Delft University of Technology

Master of Science Thesis

Computer Science (IN5000) / Embedded Systems (ES5000)

Obtaining Smoothly Navigable Approximation Sets in Bi-Objective Multi-Modal Optimization with an Application to Prostate HDR Brachytherapy Automated Treatment Planning

---

Author:
Renzo Johan Scholman
Student number 4952308

To obtain the degree of Master of Science in:
Computer Science, track Data Science and Technology, and
Embedded Systems, track Computer Architecture,
at the Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science.

Thesis committee:

Prof. dr. Peter A.N. Bosman	CWI, TU Delft, supervisor
Dr. Tanja Alderliesten	LUMC, supervisor
Dr. Annibale Panichella	TU Delft
Anton Bouter, MSc.	CWI
Leah R.M. Dickhoff, MSc.	LUMC

April 25, 2022

TU Delft Delft University of Technology

CWI Centrum Wiskunde & Informatica

Preface

It has been a long journey to get here, but this thesis is the culmination of my many years at the university. Over the past years, at both the Leiden University for my bachelor and the Delft University of Technology for my masters, I have learned a great deal about various aspects of Computer Science and Embedded Systems. Most of which came together during the last year, in which I have written my masters thesis on a topic that is important to me at the Centrum Wiskunde & Informatica. I am very grateful I had the chance to work on research that can help improve the treatment of cancer patients. Moreover, I am definitely looking forward to the next couple of years, in which I hope to achieve much more in this direction.

Even though COVID spoiled the party as I had to work from home for the majority of the time spent on my thesis, it has been an awesome experience. The weekly zoom-calls, many hours spent working on fixing bugs, producing results, writing texts, I think I can finally say I really do know every nook an cranny of my room now. However, I would not have been able to do it if it were not for the exceptional guidance and support I have received from Prof. dr. Peter Bosman, Dr. Tanja Alderliesten, MSc. Leah Dickhoff and MSc. Anton Bouter. Their guidance, extensive knowledge, and keen eye for detail is what helped me do the research and learn a great deal!

Finally I would like to thank my family and all of my friends, who have helped me through the past couple of years. The moments I got to spend with them and support I have received are unforgettable.

Renzo Johan Scholman
Leiden, April 2022

Abstract

Even if a Multi-modal Multi-Objective Evolutionary Algorithm (MMOEA) is designed to find all locally optimal approximation sets of a Multi-modal Multi-objective Optimization Problem (MMOP), there is a risk that the found approximation sets are not smoothly navigable because the solutions belong to various niches, which reduces the insight for decision makers. Moreover, when the multi-modality of MMOPs increases, this risk grows and the trackability of finding all locally optimal approximation sets decreases. One example where this issue occurs is that of High-Dose-Rate (HDR) brachytherapy for prostate cancer. In HDR brachytherapy a treatment plan is to be optimized that irradiates a tumour with a prescribed dose, whilst sparing all of the healthy organs surrounding the tumour. The radiation is administered through a radioactive source that is stopped at certain dwell positions within a set of hollow catheters that have been implanted into the patient. In a treatment plan, each of the dwell positions is given a specific dwell time for which the source is kept at that location in order to irradiate the surrounding tissue.

To tackle the navigability issues, two new MMOEAs are proposed: Multi-Modal Bézier Evolutionary Algorithm (MM-BezEA) and Set Bézier Evolutionary Algorithm (Set-BezEA). Both MMOEAs produce approximation sets that cover individual niches and exhibit inherent decision-space smoothness as they are parameterized by Bézier curves. MM-BezEA combines the concepts behind the recently introduced BezEA and MO-HillVallEA to find all locally optimal approximation sets. Set-BezEA employs a novel multi-objective fitness function formulation to find limited numbers of diverse, locally optimal, approximation sets for MMOPs of high multi-modality.

Both algorithms, but especially MM-BezEA, are found to outperform the MMOEAs MO_Ring_PSO_SCD and MO-HillVallEA on MMOPs of moderate multi-modality with linear Pareto sets. Moreover, for MMOPs of high multi-modality, Set-BezEA is found to indeed be able to produce high-quality approximation sets, each pertaining to a single niche. Set-BezEA is also shown to be comparable to the current BRIGHT approach used in the Amsterdam UMC for the optimization of treatment plans for prostate cancer HDR brachytherapy, which opens the way for it to be introduced in the clinical practice in the future.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

