > 0$ *that is used as a selection criterion in the algorithm. Then for any* $\epsilon > 0$ *there exists a number* $n_0$ *such that for* $n \geq n_0$ *it holds*

$$\delta^S(P^n, Q^n) \leq (1+\epsilon)\left(b_1 K(n)^{1/(d-1)}\right)^{-1}$$
$$\delta^H(P^n, Q^n) \leq (1+\epsilon)\left(b_2 K(n)^{1/(d-1)}\right)^{-1}$$
$$\delta(P^n, Q^n) \leq (1+\epsilon)\left(c_1 b_2 K(n)^{1/(d-1)}\right)^{-1}$$

*where*

$$b_1 = \frac{1}{2}\left((d-1)\frac{\pi_{d-1}}{d}\frac{1}{\sigma(C)^d}\right)^{1/(d-1)}(\alpha(C)^2-1)^{-1/2}\alpha(C)^{2d/(1-d)}\left(\frac{c_1}{c_2}\right)^{d/(d-1)},$$
$$b_2 = \frac{1}{2}\left(\frac{d-1}{d}\frac{\pi_{d-1}}{d}\frac{1}{\sigma(C)}\right)^{1/(d-1)}(\alpha(C)^2-1)^{-1/2}\alpha(C)^{2d/(1-d)}\left(\frac{c_1}{c_2}\right)^{d/(d-1)}$$

*and* $K(n)$ *can stand for the number of iterations* $n$*, the number of vertices of the inner approximation* $n^v(P^n)$ *or the number of facets of the outer approximation* $n^f(Q^n)$*.*

*Proof.* The proof is similar to that of Theorem 2 of [17] and given in the Appendix. □

*Remark* 2. The assumption $\alpha(C) \neq 1$ excludes sets with an asphericity of one, i.e. d-dimensional spheres. These sets belong to the class $\mathscr{C}^2$ for which a stronger result will be shown in Theorem 4.15.

### Convergence rate of the Sandwiching algorithm for convex compact sets with twice continuously differentiable boundary

After showing the convergence rate for general convex compact sets, the result is improved for convex compact sets with twice continuously differentiable boundary. When the set $C$ to be approximated is in $\mathscr{C}^2$, for every point on the boundary of $C$, a sphere can be placed inside of $C$ that touches this point. Then, Lemma 4.13 can be applied which adds a factor $\delta^H(P^n, Q^n)$ to the bound that ultimately leads to the improved convergence rate.

**Lemma 4.13** (Blaschke's rolling theorem (as stated in [20])).
*Let $C \in \mathscr{C}^2$ and $x \in \partial C$. Then, for any $r \leq \rho_{\min}(C)$ there exists $z \in C$ such that $B_r(z) \subset C$ and $x \in B_r(z)$. Moreover, for any $r \geq \rho_{\max}(C)$, there exists a $z \in \mathbb{R}^d$ such that $C \subset B_r(z)$ and $x \in B_r(z)$.*

**Lemma 4.14.** *Let $\{P^n, Q^n\}_{n=0,1,\ldots}$ be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}^2$. Let $\delta$ be the metric strongly equivalent to the Hausdorff metric with constants $c_1, c_2 > 0$ that is used as a selection criterion in the algorithm. Then for any $\epsilon$, $1 > \epsilon > 0$, there exists a value $\gamma(\epsilon)$, $1 > \gamma(\epsilon) > 0$, such that for $\delta^H(P^n, Q^n) \leq \gamma(\epsilon)\rho_{\min}(C)$ (with the minimal radius of curvature $\rho_{\min}(C)$, see Definition 3.2) it holds*

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \geq \xi_2(\epsilon, \rho_{\min}(C))[\delta^H(P^n, Q^n)]^{(d+1)/2}$$

$$\text{where } \xi_2(\epsilon, \rho_{\min}) = (1-\epsilon)\frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{\rho_{\min}}{2}\right)^{(d-1)/2}\left(\frac{c_1 r_{inner}(P^0)}{c_2 r_{outer}(C)}\right)^d.$$

*Proof.* The proof is similar to that of Lemma 7 in [17] and given in the Appendix. □

**Theorem 4.15.** *Let $\{(P^n, Q^n)\}_{n=0,1,\ldots}$ be a sequence of pairs of polyhedra, generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}^2$. Let $\delta$ be the metric strongly equivalent to the Hausdorff metric with constants $c_1, c_2 > 0$ that is used as the selection criterion in the algorithm. Then for any $\epsilon > 0$ there exists a number $n_0$ such that for $n \geq n_0$ it holds*

$$\delta^S(P^n, Q^n) \leq (1+\epsilon)\left(b_3 K(n)^{2/(d-1)}\right)^{-1}$$

$$\delta^H(P^n, Q^n) \leq (1+\epsilon)\left(b_4 K(n)^{2/(d-1)}\right)^{-1}$$

$$\delta(P^n, Q^n) \leq (1+\epsilon)\left(c_1 b_4 K(n)^{2/(d-1)}\right)^{-1}$$

$$\text{where } b_3 = \left(\frac{d-1}{2}\frac{\pi_{d-1}}{d}\left(\frac{c_1 r_{inner}(P^0)}{c_2 r_{outer}(C)}\right)^d \frac{1}{\sigma(C)^{(d+1)/2}}\right)^{2/(d-1)} \frac{\rho_{\min}(C)}{8},$$

$$b_4 = \left(\frac{d-1}{d+1}\frac{\pi_{d-1}}{d}\left(\frac{c_1 r_{inner}(P^0)}{c_2 r_{outer}(C)}\right)^d \frac{1}{\sigma(C)}\right)^{2/(d-1)} \frac{\rho_{\min}(C)}{8}$$

*and $K(n)$ can stand for the number of iterations $n$, the number of vertices of the inner approximation $n^v(P^n)$ or the number of facets of the outer approximation $n^f(Q^n)$.*

*Proof.* The proof is similar to that of Theorem 3 of [17] and is given in the Appendix. □

# 5. Applying the convergence results to the Sandwiching algorithm from multiobjective optimization

We now want to apply the convergence results developed in Section 4 for Algorithm 3 that approximation convex compact sets to the Sandwiching algorithm (Algorithm 1) for the approximation of certain nondominated sets in multiobjective optimization.

We first compare Algorithm 3 to the Sandwiching Algorithm 1 for the approximation of convex bounded nondominated sets.

The Sandwiching Algorithm 1 will always select a facet of the inner approximation as a new weight for the weighted sum scalarization problem if the solution of the quality calculation is nondegenerate (see Lemmata 2.3.1 and 2.3.2 of [22]). In the degenerate case, the normal of a lower-dimensional facet may also be an optimal dual solution. In this case, the normal of a facet can still be obtained using strategies outlined in Section 2.3 of [22].

The new nondominated point computed in an iteration of our Sandwiching algorithm fulfils $p^* \in T_C(u^*)$ (Lemma 2.3.3 of [22]).

Finally, the update process of inner and outer approximation is the same in Algorithms 1 and 3.

## 5.1. Applying results for convex compact sets to multiobjective optimization

We demonstrated that the Algorithms 3 and 1 are comparable. To apply the convergence rate results from Section 4 it therefore remains to show which requirements on the multiobjective optimization problem are necessary to fulfil the assumptions of the theorems. In particular, under which requirements does the nondominated set have a twice continuously differentiable boundary?

**Lemma 5.1** (Chapter 4.1 of [37]). *In a convex multiobjective optimization problem with $\mathcal{X} \neq \emptyset$, the set $cl(\mathcal{Y} + \mathcal{C})$ is convex and closed.*

By assuming that the multiobjective optimization problem is bounded and defining the set to be approximated (see Equation 3) as

$$\mathcal{P} := (\mathcal{Y} + \mathcal{C}) \cap (\tilde{y}^N - \mathbb{R}^d_{\geq}),$$

we obtain that $\mathcal{P}$ is a convex compact set.

Under the additional condition that the extreme compromise solutions must not coincide, $\mathcal{P}$ has nonempty interior.

Using the following lemma, we can also note that the nondominated set of such a multiobjective optimization problem will be connected.

**Definition 5.2** (Definition 3.31 of [10]). A set $\mathcal{S} \subset \mathbb{R}^d$ is called *not connected* if it can be written as $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ with $\mathcal{S}_1, \mathcal{S}_2 \neq \emptyset$, $\text{cl}\mathcal{S}_1 \cap \mathcal{S}_2 = \mathcal{S}_1 \cap \text{cl}\mathcal{S}_2 = \emptyset$.
Otherwise, $\mathcal{S}$ is called *connected*.

**Lemma 5.3** (Theorem 3.35 of [10]). *If $\mathcal{Y}$ is closed, convex and $\mathbb{R}^d_{\geq}$-compact then $\mathcal{Y}_N$ is connected.*

In Section 6.3 of [28], the application of algorithms designed for the approximation of convex compact sets to the approximation of bounded convex nondominated sets is discussed. In the context of augmentation methods (which are inner approximation methods for convex compact sets), it is noted that every convergent augmentation scheme can be directly applied to the approximation of bounded convex nondominated sets if the initial approximation is given by the cone $p + \mathbb{R}^d_{\geq}$ for a nondominated point $p$.

Note that this assumption on the initial approximation leads to a full-dimensional (i.e. $d-1$-dimensional) approximation of the nondominated set even if the nondominated set itself is of a lower dimension. The Sandwiching algorithm, both the variants for sets (Algorithm 2/3) and for multiobjective optimization (Algorithm 1) can in principle also be applied to lower-dimensional sets. In this article, however, we assume that the nondominated set is of full dimension. Requirements for a full-dimensional nondominated set have been derived in [15].

In summary, we obtain the following result.

**Theorem 5.4.** *The Sandwiching algorithm (Algorithm 1) using a quality criterion that is strongly equivalent to the Hausdorff metric to approximate the full-dimensional nondominated set of a convex bounded multiobjective optimization problem in which the extreme compromises do not coincide, converges (see Theorem 4.10) and the convergence rate result of Theorem 4.12 applies.*

Note that this result holds for every pointed, convex domination cone $\mathcal{C} \supset \mathbb{R}^d_{\geq}$.

### Requirements for a nondominated set with twice continuously differentiable boundary

In Theorem 4.15 we showed an improved convergence rate for a Sandwiching approximation of convex compact sets with twice continuously differentiable boundary (see Definition 3.2). We investigate conditions under which the set to be approximated $\mathcal{P}$ of a multiobjective optimization problem has twice continuously differentiable boundary.

Let $\min f(x)$ s.t. $x \in \mathcal{X}$ be a multiobjective optimization problem with convex objective functions $f : \mathbb{R}^k \to \mathbb{R}^d$ and a convex decision space

$$X := \{x \in \mathbb{R}^k : \ g(x) \leq 0, \ h(x) = 0\}$$

defined using convex functions $g : \mathbb{R}^k \to \mathbb{R}^{|I|}$ and linear functions $h : \mathbb{R}^k \to \mathbb{R}^{|J|}$ with $|I| < \infty$ and $|J| < \infty$. Let $f$, $g$, $h$ be three times continuously differentiable.