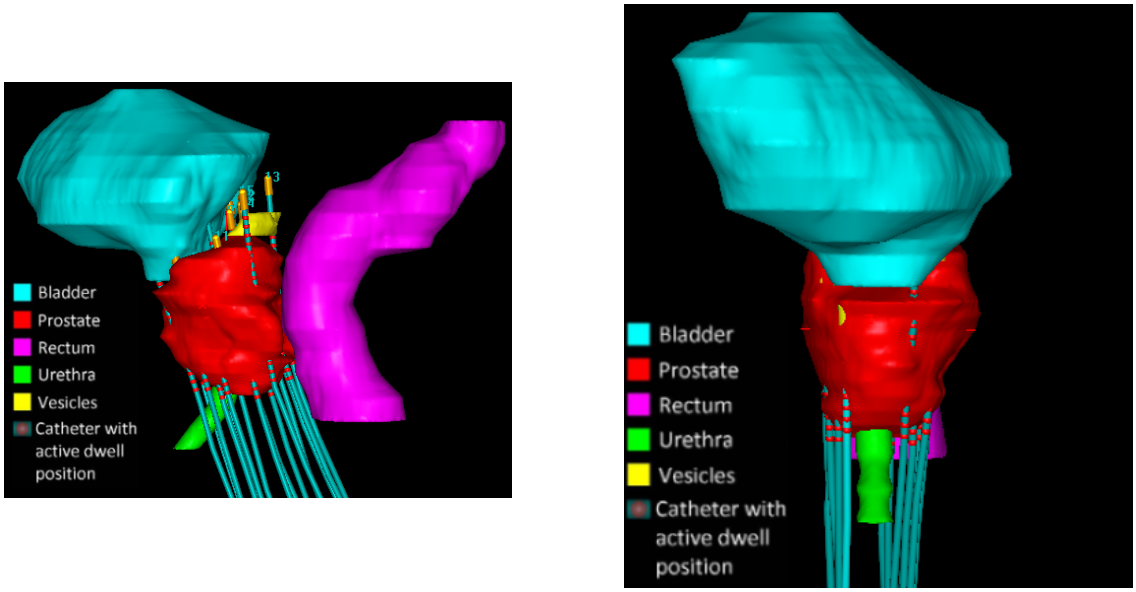
Introduction

Many people in our world suffer from terrible diseases like cancer. Last year alone, an estimated 19.3 million new cases were diagnosed around the world [62]. Out of all of the different types of cancer, prostate and cervix cancer are two of the most common forms of cancer. Taking both cancer forms together, they accounted for a total of just over 2 million new cases that were recorded in the last year. Fortunately, current advances in cancer treatments allow us to better treat patients and give them a higher chance of surviving [40].

1.1 High-Dose-Rate Brachytherapy

Both prostate and cervix cancer can be treated with a method called High Dose-Rate (HDR) brachytherapy [68], which is a type of internal radiation treatment that is generally administered in addition to external beam radiation. This treatment method makes use of hollow catheters that are implanted into the target volumes, i.e., the region(s) of interest targeted by the treatment, through which a very small and highly radioactive source is guided. This radioactive source is stopped at specific positions within the catheters for a specific amount of time in order to irradiate the tumour with a prescribed amount of radiation. At the Amsterdam University Medical Centers (Amsterdam UMC), location Academic Medical Center (AMC), a Magnetic Resonance Imaging (MRI) scan is made after the insertion of the catheters during an operation. This scan is used to annotate the location of the catheters and the target volumes of the surrounding healthy organs, called the Organs At Risk (OARs). Subsequently, the doctors can create treatment plans with the aid of software that takes this scan into account. Within the planning software, doctors can manually change the planned radiation dose in various parts of the patients body. The result is a treatment plan that describes how much radiation is given off by the source at each location in the catheters. After planning, the treatment would be performed directly after the operation in which the catheters were inserted. Figure 1.1 shows a 3D rendering of a patient. It shows the target volumes, organs at risk, and catheters as rendered from the front view and side view.

The planning process used to be very time consuming. Even with the use of software that aids the doctors in the planning process, it could take up to 1 hour to complete [35, 51]. This long delay is not desired. Furthermore, when making a plan, the position of the catheters compared to the organs and tumour is very important as each of these have a prescribed minimum or maximum radiation dose. When the planning phase takes a longer time, there is a higher chance that the positions of the catheters can change due to patient movements. This can result in a mismatch between where the planner believes the catheters are as seen on the scan and where they actually are after the movements. Subsequently, this leads to the undesirable scenario in which the planner believes to have met all prescribed radiation doses, whereas they actually did not due to the changed positions. Therefore, it is important to keep the planning phase as short as possible to reduce the total treatment time and chance of a mismatch between the software and reality.



(a) Viewed from the left side of the patient

(b) Viewed from the front side of the patient

Figure 1.1: 3D rendering of target volumes (prostate and vesicles) and organs at risk (bladder, urethra, and rectum) for a patient. Small red dots denote the dwell positions within the catheters (in cyan) that enter each plot from the bottom and run through the various organs.

1.1.1 Optimizing Treatment Plans

There is a long history of various methods and models that have been created to optimize the planning process, with some of the first research dating back to 1990 [16]. In order to optimize treatment planning, it can be mathematically described in terms of constraints that need to be met [22, 32]. These constraints are modelled according to a clinical protocol that is based on various clinical studies. Subsequently, such a mathematical model can be used to solve a problem optimally or approximate a good result.

In essence, the problem boils down to a bi-objective optimization problem that is a trade-off between tumour coverage and organ sparing. One particular bi-objective problem formulation employs two objectives called the Least Sparing Index (LSI) and Least Coverage Index (LCI) [43]. The LSI measures how the healthy OARs around the tumour are spared by keeping them under a maximum tolerable radiation dose. In order to preserve the OARs not all possible positions in the catheters are used in a treatment plan, since that would result in too much radiation. On the other hand, the contrasting objective of the LCI measures how much of the tumour is given the minimum required radiation dose to kill the cancer cells. As these two objectives are adversaries of each other, it is difficult to find good plans that maximize both objectives.

Using the LSI and LCI objectives, treatment planning for HDR brachytherapy is a problem that can be modelled as a multi-modal multi-objective optimization problem with real-valued variables. For these types of optimization problems there are various approaches that one can employ. One of the most commonly used approaches for this specific optimization problem is that of Evolutionary Algorithms (EAs). An EA is an algorithm that takes inspiration from nature, especially from the idea of biological evolution and survival of the fittest. Many forms of EAs exist, of which one type is that of the population-based meta-heuristic optimization algorithms. These algorithms have shown to perform really well on this HDR brachytherapy problem [43].

1.1.2 BRIGHT

The Multi-Objective Real-Valued Gene-pool Optimal Mixing Evolutionary Algorithm (MO-RV-GOMEA) [13] is an EA that is being used at the AMC together with the bi-objective problem formulation for HDR prostate cancer. In this form it is referred to as BRachytherapy via artificially Intelligent GOMEA-Heuristic based Treatment planning (BRIGHT) and has successfully been introduced into clinical practice since March 2020 [5].

Just a few years ago, this approach has been applied and benchmarked on several prostate cancer patients. In an observer study it showed such promising results that those plans that are created with BRIGHT have been chosen over those created by the clinical practice at the time in 53/54 cases and 17/18 patients [50]. The use of the new BRIGHT approach not only resulted in a diverse set of highly quality plans, but also in a lot less computation time. A recent study into parallelizing partial evaluations for EAs shows that significant speedups can be achieved [10]. This parallelization technique has been implemented for BRIGHT, which resulted in a computation time of just 30 seconds instead of one hour [11].

1.1.3 Approximation Front Navigability

With the use of BRIGHT, the results of the approach can be seen as a trade-off curve or the so-called approximation front. An example of an approximation front as produced by BRIGHT is given in Figure 1.2. On this curve lie all the different solutions that were found by the approach, each of which represents a high-quality trade-off between the opposing objectives. A pre-selection of a small number of plans with desirable trade-offs in the LCI and LSI objectives was made based on the visualized trade-off curve. Subsequently, the dose distributions of the selected plans were visually inspected by a doctor, after which a single desirable plan is chosen. Even though radiation oncologists appreciated the gained insight from inspecting and comparing several plans [50], some issues remained.

Although the solutions on the approximation front represent high-quality trade-offs in terms of the objectives, the neighbouring solutions on this front differ quite substantially in terms of decision variable values. In the case of HDR brachytherapy these differences amount to radiation being administered at completely different locations in the catheters, since not all locations are actively used in each treatment plan to prevent too much radiation. This dissimilarity in dose distributions could result in a time consuming inspection for the doctors in order for them to choose the best plan for the patient. Since, if the plans do not vary smoothly in dose distribution, there might be a necessity to inspect all plans before the doctors can be satisfied that the most desirable plan has been selected. If plans along the trade-off curve would indeed vary smoothly, there can be an intuitive variation in the properties of these plans. This should result in making the inspection of plans with similar trade-offs very user friendly and intuitive.

In a recent study, MO-RV-GOMEA was adapted with a new fitness function that parameterizes approximation sets as Bézier curves. The resulting algorithm is called the Bézier Evolutionary Algorithm (BezEA) [72]. This adaptation was created specifically for bi-objective problems, making it applicable to the problem of HDR brachytherapy. Results showed that the created plans were of slightly lower quality in terms of obtained LCI and LSI values compared to MO-RV-GOMEA. However, it did create approximation fronts in which the treatment plans have dose distributions that vary perfectly smoothly when navigating the front from one extreme to the other, which was previously unattainable.

One downside to coming up with a single approximation set that has smoothly varying dose distributions is that the doctors lose the differentiation of plans that was present beforehand. Even though this differentiation itself had its downside in terms of a time consuming inspection to select the best plan for the patient, it did allow doctors to choose between plans that held the radioactive source at completely different positions in the catheters. This gives the doctors the opportunity to search for plans that suit the anatomy of a patient well. However, if a plan suits the patient nearly perfectly but needs just a little more sparing, the doctors are back at the previous issue of finding a completely different plan along the approximation front. Therefore, a diverse set of smoothly varying approximation fronts would be preferable.

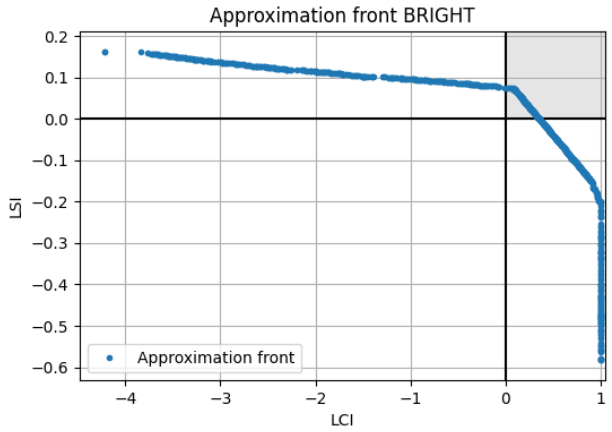


Figure 1.2: Approximation front produced by BRIGHT. In grey corner all clinical aims have successfully been met, as indicated by positive LCI and LSI. Blue circles denote treatment plans, each with a different trade-off.

1.2 Multi-Modal Multi-Objective Optimization problems

Many real-world optimization problems have multiple conflicting objectives, similar to HDR brachytherapy. Multi-Objective EAs (MOEAs), like MO-RV-GOMEA, the Strength Pareto Evolutionary Algorithm (SPEA) [78], and Non-dominated Sorting Evolutionary Algorithm II (NSGA-II) [18], are widely accepted to be well-suited to solve such Multi-objective Optimization Problems (MOPs) [71]. The aim of MOEAs is to obtain a set of solutions, called the approximation set, such that all solutions are non-dominated and are close to the set of Pareto-optimal solutions. Here, a solution $\boldsymbol{x}_0$ dominates $\boldsymbol{x}_1$ ($\boldsymbol{x}_0 \succ \boldsymbol{x}_1$) in an MOP with m objectives if $\forall i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) \leq f_i(\boldsymbol{x}_1)$ and $\exists i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) < f_i(\boldsymbol{x}_1)$. The Pareto Set (PS) is $\mathcal{P}_S = \{\boldsymbol{x}_i | \neg\exists \boldsymbol{x}_j : \boldsymbol{x}_j \succ \boldsymbol{x}_i\}$ and the Pareto Front (PF) is $\mathcal{P}_F = \{(f_0(\boldsymbol{x}), \cdots, f_{m-1}(\boldsymbol{x})) \,|\boldsymbol{x} \in \mathcal{P}_S\}$.

A more complex type of MOPs is that of Multi-modal MOPs (MMOPs). In MMOPs, the goal is not to find just one, but multiple, if not all, (local) Pareto Sets (PSs). Each of these concerns a single *niche*, which is a subset of the search space, where a single mode resides. The PSs may, however, well map to the same Pareto Front (PF) in objective space. The difficulty for MOEAs to solve MMOPs lies in the fact that they have to perform a multi-modal search in order to uniformly sample the Pareto-optimal sets and front [78].

1.2.1 Diversity Preservation

In order to have MOEAs solve MMOPs, they need additional tools that prevent their premature convergence to a single solution or niche of the landscape and thus preserve the diversity of solutions throughout the entire search space [46]. One of such tools that allows MOEAs to solve MMOPs efficiently is an elitist archive, which was originally proposed for MOPs and significantly improves most MOEAs [77]. The inclusion of an elitist archive in SPEA suggested promising improvements over other algorithms in its capability to obtain solutions along the entire Pareto front [78]. Furthermore, the implementation of elitism in other MOEAs resulted in significant improvements over their basic version [77].