We define $w : \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}$. $w(x, \gamma) := \sum_{i=1,\dots,d} \gamma_i f_i(x)$ and consider the weighted sum scalarization problem (see Section 2.1) as a parametric optimization problem

$$\min_{x \in \mathcal{X}} w(x, \gamma). \tag{WS($\gamma$)}$$

Every nondominated point of the convex multiobjective optimization problem is the solution of an instance of (WS($\gamma$)) for some $\gamma \in \mathcal{C}^* \backslash \{0\}$ (Theorem 1.24 of [36]). Since the $f_i$ are three times continuously differentiable, $w(x, \gamma)$ is also three times continuously differentiable in $x$.

Our aim is now to show that the function mapping a weighted sum parameter to a weakly nondominated point is twice continuously differentiable. To prepare for the result, we state some definitions from nonlinear optimization. These can be found e.g. in [4], we follow the notation of [35].

**Definition 5.5.** A feasible point $x^* \in \mathcal{X}$ is called stationary point for problem (WS($\gamma$)) if there exist $\lambda^* \in \mathbb{R}^{|I|}$ and $\mu^* \in \mathbb{R}^{|J|}$ with

$$\begin{aligned}
&D_x w(x^*, \gamma) + \sum_{i \in I} \lambda_i D g_i(x^*) + \sum_{j \in J} \mu_j D h_j(x^*) = 0, \\
&\lambda_i \geq 0 \text{ for all } i \in I, \\
&\lambda_i g_i(x^*) = 0 \text{ for all } i \in I.
\end{aligned}$$

The triple $(x^*, \lambda^*, \mu^*)$ is called *Karush-Kuhn-Tucker point* (short: KKT point). The *Lagrange function* $\mathcal{L} : \mathbb{R}^k \times \mathbb{R}^{|I|} \times \mathbb{R}^{|J|} \to \mathbb{R}$ is defined as

$$\mathcal{L}(x, \lambda, \mu) := w(x, \gamma) + \sum_{i \in I} \lambda_i g_i(x) + \sum_{j \in J} \mu_j h_j(x).$$

If, for every $i \in I$, exactly one of the values $g_i(x^*)$ and $\lambda_i$ equals zero, one says that *strict complementarity slackness* holds.

**Definition 5.6.** A feasible point $x \in \mathcal{X}$ satisfies the *Linear Independence Constraint Qualification* (short: LICQ) for problem (WS($\gamma$)), if for the set of active indices $I_0(x) := \{i \in I : g_i(x) = 0\}$ the vectors $Dg_i(x)$, $i \in I_0(x)$ and $Dh_j(x)$, $j \in J$ are linearly independent.

**Definition 5.7.** A KKT point $(x^*, \lambda^*, \mu^*)$ is said to satisfy the *Second-Order Sufficient Condition* (short: SOSC), if $d^T D_1^2 \mathcal{L}(x^*, \lambda^*, \mu^*) d > 0$ for every $d \in T(x^*)$, $d \neq 0$,

$$\text{where } \; T(x^*) := \left\{ d \in \mathbb{R}^k : \begin{array}{l} Dg_i(x^*)d \leq 0 \text{ for } i \in I_0(x^*) \text{ with } \lambda_i = 0, \\ Dg_i(x^*)d = 0 \text{ for } i \in I_0(x^*) \text{ with } \lambda_i > 0, \\ Dh_j(x^*)d = 0 \text{ for } j \in J \end{array} \right\}.$$

Applying the implicit function theorem for differentiable functions, we can obtain the following result.

**Lemma 5.8** (Modification of Theorem 2.13, together with Theorem 2.5 of [35])**.** *For a given parameter vector $\gamma^* \in \mathcal{C}^* \backslash \{0\}$ let $(x^*, \lambda^*, \mu^*)$ be a KKT point of WS($\gamma^*$). Moreover, assume that for WS($\gamma^*$) the following is true:*

- *at $x^*$ LICQ is satisfied,*
- *$x^*, \lambda^*$ fulfil the strict complementary slackness condition*
- *$x^*, \lambda^*, \mu^*$ satisfy the second-order sufficient condition SOSC.*

*Then there are $\delta > 0$, $\epsilon > 0$ and a twice continuously differentiable function $\nu = (x, \lambda, \mu) : B_\delta(\gamma^*) \to \mathbb{R}^d \times \mathbb{R}^{|I|} \times \mathbb{R}^{|J|}$, such that, for every $\gamma \in B_\delta(\gamma^*)$ the point $x(\gamma)$ is a local minimum of* (WS($\gamma$)) *which is unique in $B_\epsilon(x^*)$. Since* (WS($\gamma$)) *is convex, $x(\gamma)$ is also the global minimum of* (WS($\gamma$)).

*Remark* 3. If the objective functions $f$ are strictly convex, then SOSC is fulfilled. Otherwise, SOSC has to be verified for the Lagrange function of (WS($\gamma$)). Strict complementarity does not hold when the set of active constraints changes, i.e. in vertices of the feasible set.

What remains to discuss is whether the conditions of Theorem 4.15 are met, i.e. if the set that is approximated by the algorithm has a twice continuously differentiable boundary.

The Sandwiching algorithm approximates the set (see Equation 3) $\mathcal{P} = (\mathcal{Y} + \mathcal{C})_N \cap (\tilde{y}^N - \mathbb{R}^d_{\geq})$. Obviously, the set has a non-differentiable boundary in $\tilde{y}^N$ and is potentially non-differentiable on the relative boundary of the nondominated set. However, the $(\tilde{y}^N - \mathbb{R}^d_{\geq})$ part of $\mathcal{P}$ is already fully approximated by the initial approximation containing the extreme compromise solutions and the nadir point approximation $\tilde{y}^N$. In the following iterations of the Sandwiching approximation, the set $\mathcal{P}$ is approximated by only computing nondominated solutions. Thus, in practice the convergence behaviour of the Sandwiching algorithm is not impaired by this non-differentiability.

If a twice continuously differentiable boundary is needed for theoretical studies, the edges can be "rounded" by adding a sphere to the non-differentiable parts. Since the radius of the sphere is known, points approximating this part of $\mathcal{P}$ can be computed without solving an optimization problem.

It remains to show that the part of the boundary of $\mathcal{P}$ that is formed by the nondominated set is twice continuously differentiable (see Definition 3.2).

**Lemma 5.9.** *Let the function $x$ be defined as in Lemma 5.8. Let us assume that $f(x(\gamma))$ is injective for $\gamma \in B_\delta(\gamma^*)$. For every $\gamma \in B_\delta(\gamma^*)$ for the open ball $B_\delta(\gamma^*)$, the point $f(x(\gamma))$ is a twice continuously differentiable boundary point of $G := f(x(B_\delta(\gamma^*))) + \mathcal{C}$ (with the domination cone $\mathcal{C}$), i.e. there exists an open environment $U = U(f(x(\gamma)))$ and a twice continuously differentiable function $\rho : U \to \mathbb{R}$ that fulfils $U \cap G = \{x \in U : \rho(x) < 0\}$ and $\frac{d}{dx}\rho(x) \neq 0 \ \forall x \in U$.*

*Proof.* We define the function $p : B_\delta(\gamma^*) \times \mathbb{R} \to \mathbb{R}^d$ as $p(\gamma, t) := f(x(\gamma)) - te$ where $e$ denotes the vector of ones. Since every solution $x(\gamma)$ is a weakly efficient point, there cannot exist $\gamma_1$ and $\gamma_2$ and a $t > 0$ with $f(x(\gamma_1)) = f(x(\gamma_2)) - te$ or $x(\gamma_1)$ would not be efficient. Therefore, $p$ is injective if the function $f(x(\gamma))$ is injective. Then, $p$ is bijective on its image.

Therefore, there exists a twice continuously differentiable inverse function $p^{-1} : p(B_\delta(\gamma^*) \times \mathbb{R}) \to (B_\delta(\gamma^*)) \times \mathbb{R}$ of $p$. When projecting the image of $p^{-1}$ to the $t$ component, we obtain a function with the required properties: its value is zero on the boundary of $G$, i.e. on the nondominated set, smaller than zero in the interior of $G$, i.e. the dominated region and greater than zero outside of $G$, i.e. in the unattainable region. Since it is linear in $t$ with coefficient $e$, its derivative is never zero. □

In summary, we obtain the following result.

**Theorem 5.10.** *Consider the Sandwiching algorithm for convex bounded multiobjective optimization (Algorithm 1) using a quality criterion that is strongly equivalent to the Hausdorff metric. Let the multiobjective optimization problem fulfil*

- *the multiobjective optimization problem is convex and bounded, the extreme compromises do not coincide and the nondominated set is of full dimension,*
- *the objective functions $f$ and the functions $g$, $h$ defining the feasible set are three times continuously differentiable,*
- *for the KKT points $(x^*, \lambda^*, \mu^*)$ of the weighted sum problem for every weight $\gamma \in \mathcal{C}^* \setminus \{0\}$ the strict complementarity slackness condition, LICQ and SOSC are fulfilled,*
- *the function mapping a weight $\gamma$ to a weakly nondominated point is injective.*

*Then the algorithm converges (see Theorem 4.10), the set $\mathcal{P}$ has twice continuously differentiable boundary and therefore the improved convergence rate result of Theorem 4.15 applies.*

Until now, the quality indicator was not specified. To apply the convergence results, it is necessary that the quality criterion is a metric that is strongly equivalent to the Hausdorff metric. In the next two subsections, we will show for two quality indicators, the polyhedral gauge and the epsilon indicator, that they are strongly equivalent to the Hausdorff metric and determine the corresponding equivalence constants. Then, we can formulate the convergence rate results for these specific Sandwiching algorithms.

## 5.2. Convergence rate of Sandwiching with polyhedral gauge quality

To assess the approximation quality of the Sandwiching algorithm, we measure the distance between the inner and outer approximation. A way to measure the approximation quality using the polyhedral gauge introduced in Section 2.2.2 is the value $|\gamma_{I_R^n}(\bar{q}) - 1|$ with $\bar{q}$ the solution of $\max\{\gamma_{I_R^n}(q) \text{ s.t. } q \in O_R^n\}$. We will also denote $\gamma_{I_R^n}(O_R^n) := \gamma_{I_R^n}(\bar{q})$. The general idea of this distance measure is to determine the factor by which the inner approximation would need to be scaled to fully contain the outer approximation.

In the following, we will show that this distance measure is strongly equivalent to the Hausdorff metric. Then, we formulate the convergence rate results for the Sandwiching algorithm that uses the polyhedral gauge to determine the approximation quality and to select the location of the next approximation point.

Since we are considering the Sandwiching algorithm for the approximation of convex bounded nondominated sets (in comparison to the algorithm for convex compact sets of Sections 3 and 4), we denote the inner approximation by $I$, and the outer approximation by $O$.

### Strong equivalence to the Hausdorff metric

To discuss the polyhedral gauge, we denote the reflection set of the inner approximation (see Section 2.2.2) by $I_R^n := R(I^n)$.