Most Multi-modal MOEAs (MMOEAs) do not explicitly model multiple approximation sets, but include diversity preserving techniques to ensure that solutions from multiple niches are maintained. The result of these MMOEAs is usually given in the form of a single approximation front, often derived from (a subset of) the elitist archive. A decision maker can then investigate this front by traversing all the solutions for desired trade-offs, in the intuitively implied navigational order from one extreme of the front to the other. However, the solutions are taken from several distinct niches of the fitness landscape. This results in observing a counterintuitive change in decision variable values when navigating the approximation front, similar to the counterintuitive changes of the dose distribution in HDR brachytherapy.

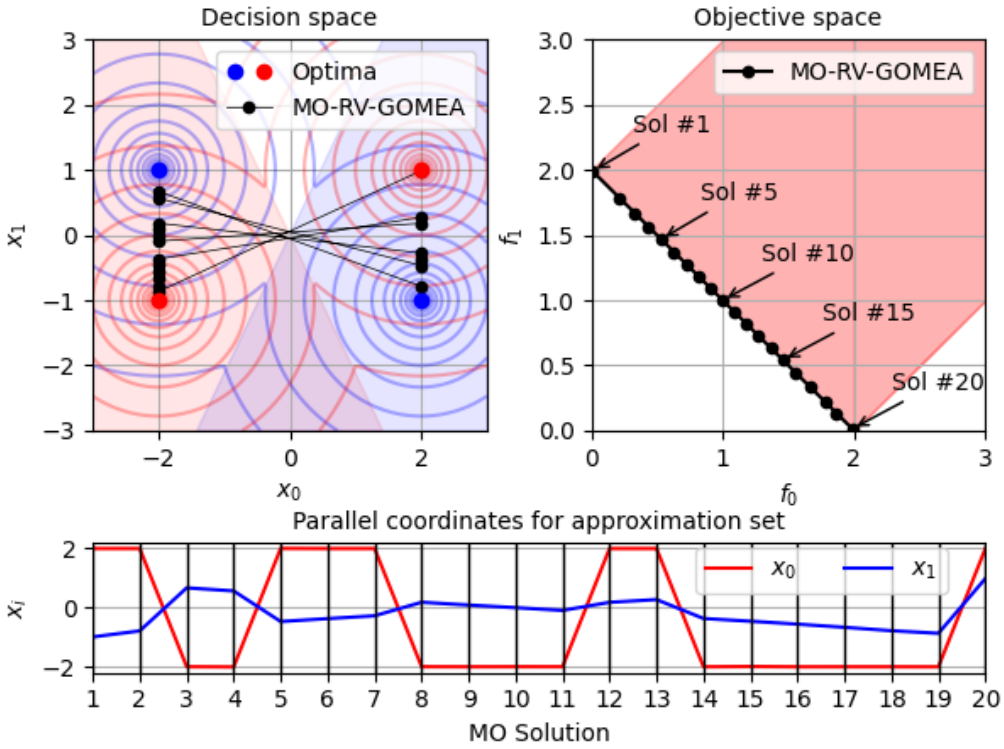


Figure 1.3: Approximation set and front with parallel coordinates plot as produced by MO-RV-GOMEA on MinDist problem. It is an instance of the genMed problem [8, 29] where the distance to k centres has to be minimized. Here $k = 2$, with $f_0(\boldsymbol{x}) = min(||\boldsymbol{x} - [1, -1]||, ||\boldsymbol{x} - [-1, 1]||)$ and $f_1(\boldsymbol{x}) = min(||\boldsymbol{x} - [1, 1]||, ||\boldsymbol{x} - [-1, -1]||)$. The shaded blue and red regions correspond to niches with global PSs.

1.2.2 Smoothness

Figure 1.3 shows an approximation set and corresponding front which contains solutions from both modes on the MinDist problem [47]. It demonstrates the counterintuitive changes in decision variable values in the parallel coordinates plot, i.e., if one were to traverse the front and inspect solutions. The figure shows that a front is found in objective space which looks to have approximated the PF to (near) optimality. However, the solutions jump around throughout decision space as seen by the abrupt changes in decision variable values in the parallel coordinates plot, thereby clearly denoting the unintuitive navigability of approximation sets. Especially in highly multi-modal problems, solutions can originate from many different modes, which may make it more insightful for decision makers to be able to request approximation sets for a specific number of modes.

1.2.3 Related Work

The issue of counterintuitive navigation along the approximation front has been explored in recent work, which introduced the new indicator-based MOEA for bi-objective optimization called BezEA [72]. A new problem formulation for population-based MOEAs was introduced where approximation sets were parameterized as Bézier curves. This ensured the navigational smoothness of an approximation set, whilst still being able to find good approximation sets when using the uncrowded hypervolume indicator [66] in order to calculate their fitness. By design, BezEA disallowed curves to dominate parts of themselves to ensure that the approximation set constitutes a single niche in the landscape. Thereby it foregoes the multi-modal search required to approximate most or all of the Pareto-optimal sets and corresponding front as it converges to a single niche.

The recently introduced population-based MMOEA called the Multi Objective Hill-Valley Evolutionary Algorithm (MO-HillVallEA) [47] niches the population into several clusters and showed promising results in maintaining multiple approximation sets. The authors took the concept of

Hill-Valley Clustering (HVC) for MOPs from the Hill-Valley Evolutionary Algorithm (HillVallEA) [48], which had been created for single-objective problems, and extended it to Multi-Objective HVC (MO-HVC) for MMOPs. MO-HillVallEA is capable of finding and preserving approximation sets, one for each niche, in parallel over time by considering the Pareto dominance relationship per niche. Figure 1.4 shows the various niches as a result of clustering a population of 5000 individuals with MO-HVC on the OmniTest problem [20]. Since MO-HillVallEA is a domination-based algorithm, it suffers from the fact that it converges further away from the PS than hypervolume based methods [6] like BezEA. Furthermore, no limit on the number of approximation sets can currently be set, which degrades the quality of the approximation sets in highly multi-modal problems as the population has to be divided over all niches. Figure 1.4 indicates the problem of dividing the population over all niches in highly multi-modal problems.