The reflected inner approximation $I_R^n$ is a polytope in $\mathbb{R}^d$ that contains the origin in its interior and is symmetric with respect to the origin. From Lemma 6 of [34] we obtain that the polyhedral gauge with unit ball $I_R^n$ is a norm in $\mathbb{R}^d$. Since in a finite-dimensional vector space all norms are equivalent (see Theorem 1.18 of [14]), the block norm $\gamma$ is in particular equivalent to the Euclidean norm $||\cdot||_2$. Let us denote the equivalence constants by $\tilde{c_1}$ and $\tilde{c_2}$.

**Lemma 5.11.** *For $I_R^n$ and $O_R^n$ the reflection sets of inner and outer approximations generated by the Sandwiching algorithm (Algorithm 1), it holds*

$$\tilde{c_1}|\gamma_{I_R^n}(\bar{q}) - 1| \leq \delta^H(I_R^n, O_R^n) \leq \tilde{c_2}|\gamma_{I_R^n}(\bar{q}) - 1|$$

*with the norm equivalence constants $\tilde{c_1}$ and $\tilde{c_2}$ of the polyhedral gauge and the Euclidean norm, $\tilde{c_1}\gamma_{I_R^n}(x) \leq ||x||_2 \leq \tilde{c_2}\gamma_{I_R^n}(x)$.*

*Proof.* The first bound is obtained using norm equivalence and the reversed triangle inequality and that $\sup_{p\in I_R^n} \inf_{q\in O_R^n} ||q-p|| = 0$ due to $I_R^n \subset O_R^n$,

$$\begin{aligned}\delta^H(I_R^n, O_R^n) &= \max\{\sup_{p\in I_R^n} \inf_{q\in O_R^n} ||q-p||, \sup_{q\in O_R^n} \inf_{p\in I_R^n} ||q-p||\} = \sup_{q\in O_R^n} \inf_{p\in I_R^n} ||q-p|| \\ &\geq \tilde{c_1} \sup_{q\in O_R^n} \inf_{p\in I_R^n} \gamma_{I_R^n}(q-p) \geq \tilde{c_1} \sup_{q\in O_R^n} \inf_{p\in I_R^n} |\gamma_{I_R^n}(q) - \gamma_{I_R^n}(p)| \\ &= \tilde{c_1} \sup_{q\in O_R^n} |\gamma_{I_R^n}(q) - 1| = \tilde{c_1}|\gamma_{I_R^n}(\bar{q}) - 1|.\end{aligned}$$

For the second bound we can derive with the norm equivalence

$$\delta^H(I_R^n, O_R^n) = \sup_{q\in O_R^n} \inf_{p\in I_R^n} ||q-p|| \leq \tilde{c_2} \sup_{q\in O_R^n} \inf_{p\in I_R^n} \gamma_{I_R^n}(q-p). \tag{9}$$

Now, using that $\gamma_{I_R^n}(q) \geq 1\ \forall q \in O_R^n$ because $I_R^n \subset O_R^n$, that $\gamma_{I_R^n}(p) = 1\ \forall p \in I_R^n$ and that $\bar{q}$ solves $\max \gamma_{I_R^n}(q)$ s.t. $q \in O_R^n$, we obtain

$$|\gamma_{I_R^n}(\bar{q}) - 1| = \gamma_{I_R^n}(\bar{q}) - 1 = \max_{q\in O_R^n} \gamma_{I_R^n}(q) - 1 = \max_{q\in O_R^n} \min_{p\in I_R^n} \left(\gamma_{I_R^n}(q) - \gamma_{I_R^n}(p)\right). \tag{10}$$

To obtain the bound, it remains to show that

$$\max_{q\in O_R^n} \min_{p\in I_R^n} \gamma_{I_R^n}(q-p) = \max_{q\in O_R^n} \min_{p\in I_R^n} \left(\gamma_{I_R^n}(q) - \gamma_{I_R^n}(p)\right). \tag{11}$$

To show Equation 11, we take a closer look at the left side of the equation. In particular, for a given $\tilde{q} \in O_R^n$, we determine the value of $\min_{p\in I_R^n} \gamma_{I_R^n}(\tilde{q}-p)$.

First, note that the supremum over $O_R^n$ will always be attained on the boundary, i.e. by some $q \in \partial O_R^n$ and that the infimum of the difference to $O_R^n$ over $I_R^n$ will also be attained by some $p \in \partial I_R^n$. Since $0 \in I_R^n \subset O_R^n$, for every $\tilde{q} \in \partial O_R^n$ there is a $\tilde{p} \in \partial I_R^n$ that fulfils $\tilde{q} = \lambda\tilde{p}$ for some $\lambda \geq 1$. We write $\lambda = \lambda(\tilde{q})$. Since for every $p \in \partial I_R^n$ it holds $\gamma_{I_R^n}(p) = 1$, the difference between $\tilde{q}$ and $\tilde{p}$ in $\gamma_{I_R^n}$ is

$$\gamma_{I_R^n}(\tilde{q} - \tilde{p}) = \gamma_{I_R^n}(\lambda(\tilde{q})\tilde{p} - \tilde{p}) = (\lambda(\tilde{q}) - 1)\gamma_{I_R^n}(\tilde{p}) = \lambda(\tilde{q}) - 1.$$

In comparison to this, examine the difference between $\tilde{q}$ and an arbitrary point $p \in \partial I_R^n$. It holds, $\gamma_{I_R^n}(\tilde{q}) = \gamma_{I_R^n}(\lambda(\tilde{q})\tilde{p}) = \lambda(\tilde{q})\gamma_{I_R^n}(\tilde{p}) = \lambda(\tilde{q})$.and using the reversed triangle inequality (Remark 1.5 of [14]),

$$\gamma_{I_R^n}(\tilde{q} - p) \geq |\gamma_{I_R^n}(\tilde{q}) - \gamma_{I_R^n}(p)| = |\lambda(\tilde{q}) - 1| = \lambda(\tilde{q}) - 1.$$

So, $\gamma_{I_R^n}(\tilde{q} - p)$ attains its minimum at $\tilde{p}$ defined by $\tilde{q} = \lambda(\tilde{q})\tilde{p}$ for some $\lambda(\tilde{q}) \geq 1$.

Using this result, we can express the left side of Equation 11 as

$$\max_{q \in O_R^n} \min_{p \in I_R^n} \gamma_{I_R^n}(q - p) = \max_{q \in O_R^n}(\lambda(q) - 1)$$

with $\lambda(q) \geq 1$ such that $q = \lambda(q)p$ for some $p \in I_R^n$. Following the same argument with the right side of Equation 11 yields

$$\max_{q \in O_R^n} \min_{p \in I_R^n} \left(\gamma_{I_R^n}(q) - \gamma_{I_R^n}(p)\right) = \max_{q \in O_R^n} \gamma_{I_R^n}(q) - 1 = \left(\max_{q \in O_R^n} \lambda(q)\right) - 1.$$

Combining Equations 9, 10 and 11, we obtain the second bound. □

Next, we determine the values of $\tilde{c_1}$ and $\tilde{c_2}$.

**Lemma 5.12.** *The constants*

$$r_{inner} = \min_{\gamma_{I_R^n}(u)=1} ||u||, \quad r_{outer} = \max_{\gamma_{I_R^n}(u)=1} ||u||$$

*fulfil* $r_{inner}\gamma_{I_R^n}(x) \leq ||x||_2 \leq r_{outer}\gamma_{I_R^n}(x)\ \forall x \in \mathbb{R}^d$.

*The constant* $r_{inner}$ *is the largest radius of a ball inside of* $I_R^n$ *and centred in the origin. Analogously,* $r_{outer}$ *is the smallest radius of a ball that contains* $I_R^n$ *and that is also centred in the origin.*

*Proof.* For $x \in \mathbb{R}^d$, we define $u := x/\gamma_{I_R^n}(x)$. Then, it holds $||x||_2/\gamma_{I_R^n}(x) = ||\frac{x}{\gamma_{I_R^n}(x)}||_2 = ||u||_2$. Since $u$ fulfils $\gamma_{I_R^n}(u) = 1$, we obtain $||x||/\gamma_{I_R^n}(x) \geq r_{\text{inner}}$.

Analogously, we obtain $r_{\text{outer}} = \max_{\gamma_{I_R^n}(u)=1} ||u||$. □

## Convergence rate of Sandwiching with polyhedral gauge selection

We denote by $g_{I_R^n}(O^n) := |\gamma_{I_R^n}(O^n) - 1|$ the quality criterion based on the polyhedral gauge. In the previous subsection we have shown that the Hausdorff metric between inner and outer approximations created by the Sandwiching algorithm is strongly equivalent to this quality criterion, $r_{\text{inner}}\ g_{I_R^n}(O^n) \leq \delta^H(I_R^n, O_R^n) \leq r_{\text{outer}}\ g_{I_R^n}(O^n)$.

Therefore, the convergence results of Section 4 hold for the Sandwiching algorithm with polyhedral gauge selection. In particular, the algorithm converges (Theorems 4.10 and 5.4).

As one of the main results of this article, we can now show the convergence rate of the Sandwiching algorithm with polyhedral gauge quality.

**Theorem 5.13.** *Let $\{(I^n, O^n)\}_{n=0,1,\dots}$ be a sequence of inner and outer approximations generated by the Sandwiching algorithm (Algorithm 1) with polyhedral gauge selection for a convex bounded multiobjective optimization problems in which the extreme compromises do not coincide and which has a full-dimensional nondominated set. Then for any $\epsilon > 0$ there exists a number $n_0$ such that for $n \geq n_0$ it holds*

$$\delta^S(I^n, O^n) \leq (1+\epsilon)C_1 \; n^{-1/(d-1)}$$
$$\delta^H(I^n, O^n) \leq (1+\epsilon)C_2 \; n^{-1/(d-1)}$$
$$g_{I_R^n}(O^n) \leq (1+\epsilon)C_2 r_{inner}^{-1} \; n^{-1/(d-1)}$$

*with $C_1 := 1/b_1$, $C_2 := 1/b_2$ with $b_1, b_2$ defined in Theorem 4.12.*

*Proof.* Since the polyhedral gauge is strongly equivalent to the Hausdorff metric (Lemma 5.11), the result follows directly from Theorems 4.12 and 5.4. □

Under additional regularity assumptions, we obtain an improved convergence rate of the Sandwiching algorithm with polyhedral gauge quality.

**Theorem 5.14.** *Consider a convex bounded multiobjective optimization problem in which the extreme compromises do not coincide and which has a full-dimensional nondominated set. Let additionally the regularity assumptions of Theorem 5.10 hold, i.e. the objective functions and the functions forming the feasible set are three times continuously differentiable, the function mapping a weight $\gamma$ to a weakly nondominated point is locally injective, and for the KKT points of the weighted sum problem the strict complementary slackness condition, LICQ and SOSC are fulfilled. Let the nondominated set of this MOP be approximated by the Sandwiching algorithm (Algorithm 1) with polyhedral gauge quality.*

*Then for any $\epsilon > 0$ there exists a number $n_0$ such that for $n \geq n_0$ it holds*

$$\delta^S(I^n, O^n) \leq (1+\epsilon)C_3 \; n^{-2/(d-1)}$$
$$\delta^H(I^n, O^n) \leq (1+\epsilon)C_4 \; n^{-2/(d-1)}$$
$$g_{I_R^n}(O^n) \leq (1+\epsilon)C_4 r_{inner}^{-1} \; n^{-2/(d-1)}$$

*with $C_3 := 1/b_3$, $C_4 := 1/b_4$ with $b_3$, $b_4$ defined in Theorem 4.15.*

*Proof.* Since the polyhedral gauge is strongly equivalent to the Hausdorff metric (Lemma 5.11), the result follows directly from Theorems 4.15 and 5.10. □

## 5.3. Convergence rate of Sandwiching with epsilon indicator selection

The epsilon indicator has been introduced in Section 2.2.2. We first show that the epsilon indicator is strongly equivalent to the Hausdorff metric. Then, we formulate the convergence rate results for the Sandwiching algorithm that uses the epsilon indicator to determine the approximation quality and to select the location of the next approximation point.