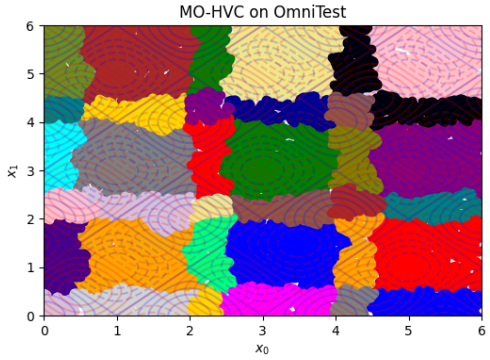


Figure 1.4: Figure showing result of MO-HVC on multi-modal multi-objective problem OmniTest for a population of 5000 solutions. Each colored area corresponds to a niche as found by MO-HVC. Transparent contour lines denote the fitness landscape of the two objectives.

1.3 Problem Statement

In this thesis, the notions of niching through (MO-)HVC and Bézier curve parameterizations will be combined. The use of niching allows to effectively search the multi-modal landscape, whereas the use of Bézier curve parameterizations not only enforces the smooth and intuitive navigability that is desired by decision makers, but also enforces each approximation set to be within a single niche. In order to combine these concepts, two sub questions of the first main research question have to be answered. First: "How can the concept of Hill Valley Clustering be applied to Bézier parameterizations of solution sets?". An EA also needs to be able to keep track of the found approximation sets in an elitist archive for diversity preservation, which results in the second sub question of: "How to preserve Bézier parameterizations for each mode in an elitist archive?" With both sub questions answered, the first main research questions of "How to combine the concepts behind MO-HillVallEA and BezEA in order to create an EA that is capable of finding all smoothly navigable approximation sets?" can be answered. The answer of this research question brought about a new algorithm called the Multi-Modal Bézier Evolutionary Algorithm (MM-BezEA).

The concept of (MO-)HVC tends to get problematic in highly multi-modal problems since it requires a very large population size to effectively search each mode [47]. Furthermore, a decision maker cannot inspect all of the approximation sets in such cases. This leads to the second research question: "How can a preset number of smoothly navigable approximation sets in highly multi-modal bi-objective optimization problems be found?" Each approximation set should constitute a single niche, but they should be dissimilar in order to preserve the diversity of solutions from various niches. For example, the preservation of diversity is desirable in optimization problems like the HDR brachytherapy problem, since the physicians can then choose a treatment plan that suits the patients anatomy well from one of the various approximation sets. Therefore, the following sub question has been formulated in order to answer the second main research question: "How to preserve variation between different solution sets, but enforce high-quality solution sets at the same time?". The answers to these questions have led to a new algorithm, called Set Bézier Evolutionary Algorithm (Set-BezEA), that can find a set of distinct yet high-quality approximation sets.

Both newly created algorithms will be tested on several benchmark problems to empirically test their behaviour and substantiate the answers to the two main research questions. They are benchmarked against a relatively new and much cited MMOEA called the Multi-Objective Particle Swarm Optimizer using Ring topology and Special Crowding Distance algorithm (MO_Ring_-PSO_SCD) [74] and against MO-RV-GOMEA [13] and MO-HillVallEA [47]. The indicators Pareto Set Proximity [74], Hypervolume [66] and Smoothness [72] will be employed for the results to determine if all Pareto sets are approximated, what the quality of the approximation sets in objective space is, and how smooth the approximation sets are.

The main goal is to create a new algorithm that is capable of providing smoothly navigable, yet distinct approximation sets for HDR brachytherapy. In order to evaluate whether this goal has been met, a comparison will be set up between the BRIGHT approach, that is currently used in the clinical practice, and what is researched in this thesis. MM-BezEA will explicitly be omitted from this comparison as it is intractable to obtain multiple distinct approximation sets with it, because of the requirement of very large population sizes for the effective niching with MO-HVC in highly multi-modal problems [47]. The third and last main research question is: "How does the new algorithm compare to BRIGHT on the HDR brachytherapy problem?" In order to provide an answer to the research question, a clinical evaluation of a few selected treatment plans is given by physicians. For seven patients, one plan from BRIGHT is selected together with one from Set-BezEA and visualized using commercial radiation therapy treatment planning software. The physicians can then determine, based on their expertise, if the plans from Set-BezEA are clinically acceptable and/or comparable to those from BRIGHT. One of the goals of the clinical evaluation is to answer the sub research question: "Are the plans from the newly created algorithm clinically acceptable?". Even though the treatment plans from the new algorithm can be clinically acceptable, BRIGHT might still produce better treatment plans. Therefore, the second sub question arises: "Does BRIGHT produce plans that are preferable in the clinical setting when compared to plans generated by the newly created algorithm?" Lastly, both BRIGHT and Set-BezEA are empirically benchmarked on several patients, where performance indicators like the hypervolume indicator [66] will be used to denote the results. Combined with the results from the clinical evaluation, the results from the empirical benchmarks can be used to substantiate the answer to the third main research question.

1.4 Overview of Research Questions

Research Question 1

How to combine the concepts behind MO-HillVallEA and BezEA in order to create an evolutionary algorithm that is capable of finding all smoothly navigable approximation sets?

(1.1) How can the concept of Hill Valley Clustering be applied to Bézier parameterizations of solution sets?

(1.2) How to preserve Bézier parameterizations for each mode in an elitist archive?

Research Question 2

How can a preset number of smoothly navigable approximation sets in highly multi-modal bi-objective optimization problems be found?

(2.1) How to preserve variation between different solution sets, but enforce high-quality solution sets at the same time?

Research Question 3

How does the new algorithm compare to BRIGHT on the HDR brachytherapy problem?

(3.1) Are the plans from the newly created algorithm clinically acceptable?

(3.2) Does BRIGHT produce plans that are preferable in the clinical setting when compared to plans generated by the newly created algorithm?

1.5 Overview of Thesis

In this thesis, the BezEA method is researched further and expanded upon. One part of the research focuses on the possibility of improving the BezEA method for the HDR brachytherapy problem, since the solutions are to become of equal quality as compared to MO-RV-GOMEA. Furthermore, several different techniques are researched which allow for the niching of different Bézier curves, thereby enabling the simultaneous optimization of Bézier curves that are in different modes of the fitness landscape. The result of combining these ideas is that each of the resulting curves is a distinct and high-quality approximation set that is perfectly smooth, and on which clinically acceptable treatment plans lie.

In order to combine the techniques of Bézier curve parameterizations and diversity preservation techniques into the two proposed new algorithms, several contributions are made in this thesis. However, before doing so, Chapter 2 introduces the reader to all concepts included in the new algorithms MM-BezEA and Set-BezEA. First, the chapter starts with a general introduction to the type of optimization problem that is solved, followed by a high level overview of an EA. Second, it introduces several diversity preserving methods required to find solutions pertaining to different niches. Third, it dives deeper into how certain variables can be grouped in order to gain information about the structure of problems. Fourth, the algorithms that have been combined into MM-BezEA and Set-BezEA are introduced, i.e. MO-RV-GOMEA, HillVallEA, MO-HillVallEA, and BezEA. Last, a detailed explanation of the HDR brachytherapy problem for prostate cancer is given.

In Chapter 3 some of the revisions that have been performed for BezEA are explained as a few issues have arisen during further research into the algorithm. In Chapter 4 the MM-BezEA algorithm and the concepts behind it are presented. In this chapter the focus is on the sub questions of Research question 1. In Chapter 5 the Set-BezEA algorithm and its newly defined fitness function are introduced, which a focus on the sub questions of Research question 2.

In Chapter 6 the empirical results for Set-BezEA and MM-BezEA on several benchmark problems are provided. These problems are used to clearly indicate the strengths and weaknesses of both algorithms, due to which Set-BezEA was chosen for the HDR brachytherapy problem. The results can be used to verify whether research questions 1 and 2 have been successfully answered. In Chapter 7 the results of Set-BezEA are compared to the current BRIGHT approach, in order to see if Set-BezEA could be the replacement for MO-RV-GOMEA as the optimization algorithm that is currently used in the clinic. The focus in this chapter is to answer the last research questions through the clinical evaluation by physicians as well as empirical benchmarks.

Lastly, Chapters 8 and 9 are the discussion and conclusion of this thesis as a whole.

Chapter 2

Background

In order to give information about which techniques are used in the algorithms created in this thesis, and how they compare to others, this chapter gives an overview of a part of the related research. First of the topics that will be explained is the types of problems that EAs can encounter. This is useful to begin with as algorithms can adapt themselves to the type of problem they come across. Furthermore, some basic definitions in the field of EAs will be introduced as well as a logical structure that can be used to find dependencies within the parameters of the encountered problems. These logical structures can be built up in various ways, of which one will be highlighted that builds a structure in an iterative fashion. Most of the information presented in this chapter is combined in an algorithm that is used at the core of the algorithms created in this thesis. This algorithm is called MO-RV-GOMEA [13], which is an adaptation from the discrete MO-GOMEA algorithm [45]. It has been created based on various concepts from the field of evolutionary computation, of which some of the most important concepts will be introduced.

2.1 Optimization Problems

Considering a specific problem, the goal in optimization problems is to optimize a given task by finding or approximating the best solution. This can mathematically be described as:

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \text{ or } \max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \tag{2.1}$$

The task that is to be solved can be evaluated with an *optimization function* $f(\boldsymbol{x})$. In biological terms, this is also called the *fitness function*. It is used to score a *solution* $\boldsymbol{x}$, and results in what is called an objective value.

2.1.1 Types of Optimization

Three types of optimization can be distinguished depending on what is known about the optimization problem. These types are as follows:

- Black Box Optimization
- Grey Box Optimization
- White Box Optimization

First, there is White Box Optimization, in which case everything about the problem at hand is known and sufficiently understood. Using this knowledge can be an advantage as it enables the building of problem-specific algorithms to analytically solve the problem to optimality. Second, there is Black Box Optimization (BBO). In BBO there is absolutely no knowledge of the problem besides an evaluation function, which has to be optimized. Thus, here the objective is to find as good a solution as possible in a minimal amount of time. Last, there is Grey Box Optimization (GBO), where the optimal solution still cannot be analytically determined with a problem specific algorithm. In general there is limited domain knowledge available in GBO. In certain cases, so-called partial evaluations are possible since the algorithm has knowledge as to how the problem decomposes into smaller parts, through which incremental updates become possible. Subsequently, these partial updates make the optimization more efficient and faster.

Out of all three of these types, only the cases of BBO and GBO will be considered in this thesis. This is due to the fact that, in the case of WBO, problem specific algorithms can be created by hand to solve the problem to optimality. For HDR brachytherapy it is impossible to solve the problem to optimality due to the dimensionality of the problem and its multi-objective and multi-modal nature. The dimensionality makes it computationally expensive to evaluate and the multi-objective and multi-modal nature makes it so that there are multiple sets of best possible solutions for which no analytical formula exists to compute these. Thus, for the cases of GBO and BBO, in which HDR brachytherapy falls, more general optimization algorithms need to be used to quickly and efficiently search for a good solution.

2.1.2 Search Spaces

Optimization problems have a so-called *search space* $\mathcal{X}$. It denotes the domain and its boundaries in which an algorithm has to look for solutions, e.g. $\boldsymbol{x} \in \mathcal{X}$. Many forms of search spaces exist. For example, when trying to play a game of chess, the search space is the set of all legal moves. Likewise there are many problems, and games, for which different search spaces can be defined, but some of the more common are:

- Binary
- Permutation
- Real-valued

Binary and permutation search spaces are not used in this thesis, but simple examples of them could be the well-known 0-1 knapsack problem and the travelling salesman problem.