#### Strong equivalence of the epsilon indicator to the Hausdorff metric

Both [25] and [5] state that the calculation of the $\epsilon$-indicator value is equivalent to the Hausdorff metric using a one-sided maximum metric

$$\epsilon(Z, A) = \delta_{\tilde{d}}^H(Z, A) = \max_{z \in Z} \min_{a \in A} \tilde{d}(a, z) \tag{12}$$

with $\tilde{d}(a, b) = \max_{i=1,\dots,d} \max\{b_i - a_i, 0\}$.

We want to select a new facet to be improved by measuring the distance between the outer and inner approximation using the epsilon indicator $\epsilon(I^n, O^n)$.

Lemma 1 of [7] states that the $\epsilon$-indicator $\epsilon(Z, A)$ of a closed, non-empty approximation $A \subseteq Z + \mathbb{R}^d_{\geq}$ of the closed set of attainable objective vectors $Z \subseteq \mathbb{R}^d$ can be represented as $\epsilon(Z, A) = \delta_\infty^H(Z, A + \mathbb{R}^d_{\geq}) := \sup_{z \in Z} \inf_{z' \in A + \mathbb{R}^d_{\geq}} ||z' - z||_\infty$.

Remember that due to the definition of the set to be approximated $\mathcal{P}$, adding the domination cone to the inner approximation does not change the approximation and therefore $\delta(O^n, I^n) = \delta(O^n, I^n + \mathbb{R}^d_{\geq})$ for every quality measure $\delta$. Thus, we obtain that the $\epsilon$-indicator between the inner and outer approximation calculated by the Sandwiching algorithm $\epsilon(I^n, O^n)$ is equivalent to the Hausdorff distance with maximum metric.

$$\epsilon(I^n, O^n) = \delta_{\tilde{d}}^H(I^n, O^n) = \delta_\infty^H(I^n, O^n + \mathbb{R}^d_{\geq}) = \delta_\infty^H(I^n, O^n). \tag{13}$$

**Lemma 5.15.** *The epsilon indicator is strongly equivalent to the Hausdorff metric with constants $c_1 = 1$ and $c_2 = \sqrt{d}$.*

*Proof.* The maximum norm is equivalent to the Euclidean norm with the constants (see [14]) $||x||_\infty \leq ||x||_2 \leq \sqrt{d}||x||_\infty, \quad x \in \mathbb{R}^d$. Using (13) we obtain $\delta_\infty^H(P, C) \leq \delta^H(P, C) \leq \sqrt{d}\ \delta_\infty^H(P, C)$ for $C \in \mathscr{C}$ and $P \in \mathscr{P}^i(C)$. □

#### Convergence rate of Sandwiching with epsilon indicator selection

In the previous section we have shown that the $\epsilon$-indicator between the inner and outer approximations of the Sandwiching algorithm is strongly equivalent to the Hausdorff metric. Therefore, the convergence results of Section 4 hold for the Sandwiching algorithm with epsilon indicator selection. In particular, the algorithm converges (Theorems 4.10 and 5.4).

As one of the main results of this article, we can now show the convergence rate of the Sandwiching algorithm with epsilon indicator quality.

**Theorem 5.16.** *Let $\{(I^n, O^n)\}_{n=0,1,\dots}$ be a sequence of inner and outer approximations generated by the Sandwiching algorithm (Algorithm 1) with epsilon indicator selection for a convex bounded multiobjective optimization problem in which the extreme compromises do not coincide and which has a full-dimensional nondominated set. Then for any $\epsilon > 0$*

*there exists a number $n_0$ such that for $n \geq n_0$ it holds*

$$\delta^S(I^n, O^n) \leq (1+\epsilon)C_1 \; n^{-1/(d-1)}$$
$$\delta^H(I^n, O^n) \leq (1+\epsilon)C_2 \; n^{-1/(d-1)}$$
$$\delta^H_\infty(I^n, O^n) \leq (1+\epsilon)C_2 \; n^{-1/(d-1)}$$

*with $C_1 := 1/b_1$, $C_2 := 1/b_2$ with $b_1, b_2$ defined in Theorem 4.12.*

*Proof.* Since the epsilon indicator is strongly equivalent to the Hausdorff metric (Lemma 5.15), the result follows directly from Theorems 4.12 and 5.4. □

Under additional regularity assumptions, we obtain an improved convergence rate of the Sandwiching algorithm with epsilon indicator quality.

**Theorem 5.17.** *Consider a convex bounded multiobjective optimization problem in which the extreme compromises do not coincide and which has a full-dimensional nondominated set. Let additionally the regularity assumptions of Theorem 5.10 hold, i.e. the objective functions and the functions forming the feasible set are three times continuously differentiable, the function mapping a weight $\gamma$ to a weakly nondominated point is locally injective, and for the KKT points of the weighted sum problem the strict complementary slackness condition, LICQ and SOSC are fulfilled. Let the nondominated set of this MOP be approximated by the Sandwiching algorithm (Algorithm 1) with epsilon indicator quality.*

*Then for any $\epsilon > 0$ there exists a number $n_0$ such that for $n \geq n_0$ it holds*

$$\delta^S(I^n, O^n) \leq (1+\epsilon)C_3 \; n^{-2/(d-1)}$$
$$\delta^H(I^n, O^n) \leq (1+\epsilon)C_4 \; n^{-2/(d-1)}$$
$$\delta^H_\infty(I^n, O^n) \leq (1+\epsilon)C_4 \; n^{-2/(d-1)}$$

*with $C_3 := 1/b_3$, $C_4 := 1/b_4$ with $b_3$, $b_4$ defined in Theorem 4.15.*

*Proof.* Since the epsilon indicator is strongly equivalent to the Hausdorff metric (Lemma 5.15), the result follows directly from Theorems 4.15 and 5.10. □

# 6. Conclusion and Discussion

In this article, we showed the convergence rate of a Sandwiching algorithm for the approximation of convex bounded nondominated sets using two different quality indicators, the epsilon indicator and polyhedral gauge. We achieved this by extending a convergence rate result published in 1996 in Russian [17] for the approximation of convex compact sets to a class of quality indicators and to the approximation of convex bounded nondominated sets in multiobjective optimization.

The convergence rate of the Sandwiching algorithm for variable selection criteria (Theorem 4.15) is optimal for the approximation of convex compact nondominated sets with twice continuously differentiable boundary, see Chapter 1.10.4 of [13].

For general convex compact nondominated sets, the convergence rate proved in this article is not optimal. While the convergence rate of the pure inner approximation is optimal (see [28], Theorems 8.29 and 8.30), the outer approximation is updated in a non-optimal way. This is intuitive, since the selection of the new approximation point in the Sandwiching algorithm (Algorithm 1) as the point whose tangential hyperplane is parallel to a facet of the inner approximation depends only on the inner approximation and therefore cannot be optimal with respect to the outer approximation.

If the inner and outer approximation are treated completely separately, then optimal strategies can be chosen for each and the overall convergence rate will be optimal, even in the general convex compact case. Two examples for these algorithms have been introduced in [18]. The cost of this improved convergence rate is the computation of a higher number of approximation points. While in our algorithm, an approximation point is used to improve both the inner and the outer approximation, in these algorithms approximation points are computed to improve only either the inner or the outer approximation.

Our Sandwiching algorithm (Algorithm 1) can be regarded as an inner approximation algorithm that also creates an outer approximation using tangent information that is automatically obtained when using the weighted sum scalarization to compute approximation points.

Our main priority in approximating a nondominated set is to obtain a good inner approximation quickly. Since it is feasible, this approximation can then be presented to the decision maker, e.g., in a navigation process [1], [31]. The outer approximation serves the important purpose to give an upper bound on the approximation error. But since the elements of the outer approximation are either points that are already known or that are infeasible, the outer approximation has no further practical purpose in our following decision-making process. Therefore, in our context it is not useful to update the inner approximation only in some iterations and the additional optimization problems that need to be solved to obtain the outer approximation with an optimal convergence rate are too costly for little added value.

## References


[1] Richard Allmendinger, Matthias Ehrgott, Xavier Gandibleux, Martin Josef Geiger, Kathrin Klamroth, and Mariano Luque. Navigation in multiobjective optimization methods. *Journal of Multi-Criteria Decision Analysis*, 24(1-2):57–70, 2017.

[2] Çağın Ararat, Firdevs Ulus, and Muhammad Umer. A norm minimization-based convex vector optimization algorithm. *Journal of Optimization Theory and Applications*, 194(2):681–712, 2022.

[3] Çağın Ararat, Firdevs Ulus, and Muhammad Umer. Convergence analysis of a norm minimization-based convex vector optimization algorithm. *arXiv preprint arXiv:2302.08723*, 2023.

[4] Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan M Shetty. *Nonlinear programming: theory and algorithms.* John Wiley & Sons, 2013.

[5] Rasmus Bokrantz and Anders Forsgren. An algorithm for approximating convex pareto surfaces based on dual techniques. *INFORMS Journal on Computing*, 25(2):377–393, 2013.

[6] Guillermo Cabrera-Guerrero, Matthias Ehrgott, Andrew J Mason, and Andrea Raith. Bi-objective optimisation over a set of convex sub-problems. *Annals of Operations Research*, pages 1–26, 2021.

[7] Erik Diessel. Precise control of approximation quality in multicriteria optimization. Preprint, `http://www.optimization-online.org/DB_HTML/2020/09/8036.html`, 2021.

[8] Daniel Dörfler, Andreas Löhne, Christopher Schneider, and Benjamin Weißing. A benson-type algorithm for bounded convex vector optimization problems with vertex selection. *Optimization Methods and Software*, pages 1–21, 2021.

[9] Herbert Edelsbrunner. *Algorithms in combinatorial geometry*, volume 10. Springer Science & Business Media, 1987.

[10] Matthias Ehrgott. *Multicriteria optimization*, volume 491. Springer Science & Business Media, 2005.

[11] Matthias Ehrgott, Lizhen Shao, and Anita Schöbel. An approximation algorithm for convex multi-objective programming problems. *Journal of Global Optimization*, 50(3):397–416, 2011.

[12] Otto Forster. *Analysis 3: Maß-und Integrationstheorie, Integralsätze im IRn und Anwendungen.* Springer Science & Business Media, 2012.

[13] Peter M Gruber and Jörg M Wills. *Handbook of convex geometry: Vol. A.* North-Holland, 1993.

[14] Matthias Hieber. *Analysis II.* Springer Spektrum, Berlin, Heidelberg, 2019.

[15] Claus Hillermeier. Generalized homotopy approach to multiobjective optimization. *Journal of Optimization Theory and Applications*, 110(3):557–583, 2001.

[16] George K Kamenev. Analysis of an algorithm for approximating convex bodies. *Computational mathematics and mathematical physics*, 34(4):521–528, 1994.

[17] George K Kamenev. An algorithm for approximating polyhedra (in russian). *Computational mathematics and mathematical physics*, 4(36):533–544, 1996.

[18] George K Kamenev. Self-dual adaptive algorithms for polyhedral approximation of convex bodies. *Computational mathematics and mathematical physics*, 8(43):1073–1086, 2003.

[19] Georgij Kirillovich Kamenev. A class of adaptive algorithms for approximating convex bodies by polyhedra. *Computational mathematics and mathematical physics*, 32(1):114–127, 1992.

[20] Georgij Kirillovich Kamenev. The initial convergence rate of adaptive methods for polyhedral approximation of convex bodies. *Computational Mathematics and Mathematical Physics*, 48(5):724–738, 2008.

[21] Kathrin Klamroth, Jørgen Tind, and Margaret M Wiecek. Unbiased approximation in multicriteria optimization. *Mathematical Methods of Operations Research*, 56(3):413–437, 2003.

[22] Ina Lammel. *Approximation of Pareto Surfaces in Multi-Criteria Optimization.* PhD thesis, RPTU Kaiserslautern, 2023.

[23] Ina Lammel, Karl-Heinz Küfer, and Philipp Süss. An approximation algorithm for multi-objective mixed-integer convex optimization, 2023. Preprint at `https://optimization-online.org/?p=22340`.

[24] Ina Lammel, Karl-Heinz Küfer, and Philipp Süss. Efficient computation of the approximation quality in sandwiching algorithms, 2023. Preprint at `https://optimization-online.org/?p=24383`.

[25] Julien Legriel, Colas Le Guernic, Scott Cotton, and Oded Maler. Approximating the pareto front of multi-criteria optimization problems. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 69–83. Springer, 2010.

[26] Kurt Leichtweiß. *Konvexe Mengen.* Springer-Verlag Berlin Heidelberg, 1979. To the best of our knowledge, this work is only available in German and Russian.

[27] Andreas Löhne, Birgit Rudloff, and Firdevs Ulus. Primal and dual approximation algorithms for convex vector optimization problems. *Journal of Global Optimization*, 60(4):713–736, 2014.

[28] Alexander V Lotov, Vladimir A Bushenkov, and Georgy K Kamenev. *Interactive decision maps: Approximation and visualization of Pareto frontier*, volume 89. Springer Science & Business Media, 2004.

[29] Anna Marciniak-Czochra. Lecture notes on Funktionalanalysis, July 2018.

[30] Kaisa Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 2012.

[31] Michael Monz, Karl-Heinz Küfer, Thomas R Bortfeld, and Christian Thieke. Pareto navigation—algorithmic foundation of interactive multi-criteria imrt planning. *Physics in Medicine & Biology*, 53(4):985, 2008.

[32] Gijs Rennen, Edwin R van Dam, and Dick den Hertog. Enhancement of sandwich algorithms for approximating higher-dimensional convex pareto sets. *INFORMS Journal on Computing*, 23(4):493–517, 2011.

[33] Günter Rote. The convergence rate of the sandwich algorithm for approximating convex functions. *Computing*, 48(3):337–361, 1992.

[34] Bernd Schandl, Kathrin Klamroth, and Margaret M Wiecek. Introducing oblique norms into multiple criteria programming. *Journal of Global Optimization*, 23(1):81–97, 2002.

[35] Tobias Seidel. *Solving Semi-infinite Optimization Problems with Quadratic Rate of Convergence*. Fraunhofer Verlag, 2020.

[36] Jorge Ivan Serna Hernandez. *Multi-objective optimization in Mixed Integer Problems with application to the Beam Selection Optimization Problem in IMRT*. PhD thesis, TU Kaiserslautern, 2011.

[37] Firdevs Ulus. Tractability of convex vector optimization problems in the sense of polyhedral approximations. *Journal of Global Optimization*, 72(4):731–742, 2018.

[38] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132, 2003.

# A. Proofs that are direct extensions of proofs by Kamenev

*Proof of Lemma 4.6.* Lemma 1 of [17] states: Let $C \in \mathscr{C}^2$, $P \in \mathscr{P}^i(C)$, $\delta^H(P,C) < \rho_{\min}(C)$ and $C' \in \mathscr{C}$, $P \in \mathscr{P}^i(C')$. Then

$$\delta_P(P,C') \geq \left(1 - \frac{\delta^H(P,C)}{\rho_{\min}(C)}\right) \delta^H(P,C').$$

Since its proof is only available in Russian, we state it here. Let $x \in \partial P$. According to Lemma 2 of [16] , $\delta_P(P,C') \geq \xi(P)\delta^H(P,C')$. In addition, for $C \in \mathscr{C}^2$, $P \in \mathscr{P}(C)$ and $\delta^H(P,C) < \rho_{\min}(C)$, according to Lemma 4 of [16] , it holds $\xi(P) \geq [\rho_{\min}(C) - \delta^H(P,C)]/\rho_{\min}(C)$, which yields

$$\begin{aligned}\delta_P(P,C') \geq \xi(P)\; \delta^H(P,C') &\geq \frac{\rho_{\min}(C) - \delta^H(P,C)}{\rho_{\min}(C)} \delta^H(P,C') \\ &= \left(1 - \frac{\delta^H(P,C)}{\rho_{\min}(C)}\right) \delta^H(P,C').\end{aligned}$$

The statement of Lemma 4.6 then follows using $\delta_P(P,C) \leq \delta^H(P,C) \leq c_2\delta(P,C)$ (from Equation 8). □

*Proof of Lemma 2 of [17], used in the proof of Lemma 4.8.* Note that, according to the algorithm, $P^{n+1} = \text{conv}\{\{p^*\} \cup P^n\} \supseteq P^n$ and
$Q^{n+1} = Q^n \cap H_C(u^*) \subseteq Q^n$. Therefore

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) = \delta^S(P^n, P^{n+1}) + \delta^S(Q^n, Q^{n+1}).$$

Denote by $h_P := g_{P^{n+1}}(u^*) - g_{P^n}(u^*)$ and $h_Q := g_{Q^n}(u^*) - g_{Q^{n+1}}(u^*)$. From the scheme of the algorithm it follows that $g_{Q^{n+1}}(u^*) = \langle u^*, p^* \rangle = g_{P^{n+1}}(u^*)$, so $h_P + h_Q = \delta_{P^n}(P^n, Q^n)$.Let $z \in C$ and $r > 0$ such that $B_r(z) \subset P^n$. The value of $\delta^S(P^n, P^{n+1})$ is not less than the volume cut off from the cone of visibility of the ball $B_r(z)$ from point $p^*$ by a hyperplane with norm $u^*$, which is supporting $P^n$. This situation is illustrated in Figure 1.

Lemma 1 of [19] estimates the volume of a pyramid defined by a sphere, an external point and a hyperplane: Suppose we are given the sphere $B_\rho(z) \subset \mathbb{R}^d$, the exterior point $y$ and the hyperplane $H$ separating the sphere and the exterior point. The hyperplane $H$ has distance $h$ to the point $y$. Then $H$ cuts a pyramid from the visibility cone of $B_\rho(z)$ from the point $y$ with a volume not less than

$$\frac{\pi_{d-1}}{d}\left(\left(\frac{||z-y||)}{\rho}\right)^2 - 1\right)^{(1-d)/2} h^d.$$

Therefore, we have

$$\delta^S(P^n, P^{n+1}) \geq \frac{\pi_{d-1}}{d}\left[\left(\frac{||p^*-z||}{r}\right)^2 - 1\right]^{(1-d)/2} h_P^d \geq \frac{\pi_{d-1}}{d}(\zeta_1^2 - 1)^{(1-d)/2} h_P^d.$$

Analogously, let $p \in T_{Q^n}(u^*)$, $z \in C$ and $r > 0$ such that $B_r(z) \subset C$. Then

$$\delta^S(Q^n, Q^{n+1}) \geq \frac{\pi_{d-1}}{d}\left[\left(\frac{||p-z||}{r}\right)^2 - 1\right]^{(1-d)/2} h_Q^d \geq \frac{\pi_{d-1}}{d}(\zeta_2^2 - 1)^{(1-d)/2} h_Q^d$$

Since for any $t, s, a > 0$ with $t + s = a$ we have $t^d + s^d \geq a^d/2^{d-1}$, it follows

$$h_P^d + h_Q^d \geq (h_P + h_Q)^d/2^{d-1} = \delta_{P^n}(P^n, Q^n)^d/2^{d-1}.$$

Therefore, we get

$$\begin{aligned}
&\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) = \delta^S(P^n, P^{n+1}) + \delta^S(Q^n, Q^{n+1}) \\
&\geq \frac{\pi_{d-1}}{d}(\zeta_1^2 - 1)^{(1-d)/2} h_P^d + \frac{\pi_{d-1}}{d}(\zeta_2^2 - 1)^{(1-d)/2} h_Q^d \\
&\geq \frac{\pi_{d-1}}{d}\left(\max\{\zeta_1^2, \zeta_2^2\} - 1\right)^{(1-d)/2}(h_P^d + h_Q^d) \geq \frac{\pi_{d-1}}{d}(\zeta^2 - 1)^{(1-d)/2}\frac{\delta_{P^n}(P^n, Q^n)^d}{2^{d-1}}.
\end{aligned}$$

□

**Lemma A.1** (Lemma 4 of [17]). *Let $\{a_n\}_{n=0,1,\ldots}$ and $\{b_n\}_{n=0,1,\ldots}$ be monotonically decreasing sequences of positive numbers, and let there exist constants $c_1, c_2 > 0$ and $\beta > 1$ such that $a_n - a_{n+1} \geq c_1 b_n^\beta$ and $c_2 b_n \geq a_n$ at $n = 0, 1, \ldots$. Then for any $n \geq 1$ it holds $a_n \leq (\lambda_a n^{1/(\beta-1)})^{-1}$, where*

$$\lambda_a = \left( \frac{(\beta-1)c_1}{c_2^\beta} \right)^{1/(\beta-1)}.$$

*In addition, for any number $n_0 > \beta$, $n \geq n_0$ it holds $b_n \leq (\lambda_b n^{1/(\beta-1)})^{-1}$, where*

$$\lambda_b = \left( \frac{(\beta-1)c_1}{\beta c_2} \left(1 - \frac{\beta}{n_0}\right)^{1/\beta} \right)^{1/(\beta-1)}.$$

*Proof of Lemma 4.9.* The proof is similar to that of Lemma 5 in [17].

Let $z \in C$ and $r, R > 0$ be such that $B_r(z) \subset P^0 \subset Q^0 \subset B_R(z)$. Then the inclusions $B_r(z) \subset P^n \subset Q^n \subset B_R(z)$ hold for any $n \geq 0$.

According to Equation 8 it holds $\delta(P, C) \geq \frac{1}{c_2 \alpha(P)} \delta^H(P, C)$. The asphericity $\alpha(P^n)$ is defined as the minimal ratio of the radii of concentric outer and inner spheres. Therefore, it holds $\alpha(P^n) \leq R/r$. Hence, $\delta(P, C) \geq \frac{r}{c_2 R}\, \delta^H(P, C)$.