Real Valued Search Spaces In the case of the real-valued search space the domain of the search space is $\mathbb{R}$. A more mathematical notation is $\mathcal{X} \subseteq \mathbb{R}^\ell$. A solution $\boldsymbol{x}$ can take on any of the values in the domain of $\mathbb{R}$, meaning that the search space can be infinitely large. Real-valued search spaces will be used in the next sections to explain more about optimization problems. Furthermore, problems with this type of search space is what the proposed algorithms are benchmarked on at the end of this thesis.

Constraints Besides having to find a solution $\boldsymbol{x}$ out of the search space $\mathcal{X}$ that is the best for a particular optimization function $f(\boldsymbol{x})$, many problems are also restricted as to what values $\boldsymbol{x}$ can take on within $\mathcal{X}$. This comes down to constraining what $\boldsymbol{x}$ can be, stating that it has to be subject to a certain condition. Mathematically this can be described as:

$$\begin{aligned} \min_{x\in\mathcal{X}} \text{ or } \max_{x\in\mathcal{X}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 && \text{for } i = 1, 2, ... \\ & h_j(x) = 0 && \text{for } j = 1, 2, ... \end{aligned} \tag{2.2}$$

In the equation, two *constraint functions*, $g_i(\boldsymbol{x})$ and $h_j(\boldsymbol{x})$, are present. These are called inequality constraints in the case of $g_i(\boldsymbol{x})$ and equality constraints for $h_j(\boldsymbol{x})$. For example, if the search space $\mathcal{X}$ is viewed as a 2D square, then the search space with a constraint can be visualized as a 2D square where one of the corners is cut off as all solutions in that corner are infeasible. A solution $\boldsymbol{x} \in \mathcal{X}$ is then said to be *feasible* if it adheres to all constraints for $g_i(\boldsymbol{x})$ and $h_j(\boldsymbol{x})$. On the other hand, if one of the constraints is not satisfied, then $\boldsymbol{x}$ is said to be *infeasible* for the problem.

2.1.3 Multi-Modal Multi-Objective Problems

Multi-Modal Multi-Objective Problems (MMOPs) are a special type of optimization problem with certain characteristics. MMOPs are the type of problems for which the methods researched in this thesis are designed. In order to explain what MMOPs exactly are, some more simple definitions and optimization problems need to be described first.

Single-Objective Problems

When $f(\boldsymbol{x})$ produces a single objective value, algorithms are said to operate in a single-objective (SO) setting. Here, the goal is to find a solution $\boldsymbol{x}$ that has the smallest possible value for $f(\boldsymbol{x})$.

Uni-Modality. SO uni-modal problems have just one global minimum and no other local minima, meaning that they have exactly one solution with the best possible objective value in the entire search space. The left diagram of Figure 2.1 shows a single-objective uni-modal problem. In this particular case, the goal is to find a solution as close to the point $[1, 1]$ as possible, which is also its global minimum. The objective function calculates the Euclidean distance to the specified point, which automatically results in lower, and thus better, objective values when a solution $\boldsymbol{x}$ is closer to $[1, 1]$.

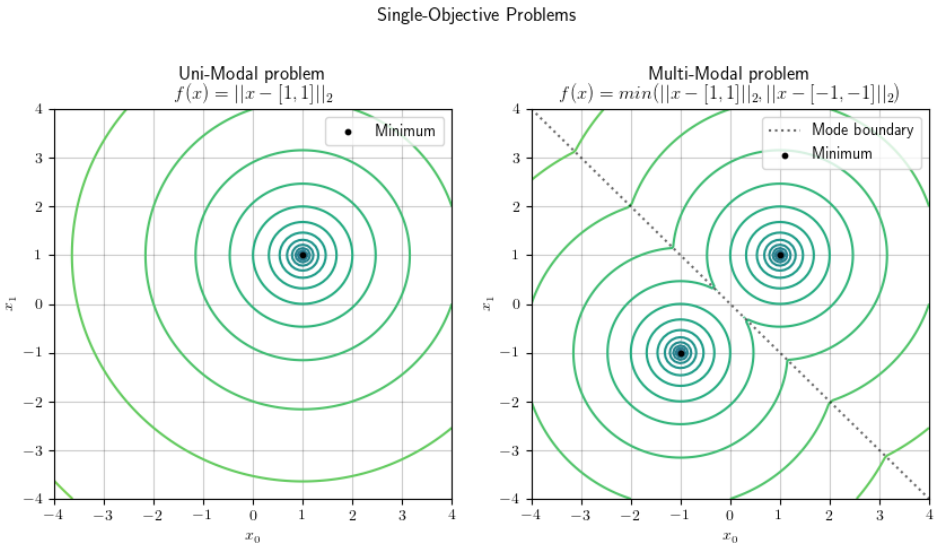


Figure 2.1: Uni- and Multi-Modal Single-Objective Problem.

Multi-Modality. SO multi-modal problems are more elaborate version as compared to their uni-modal counterpart. The objective function $f(x)$ still operates in the SO setting, but now it has at least one global minimum with possible local minima as well. A local minimum can be described as a solution that has the lowest possible objective value out of its set of neighbouring solutions, but its objective value is worse than the global minimum. However, there can also be multiple global minima, meaning that there are several solutions with the same globally minimal value for the objective function $f(\boldsymbol{x})$.

Given a minimization problem $f : \mathcal{X} \to \mathbb{R}$, a formal definition for a local and global minimum of $f(\boldsymbol{x})$ is:

Definition 1. *Local Minimum [17]: A local minimum $\boldsymbol{x}_l \in \mathcal{X}$ of a single-objective function $f : \mathcal{X} \to \mathbb{R}$ is an input vector $\boldsymbol{x}_l$ with $f(\boldsymbol{x}_l) < f(\boldsymbol{x}_i)$ for all $\boldsymbol{x}_i$ neighbouring solutions of $\boldsymbol{x}_l$ within a range ϵ.*

Definition 2. *Global Minimum [17]: A global minimum $\boldsymbol{x}_g \in \mathcal{X}$ of a single-objective function $f : \mathcal{X} \to \mathbb{R}$ is an input vector $\boldsymbol{x}_g$ with $f(\boldsymbol{x}_g) \leq f(\boldsymbol{x}_i) \forall \boldsymbol{x}_i \in \mathcal{X}$.*

The right diagram of Figure 2.1 shows a single-objective multi-modal problem. It shows a problem with two global minima with equal optimal values for $f(\boldsymbol{x})$. The goal is similar to the one of the uni-modal version, but now there are two points to which the distance has to be minimized, namely $[-1, -1]$ and $[1, 1]$. The entire search space can also be divided into several modes, one around each minimum. The separation of the two modes in the example can be seen as the dotted line that cuts diagonally through the search space.