We can see that it holds $R/r \geq \zeta$, where $\zeta = \max\{\zeta_1, \zeta_2\}$ is defined as in Lemma 4.8: For every $p \in T_{Q^n}$ we have $R \geq ||p - z||$ and therefore also $R \geq ||p^* - z||$ for every $p^* \in \partial C$. Thus we obtain

$$\frac{R}{r} \geq \frac{||p^* - z||}{r} \geq \zeta_1 \quad \text{and} \quad \frac{R}{r} \geq \frac{||p - z||}{r} \geq \zeta_2.$$

Using Lemma 4.8, we obtain that

$$\begin{aligned}
\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) &\geq \frac{\pi_{d-1}}{2^{d-1}d} (\zeta^2 - 1)^{(1-d)/2} \left( \frac{c_1}{\alpha(P^n)} \right)^d (\delta(P^n, Q^n))^d \\
&\geq \frac{\pi_{d-1}}{2^{d-1}d} \left( \left(\frac{R}{r}\right)^2 - 1 \right)^{(1-d)/2} \left(\frac{r}{R}\right)^d c_1^d \left(\frac{r}{c_2 R}\right)^d \left(\delta^H(P^n, Q^n)\right)^d \\
&\geq \underbrace{\frac{\pi_{d-1}}{2^{d-1}d} \left( \left(\frac{R}{r}\right)^2 - 1 \right)^{(1-d)/2} \left(\frac{c_1}{c_2}\right)^d \left(\frac{R}{r}\right)^{-2d}}_{=:\tilde{c}} \left(\delta^H(P^n, Q^n)\right)^d,
\end{aligned}$$

so

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \geq \tilde{c}\, \delta^H(P^n, Q^n)^d,$$

with $\tilde{c}$ defined above. In the case $\delta = \delta^H$ we obtain

$$\tilde{c} := \frac{\pi_{d-1}}{2^{d-1}d} \left( \left(\frac{R}{r}\right)^2 - 1 \right)^{(1-d)/2} \left(\frac{r}{R}\right)^d.$$

Lemma 3 of [17] states: Let $C \in \mathscr{C}$ and let $R \in \mathbb{R}$ and $z \in \mathbb{R}^d$ be chosen such that $C \subset B_R(z)$. Then for $\epsilon > 0$ $\mu((C)_\epsilon) - \mu(C) \leq \pi_d[(R+\epsilon)^d - R^d]$.

Denote $\delta^S(P^n, Q^n)$ by $x_n$ and $\delta^H(P^n, Q^n)$ by $y_n$. Since $Q^n \subset (P^n)_{y_n}$, using Lemma 3 of [17] we obtain

$$\begin{aligned} x_n =& \mu((P^n \backslash Q^n) \cup (Q^n \backslash P^n)) = \mu(Q^n \backslash P^n) = \mu(Q^n) - \mu(P^n) \\ \leq& \mu((P^n)_{y_n}) - \mu(P^n) \leq \pi_d((R+y_n)^d - R^d). \end{aligned}$$

Using the binomial theorem and $y_n < R$, we have

$$\begin{aligned} ((R+y_n)^d - R^d) &= \sum_{k=0}^{d} \binom{d}{k} R^{d-k} y_n^k - R^d = R^d + R^{d-1} y_n + \sum_{k=2}^{d} \binom{d}{k} R^{d-k} y_n^k \; - R^d \\ &\leq \left(1 + \sum_{k=2}^{d} \binom{d}{k}\right) R^{d-1} y_n = \left(\sum_{k=0}^{d} \binom{d}{k} - 1\right) R^{d-1} y_n = (2^d - 1) R^{d-1} y_n \end{aligned}$$

Therefore $x_n \leq \tilde{d} y_n$, where $\tilde{d} := \pi_d (2^d - 1) R^{d-1}$. In addition, the sequences $\{x_n\}_{n=0,1,\ldots}$ and $\{y_n\}_{n=0,1,\ldots}$ are positive and monotonically decreasing. Therefore with $b_n := y_n$ and the constant $\beta = d \geq 1$ the conditions of Lemma A.1 hold and hence for any $n \geq 1$ $x_n = \delta^S(P^n, Q^n) \leq (\lambda_a n^{1/(d-1)})^{-1}$,where $\lambda_a$ depends only on $\beta$, $c_1$ and $c_2$, and therefore only on $d$, $\delta$, $R$, and $r$. Furthermore, let $n_0 = \beta + 1$. Then, by the same Lemma A.1, for $n \geq n_0$ it holds $y_n = \delta^H(P^n, Q^n) \leq (\lambda_b n^{1/(d-1)})^{-1}$where $\lambda_b$ depends only on $\beta$, $c_1$, $c_2$ and $n_0$, and therefore only on $d$, $\delta$, $R$, and $r$. From Equation 7 we can conclude

$$\delta(P^n, Q^n) \leq \frac{1}{c_1} y_n \leq (c_1 \lambda_b n^{1/(d-1)})^{-1}.$$

□

*Proof of Lemma 4.11.* The proof is similar to that of Lemma 6 in [17].

Let $r'$ and $R'$ be the radii of concentric inner and outer balls for $C$ such that $\alpha(C) = R'/r'$. Since, by Theorem 4.10, the sequences $\{P^n\}_{n=0,1,\ldots}$ and $\{Q^n\}_{n=0,1,\ldots}$ converge to $C$, then for any $\epsilon'$, $0 < \epsilon' < 1$, we can find a $n_0 \in \mathbb{N}$ such that $\delta^H(P^n, Q^n) < \epsilon' r'$ for $n \geq n_0$. Denote by $r = (1-\epsilon')r'$ and $R = (1+\epsilon')R'$. Let $z \in C$ such that $B_{r'}(z) \subset C \subset B_{R'}(z)$. By Lemma 2 of [19] we have $B_r(z) \subset P^n$.

Using $\rho := \delta^H(C, Q^n)$ it holds $Q^n \subset (C)_\rho$. From $C \subset B_{R'}(z)$ and $\rho = \delta^H(C, Q^n) \leq \delta^H(P^n, Q^n) < \epsilon' r' < \epsilon' R'$ it follows $Q^n \subset (C)_\rho \subset B_{R'+\rho}(z) \subset B_{R'+\epsilon' R'}(z)$.Therefore, we have the inclusion $Q^n \subset B_R(z)$.

According to Equation 8 it holds $\delta(P^n, Q^n) \geq 1/(c_2 \alpha(P^n)) \delta^H(P^n, Q^n)$, therefore $\delta(P^n, Q^n) \geq r/(c_2 R) \delta^H(P^n, Q^n)$. In addition, $R/r \geq \zeta$ (see the proof of Lemma 4.9) where $\zeta$ is defined in Lemma 4.8. Hence, as in the proof of Lemma 4.9, for $n \geq n_0$,

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \geq \frac{\pi_{d-1}}{2^{d-1} d} \left(\left(\frac{R}{r}\right)^2 - 1\right)^{(1-d)/2} \left(\frac{c_1}{c_2}\right)^d \left(\frac{R}{r}\right)^{-2d} \left(\delta^H(P^n, Q^n)\right)^d.$$

Consider that $R/r = \alpha(C)(1+\epsilon')/(1-\epsilon')$ and $\alpha(C) \neq 1$. Let us choose $\epsilon' > 0$ so small that

$$\left(\left(\alpha(C)\frac{1+\epsilon'}{1-\epsilon'}\right)^2 - 1\right)^{(1-d)/2} \left(\frac{1+\epsilon'}{1-\epsilon'}\right)^{-2d} \geq (1-\epsilon)(\alpha(C)^2-1)^{(1-d)/2}. \qquad (*_1)$$

Then, for $n \geq n_0$ we obtain

$$\begin{aligned}
&\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \\
&\geq \frac{\pi_{d-1}}{2^{d-1}d}\left(\left(\frac{R}{r}\right)^2 - 1\right)^{(1-d)/2} \left(\frac{c_1}{c_2}\right)^d \left(\frac{R}{r}\right)^{-2d} \left(\delta^H(P^n, Q^n)\right)^d \\
&= \frac{\pi_{d-1}}{2^{d-1}d}\left(\left(\alpha(C)\frac{1+\epsilon'}{1-\epsilon'}\right)^2 - 1\right)^{(1-d)/2} \left(\frac{1+\epsilon'}{1-\epsilon'}\right)^{-2d} \alpha(C)^{-2d} \\
&\quad \left(\frac{c_1}{c_2}\right)^d \left(\delta^H(P^n, Q^n)\right)^d \\
&\overset{(*_1)}{\geq} \frac{\pi_{d-1}}{2^{d-1}d}\left(\alpha(C)^2 - 1\right)^{(1-d)/2} (1-\epsilon)\alpha(C)^{-2d} \left(\frac{c_1}{c_2}\right)^d \left(\delta^H(P^n, Q^n)\right)^d.
\end{aligned}$$

□

*Proof of Theorem 4.12.* The proof is similar to that of Theorem 2 of [17].

Let $\{(P^n, Q^n)\}_{n=0,1,\ldots}$ be a sequence of pairs of polyhedra generated by the Sandwiching algorithm with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}$, $\alpha(C) \neq 1$. Let us denote $\delta^S(P^n, Q^n)$ by $x_n$ and $\delta^H(P^n, Q^n)$ by $y_n$. Then, according to Lemma 4.11, for any $\epsilon_1$, $0 < \epsilon_1 < 1$, there exists a $n_1$ such that for $n \geq n_1$ we have $x_n - x_{n+1} \geq \xi_1(\epsilon_1) y_n^d$.

Define a constant $\alpha_n > 1$ such that $\delta(P^n, \alpha_n P^n) = \delta^H(P^n, Q^n)$. Due to the scheme of the algorithm, we have $Q^n \subset \alpha_n P^n$. Then

$$x_n = \delta^S(P^n, Q^n) \leq \delta^S(P^n, \alpha_n P^n) \text{ and it holds} \qquad (14)$$

$$\sigma(P^n)\delta^H(P^n, Q^n) \leq \delta^S(P^n, \alpha_n P^n). \qquad (15)$$

Note that $\lim_{n\to\infty} \delta^S(P^n, \alpha_n P^n) = 0$ and $\lim_{n\to\infty} \sigma(P^n)\delta^H(P^n, Q^n) = 0$. Using the convergence of these sequences together with the reversed triangle inequality ( [14] remark 1.5), the nonnegativity of the sequences and Equation 15 we get that for every $\tilde{\epsilon}$ there exists a number $N \in \mathbb{N}$ such that

$$\begin{aligned}
&\left|\left|\delta^S(P^n, \alpha P^n)\right| - \left|\sigma(P^n)\delta^H(P^n, Q^n)\right|\right| \leq \left|\delta^S(P^n, \alpha P^n)\right| + \left|\sigma(P^n)\delta^H(P^n, Q^n)\right| < \tilde{\epsilon} \\
&\Leftrightarrow\ \delta^S(P^n, \alpha P^n) \leq \sigma(P^n)\delta^H(P^n, Q^n) + \tilde{\epsilon}, \quad \forall n > N.
\end{aligned}$$

From this result and Equation 14 and defining $\epsilon_2 = \tilde{\epsilon}/\max_n\{\delta^H(P^n, Q^n)\} = \tilde{\epsilon}/\delta^H(P^0, Q^0)$ we obtain that for all $\epsilon_2 > 0$ there exists an $n_2 \in \mathbb{N}$ such that

$$x_n \leq \delta^S(P^n, \alpha P^n) \leq (\sigma(P^n) + \epsilon_2)\, y_n \quad \forall n > n_2.$$

For convex compact sets $A$, $B$ with $A \subset B$ it holds $\sigma(A) \leq \sigma(B)$ (Theorem 15.8 and Remark 15.4 of [26]). Therefore, from $P^n \subset C$ we obtain $\sigma(P^n) \leq \sigma(C)$ and conclude that for any $\epsilon_2 > 0$ there exists a number $n_2$, at which

$$x_n \leq (\sigma(P^n) + \epsilon_2) y_n \leq (\sigma(C) + \epsilon_2) y_n, \ n \geq n_2.$$

In addition, the terms of the sequences $\{x_n\}_{n=0,1,\dots}$ and $\{y_n\}_{n=0,1,\dots}$ are positive and monotonically decreasing. Therefore, for $n \geq n_3 = \max\{n_1, n_2\}$ the conditions of Lemma A.1 are valid with constants $c_1 = \xi_1(\epsilon_1)$, $c_2 = \sigma(C) + \epsilon_2$ and $\beta = d$. From Lemma A.1 we obtain

$$\begin{aligned} x_n =& \delta^S(P^n, Q^n) \leq \left( (d-1)\xi_1(\epsilon_1) \left[ \frac{1}{\sigma(C) + \epsilon_2} \right]^d n \right)^{1/(1-d)} \\ =& \left( (d-1)(1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d} \left( \alpha(C)^2 - 1 \right)^{(1-d)/2} \alpha(C)^{-2d} \left( \frac{c_1}{c_2} \right)^d \left[ \frac{1}{\sigma(C) + \epsilon_2} \right]^d n \right)^{1/(1-d)} \end{aligned}$$

and

$$\begin{aligned} y_n &= \delta^H(P^n, Q^n) \\ &\leq \left( \left( \frac{d-1}{d} \xi_1(\epsilon_1) \frac{1}{\sigma(C) + \epsilon_2} \left( 1 - \frac{d}{\tilde{n}} \right)^{1/d} \right)^{1/(d-1)} n^{1/(d-1)} \right)^{-1} \\ &= \left( \frac{d-1}{d} \xi_1(\epsilon_1) \frac{1}{\sigma(C) + \epsilon_2} n \right)^{1/(1-d)} \left( 1 - \frac{d}{\tilde{n}} \right)^{1/(d(1-d))} \quad \text{for any } \tilde{n} > d, \ n \geq \tilde{n}. \end{aligned}$$

Since $\left(1 - \frac{d}{\tilde{n}}\right)^{1/(d(1-d))} > 1$ is constant, we can rewrite the statement: for any $\epsilon_3 > 0$ there exists a number $n_4 \geq n_3$, for which

$$\begin{aligned} y_n &= \delta^H(P^n, Q^n) \leq (1+\epsilon_3) \left( \frac{d-1}{d} \xi_1(\epsilon_1) \frac{1}{\sigma(C) + \epsilon_2} n \right)^{1/(1-d)} \\ &= (1+\epsilon_3) \left( (d-1)(1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d^2} \left( \alpha(C)^2 - 1 \right)^{(1-d)/2} \alpha(C)^{-2d} \left( \frac{c_1}{c_2} \right)^d \frac{1}{\sigma(C) + \epsilon_2} n \right)^{1/(1-d)}, \end{aligned}$$

$n \geq n_4$. Let us choose $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$ such that

$$(1+\epsilon_3)(1-\epsilon_1)^{1/(1-d)} \left( 1 + \frac{\epsilon_2}{\sigma(C)} \right)^{d/(d-1)} \leq 1 + \epsilon. \qquad (*_2)$$

Then, using $(*_2)$ and introducing the constant $b_1$ defined in the formulation of Theorem 4.12, we get the first statement of the theorem:

$$
\begin{aligned}
x_n \leq & \left( (d-1)(1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d} \left(\alpha(C)^2 - 1\right)^{(1-d)/2} \alpha(C)^{-2d} \left(\frac{c_1}{c_2}\right)^d \left[\frac{1}{\sigma(C)+\epsilon_2}\right]^d n \right)^{1/(1-d)} \\
=& 2 \left( (d-1)\frac{\pi_{d-1}}{d} \right)^{1/(1-d)} (1-\epsilon_1)^{1/(1-d)} (\alpha(C)^2-1)^{1/2} \alpha(C)^{2d/(d-1)} \left(\frac{c_1}{c_2}\right)^{d/(1-d)} \\
& \left(1 + \frac{\epsilon_2}{\sigma(C)}\right)^{d/(d-1)} \sigma(C)^{d/(d-1)} n^{1/(1-d)} \\
=& \Bigg( \underbrace{\frac{1}{2} \left( (d-1)\frac{\pi_{d-1}}{d} \frac{1}{\sigma(C)^d} \right)^{1/(d-1)} (\alpha(C)^2-1)^{-1/2} \alpha(C)^{2d/(1-d)} \left(\frac{c_1}{c_2}\right)^{d/(d-1)}}_{=b_1} n^{1/(d-1)} \Bigg)^{-1} \\
& (1-\epsilon_1)^{1/(1-d)} \left(1 + \frac{\epsilon_2}{\sigma(C)}\right)^{d/(d-1)} \\
\overset{(*_2)}{\leq} & (1-\epsilon) \left(b_1 n^{1/(d-1)}\right)^{-1} \quad \text{for } n \geq n_0 = n_4.
\end{aligned}
$$

Similarly, we obtain the second statement

$$
\begin{aligned}
y_n \leq & (1+\epsilon_3) \left( (d-1)(1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d^2} \left(\alpha(C)^2 - 1\right)^{(1-d)/2} \alpha(C)^{-2d} \left(\frac{c_1}{c_2}\right)^d \frac{1}{\sigma(C)+\epsilon_2} n \right)^{1/(1-d)} \\
=& \left( \frac{1}{2} \left( \frac{d-1}{d} \frac{\pi_{d-1}}{d} \right)^{1/(d-1)} (\alpha(C)^2-1)^{-1/2} \alpha(C)^{2d/(1-d)} \left(\frac{c_1}{c_2}\right)^{d/(d-1)} n^{1/(d-1)} \right)^{-1} \\
& (1+\epsilon_3)(1-\epsilon_1)^{1/(1-d)} \left(1 + \frac{\epsilon_2}{\sigma(C)}\right)^{1/(d-1)} \sigma(C)^{1/(d-1)} \\
=& \Bigg( \underbrace{\frac{1}{2} \left( \frac{d-1}{d} \frac{\pi_{d-1}}{d} \frac{1}{\sigma(C)} \right)^{1/(d-1)} (\alpha(C)^2-1)^{-1/2} \alpha(C)^{2d/(1-d)} \left(\frac{c_1}{c_2}\right)^{d/(d-1)}}_{=b_2} n^{1/(d-1)} \Bigg)^{-1} \\
& (1+\epsilon_3)(1-\epsilon_1)^{1/(1-d)} \left(1 + \frac{\epsilon_2}{\sigma(C)}\right)^{1/(d-1)} \\
\overset{(*_2)}{\leq} & (1+\epsilon) \left(b_2 n^{1/(d-1)}\right)^{-1} \quad \text{for } n \geq n_0 = n_4.
\end{aligned}
$$

From Equation 7 we can conclude

$$
\delta(P^n, Q^n) \leq \frac{1}{c_1} y_n \leq (1+\epsilon) \left(c_1 b_2 n^{1/(d-1)}\right)^{-1} \quad \text{for } n \geq n_0 = n_4.
$$

Now, the assertions of Theorem 4.12 for the number of iterations $K(n) = n$ have been proven. The validity of them when $K(n)$ is the number of vertices of the inner approximation $n^v(P^n)$ or the number of facets of the outer approximation $n^f(Q^n)$ follows from Equation 4. □

*Proof of Lemma 4.14.* The proof is similar to that of Lemma 7 in [17] and given in the Appendix.

Since, by Theorem 4.10, $\{P^n\}_{n=0,1,\dots}$ and $\{Q^n\}_{n=0,1,\dots}$ converge to $C$, then for any $\gamma$, $0 < \gamma < 1$, we can find $n_0$ such that $\delta^H(P^n, C) \le \delta^H(P^n, Q^n) \le \gamma\rho_{\min}(C)$ for $n \ge n_0$. By Lemma 4.6, with $\delta^H(P^n, C) \le \gamma(\epsilon)\rho_{\min}(C)$ and Equation 8, we have

$$\delta(P^n, Q^n) \ge \frac{1}{c_2}\left(1 - \gamma(\epsilon)\right)\delta^H(P^n, Q^n).$$

By Lemma 4.8,

$$\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \ge \frac{\pi_{d-1}}{2^{d-1}d}(\zeta^2 - 1)^{(1-d)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{r_{\text{outer}}(C)}\right)^d (\delta(P^n, Q^n))^d.$$

Let us evaluate $\zeta$. Let $u^*$ and $p^*$ be the direction and point chosen for $(P^n, Q^n)$ at step 1 of the algorithm. Denote $\rho_{\min}(C)$ by $\rho$. By Blaschke's rolling theorem (Lemma 4.13), there exists a $z \in C$ such that $B_\rho(z) \subset C$ and $p^* \in B_\rho(z)$. By Lemma 2 of [19], it follows that $B_{\rho'}(z) \subset P^n$ where $\rho' = \rho - \delta^H(P^n, C)$. Therefore

$$\zeta_1 = \inf\left\{\frac{||p^* - z||}{r} \,:\, z \in C,\ B_r(z) \subset P^n\right\} \le \frac{||p^* - z||}{\rho'} = \frac{\rho}{\rho'} = \frac{\rho}{\rho - \delta^H(P^n, C)}.$$

Since $\rho > \gamma\rho \ge \delta^H(P^n, Q^n) \ge \delta^H(P^n, C)$, it holds

$$\zeta_1 \le \frac{\rho}{\rho - \delta^H(P^n, C)} \le \frac{\rho}{\rho - \delta^H(P^n, Q^n)} = \left(\frac{\rho - \delta^H(P^n, Q^n)}{\rho}\right)^{-1} = \left(1 - \frac{\delta^H(P^n, Q^n)}{\rho}\right).$$

With $x = \frac{\delta^H(P^n,Q^n)}{\rho}$ it holds $\frac{1}{1-x} \le 1 + \frac{x}{1-\gamma}$ since $x \le \gamma$ due to $\delta^H(P^n, Q^n) \le \gamma\rho$:

$$\begin{aligned}
&\frac{1}{1-x} \le 1 + \frac{x}{1-\gamma} \Leftrightarrow \frac{1}{1-x} - \frac{x}{1-\gamma} \le 1 \Leftrightarrow \frac{(1-\gamma) - x(1-x)}{(1-x)(1-\gamma)} \le 1 \\
\Leftrightarrow &1 - \gamma - x + x^2 \le 1 - \gamma - x + x\gamma \Leftrightarrow x \le \gamma.
\end{aligned}$$

We thus obtain

$$\zeta_1 \le 1 + \frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}.$$

Now let $p \in T_{Q^n}(u^*)$ and $p'$ the projection of $p$ on $C$. Since $C$ is a nonempty, closed and convex set, this projection exists and is unique (Theorem 1.46 of [29]). The point $p'$ is the element of $C$ with the smallest distance to $p$ and the (unique) tangent on $C$ in $p'$ is perpendicular to $(p - p')$ (Lemma 1.47 of [29]). By the Blaschke rolling theorem (Lemma 4.13), there exists $z \in C$ such that $B_\rho(z) \subset C$ with $p' \in B_\rho(z)$. This ball touches $C$ in $p'$. The vector $(p' - z)$ is orthogonal to the tangent of $C$ in $p'$. Therefore, $p' \in [z, p]$.

It holds $||p - p'|| \leq \delta^H(Q^n, C)$. Since we have $\frac{1}{1-\gamma} > 1$ and $\delta^H(P^n, C) \leq \delta^H(P^n, Q^n)$, it follows

$$\begin{aligned}\zeta_2 &\leq \frac{||p - z||}{\rho} = \frac{||p' - z|| + ||p - p'||}{\rho} \leq \frac{\rho + \delta^H(Q^n, C)}{\rho} = 1 + \frac{\delta^H(Q^n, C)}{\rho} \\ &\leq 1 + \frac{\delta^H(Q^n, C)}{\rho}\frac{1}{1-\gamma} \leq 1 + \frac{\delta^H(P^n, Q^n)}{\rho(1-\gamma)}.\end{aligned}$$

Since $\zeta = \max\{\zeta_1, \zeta_2\}$ we get

$$\zeta \leq 1 + \frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}.$$

Now we obtain

$$\begin{aligned}\zeta^2 - 1 &\leq \left(1 + \frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}\right)^2 - 1 = 1 + 2\frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)} + \frac{\delta^H(P^n, Q^n)^2}{\rho_{\min}(C)^2(1-\gamma)^2} - 1 \\ &\leq 2\frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)} + \frac{2\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}\frac{1}{1-\gamma}\frac{\gamma}{2} \quad \text{since } \frac{\delta^H(P^n, Q^n)}{\rho_{\min}(C)} \leq \gamma \\ &= \frac{2\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}\left(1 + \frac{1}{1-\gamma}\frac{\gamma}{2}\right)\end{aligned}$$

So, using Lemma 4.8 and $\delta(P^n, Q^n) \geq 1/c_2(1-\gamma)\delta^H(P^n, Q^n)$,

$$\begin{aligned}&\delta^S(P^n, Q^n) - \delta^S(P^{n+1}, Q^{n+1}) \\ &\geq \frac{\pi_{d-1}}{2^{d-1}d}(\zeta^2 - 1)^{(1-d)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{r_{\text{outer}}(C)}\right)^d (\delta(P^n, Q^n))^d \\ &\geq \frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{2\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}\left(1 + \frac{1}{1-\gamma}\frac{\gamma}{2}\right)\right)^{(1-d)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{r_{\text{outer}}(C)}\right)^d (\delta(P^n, Q^n))^d \\ &\geq \frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{2\delta^H(P^n, Q^n)}{\rho_{\min}(C)(1-\gamma)}\left(1 + \frac{1}{1-\gamma}\frac{\gamma}{2}\right)\right)^{(1-d)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{r_{\text{outer}}(C)}\right)^d \\ &\left(\frac{1}{c_2}\right)^d (1-\gamma)^d \left(\delta^H(P^n, Q^n)\right)^d \\ &= \frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{1}{(1-\gamma)}\left(1 + \frac{1}{1-\gamma}\frac{\gamma}{2}\right)\right)^{(1-d)/2}(1-\gamma)^d\left(\frac{\rho_{\min}(C)}{2}\right)^{(d-1)/2} \\ &\left(\frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)}\right)^d \left(\delta^H(P^n, Q^n)\right)^{(d+1)/2}.\end{aligned}$$

To obtain the statement of the lemma, we need to choose $\gamma$ such that

$$(1-\gamma)^d\left(\frac{1}{1-\gamma}\left(1 + \frac{\gamma}{2}\frac{1}{1-\gamma}\right)\right)^{(1-d)/2} \geq 1 - \epsilon$$

and put $\gamma(\epsilon) = \gamma$. □

*Proof of Theorem 4.15.* The proof is similar to that of Theorem 3 of [17].

Let $\{P^n, Q^n\}_{n=0,1,\ldots}$ be a sequence of pairs of polyhedra, generated by Sandwiching with strongly equivalent selection (Algorithm 3) for $C \in \mathscr{C}^2$. Denote $\delta^S(P^n, Q^n)$ by $x_n$ and $\delta^H(P^n, Q^n)$ by $y_n$. Then, according to Lemma 4.14, for any $\epsilon_1$, $0 < \epsilon_1 < 1$, there exists $n_1$ such that for $n \geq n_1$ we have $x_n - x_{n+1} \geq \xi_2(\epsilon_1, \rho_{\min}(C)) y_n^{(d+1)/2}$.

According to properties of the surface area $\sigma(\cdot)$ (see the proof of Theorem 4.12), for any $\epsilon_2 > 0$ there exists a number $n_2$ at which $x_n \leq (\sigma(C) + \epsilon_2) y_n$ for $n \geq n_2$. Besides, the terms of the sequences $\{x_n\}_{n=0,1,\ldots}$ and $\{y_n\}_{n=0,1,\ldots}$ are positive and monotonically decreasing. Therefore, for $n \geq n_3 = \max\{n_1, n_2\}$ the conditions of Lemma A.1 are valid with constants $c_1 = \xi_2(\epsilon_1, \rho_{\min}(C))$, $c_2 = \sigma(C) + \epsilon_2$ and $\beta = (d+1)/2$. Therefore for any $\epsilon_3 > 0$ there exists a number $n_4 \geq n_3$ for which

$$
\begin{aligned}
x_n = \delta^S(P^n, Q^n) &\leq \left( \frac{d-1}{2} \xi_2(\epsilon_1, \rho_{\min}(C)) \left[ \frac{1}{\sigma(C) + \epsilon_2} \right]^{(d+1)/2} n \right)^{2/(1-d)} \\
&= \left( \frac{d-1}{2} (1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d} \left( \frac{\rho_{\min}(C)}{2} \right)^{(d-1)/2} \left( \frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)} \right)^d \left[ \frac{1}{\sigma(C) + \epsilon_2} \right]^{(d+1)/2} n \right)^{2/(1-d)}
\end{aligned}
$$

and

$$
\begin{aligned}
y_n &= \delta^H(P^n, Q^n) \\
&\leq \left(1 - \frac{d+1}{2\tilde{n}}\right)^{2/(1-d^2)} \left( \frac{d-1}{d+1} \xi_2(\epsilon_1, \rho_{\min}(C)) \frac{1}{\sigma(C) + \epsilon_2} n \right)^{2/(1-d)}, \qquad \tilde{n} > \frac{d+1}{2} \\
&\leq (1 + \epsilon_3) \left( \frac{d-1}{d+1} \xi_2(\epsilon_1, \rho_{\min}(C)) \frac{1}{\sigma(C) + \epsilon_2} n \right)^{2/(1-d)}, \qquad n \geq n_4 \\
&\leq (1 + \epsilon_3) \left( \frac{d-1}{d+1} (1-\epsilon_1) \frac{\pi_{d-1}}{2^{d-1} d} \left( \frac{\rho_{\min}(C)}{2} \right)^{(d-1)/2} \right. \\
&\qquad \left. \left( \frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)} \right)^d \frac{1}{\sigma(C) + \epsilon_2} n \right)^{2/(1-d)}, \qquad n \geq n_4.
\end{aligned}
$$

Let's choose $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$ so that

$$
(1 + \epsilon_3)(1 - \epsilon_1)^{2/(1-d)} [1 + \epsilon_2 / \sigma(C)]^{2d/(d-1)} \leq 1 + \epsilon. \qquad (*_3)
$$

Then, using $(*_3)$ and introducing the constants $b_3$ and $b_4$ defined in the formulation of

Theorem 4.15, we get the statement of the theorem:

$$
\begin{aligned}
x_n &\leq \left(\frac{d-1}{2}(1-\epsilon_1)\frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{\rho_{\min}(C)}{2}\right)^{(d-1)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)}\right)^d\left[\frac{1}{\sigma(C)+\epsilon_2}\right]^{(d+1)/2} n\right)^{2/(1-d)} \\
&= \underbrace{\left(\frac{d-1}{2}\frac{\pi_{d-1}}{d}\left(\frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)}\right)^d \frac{1}{\sigma(C)^{(d+1)/2}}\right)^{2/(1-d)} \frac{8}{\rho_{\min}(C)}}_{=b_3^{-1}} \\
&\qquad (1-\epsilon_1)^{2/(1-d)}\left(1+\frac{\epsilon_2}{\sigma(C)}\right)^{(d+1)/(d-1)} n^{2/(1-d)} \\
&\overset{(*_3)}{\leq} (1+\epsilon)\left(b_3 n^{2/(d-1)}\right)^{-1}
\end{aligned}
$$

and

$$
\begin{aligned}
y_n &\leq (1+\epsilon_3) \\
&\quad \left(\frac{d-1}{d+1}(1-\epsilon_1)\frac{\pi_{d-1}}{2^{d-1}d}\left(\frac{\rho_{\min}(C)}{2}\right)^{(d-1)/2}\left(\frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)}\right)^d \frac{1}{\sigma(C)+\epsilon_2} n\right)^{2/(1-d)} \\
&= \underbrace{\left(\frac{d-1}{d+1}\frac{\pi_{d-1}}{d}\left(\frac{c_1 r_{\text{inner}}(P^0)}{c_2 r_{\text{outer}}(C)}\right)^d \frac{1}{\sigma(C)}\right)^{2/(1-d)} \frac{8}{\rho_{\min}(C)}}_{=b_4^{-1}} \\
&\qquad (1-\epsilon_1)^{2/(1-d)}(\sigma(C)+\epsilon_2)^{2/(d-1)} n^{2/(1-d)} \\
&\overset{(*_3)}{\leq} (1+\epsilon)\left(b_4 n^{2/(d-1)}\right)^{-1} \qquad \text{for } n \geq n_0 = n_4.
\end{aligned}
$$

From Equation 7 we can conclude $\delta(P^n, Q^n) \leq \frac{1}{c_1} y_n \leq (1+\epsilon)\left(c_1 b_4 n^{2/(d-1)}\right)^{-1}$ for $n \geq n_0 = n_4$.

Now, the assertions of Theorem 4.15 are proved for $K(n) = n$. The validity of them for $K(n)$ equal to the number of vertices of the inner approximation $n^v(P^n)$ or the number of facets of the outer approximation $n^f(Q^n)$ follows from Equation 4. □