Multi-Objective Problems

In many cases an optimization problems does not have just one single objective that has to be optimized. On the contrary, many real-world problems are multi-objective (MO) in nature. In MO problems, the fitness function scores solutions based on multiple objectives. Each of these m objectives will be denoted as $f_i(\boldsymbol{x}), i \in \{0, 1, ..., m - 1\}$.

Uni-Modality. One of the difficulties in the MO setting is that certain objectives can be contradicting. This can increase the difficulty as there is a trade-off between optimizing one objective over another. A definition will be introduced later in order to define when a solution is better in the MO setting. The goal for an EA in MO optimization is to find the best trade-off between all of the objectives. This trade-off is usually given to the decision makers in the form of an approximation front in order for them to select one or more solutions. This approximation front depicts all of the solutions that the EA has found and that have the best trade-off between the various objectives. The left plot of Figure 2.2 shows a uni-modal multi-objective problem. The black line depicts the set of all solutions with the most optimal trade-off between the two objective functions f_0 and f_1.

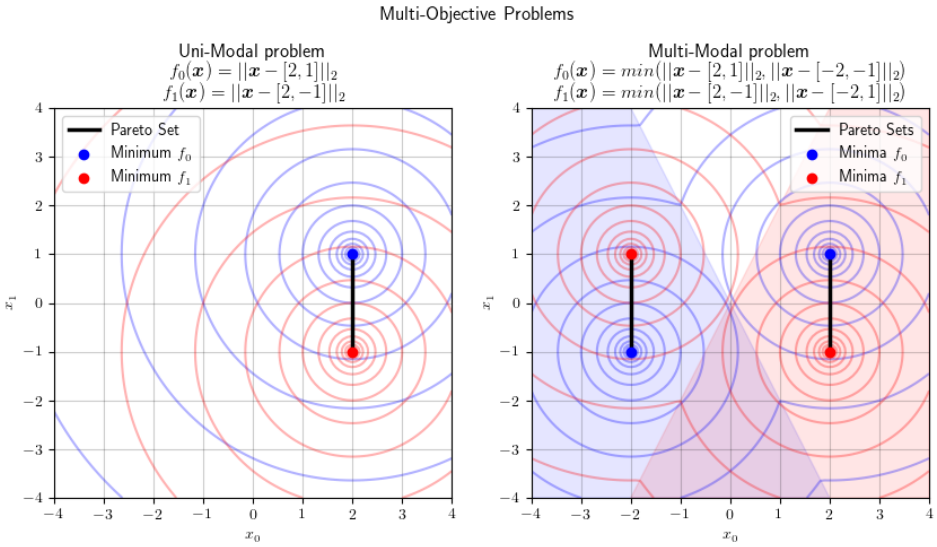


Figure 2.2: Uni- and Multi-Modal Multi-Objective Problem. Mode boundaries are shown as shaded areas for corresponding objective, depicted through the contour lines.

The contradicting objectives in MO problems can cause repercussions in the form of an improvement in one objective and subsequent worsening of one or more of the other objectives. Thus, there usually is not just one solution with the best objective values, but many solutions with different trade-offs. Therefore, another notion of a 'better' solution is required, as there are now multiple values on which solutions are scored. In order to rank solutions in the MO setting and determine which is a better solution, the concept of Pareto dominance is taken from economists amongst other definitions related to Pareto dominance. The definition of Pareto dominance between two solutions $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ is defined as follows:

Definition 3. *A solution $\boldsymbol{x}_0$ is said to Pareto dominate a solution $\boldsymbol{x}_1$, also denoted $\boldsymbol{x}_0 \succ \boldsymbol{x}_1$, if and only if:*

$$\forall i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) \leq f_i(\boldsymbol{x}_1) \quad \wedge$$
$$\exists i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) < f_i(\boldsymbol{x}_1)$$

Using the definition of Pareto dominance from above, other closely related definitions can also be given. First, Pareto optimality is defined as:

Definition 4. *A solution $\boldsymbol{x}_0$ is said to be Pareto optimal if and only if $\neg\exists \boldsymbol{x}_1 : \boldsymbol{x}_1 \succ \boldsymbol{x}_0$.*

Second, the Pareto set $\mathcal{P}_S$ of all Pareto optimal solutions can then be defined as:

Definition 5. *The Pareto set $\mathcal{P}_S = \{\boldsymbol{x}_i | \neg\exists \boldsymbol{x}_j : \boldsymbol{x}_j \succ \boldsymbol{x}_i\}$.*

Lastly, there is the definition of the Pareto front. The set $\mathcal{P}_F$ is the set of all objective function values corresponding to the solutions in $\mathcal{P}_S$:

Definition 6. *The Pareto front $\mathcal{P}_F = \{(f_0(\boldsymbol{x}), f_1(\boldsymbol{x}), ..., f_{m-1}(\boldsymbol{x})) \, | \boldsymbol{x} \in \mathcal{P}_S\}$.*

Multi-Modality. Just as in single-objective problems, multi-objective problems can have multiple global or possibly local optima. Their definitions are similar to those for the single-objective problems, but use the notion of Pareto dominance. The right plot of Figure 2.2 shows such a multi-modal multi-objective problem. It depicts two Pareto sets, each pertains to a single niche or mode of the fitness landscape.

2.2 Evolutionary Algorithms

The information given in this section explains how EAs work. EAs mimic abstractions of key features in natural evolution that are assumed to be responsible for the creation of the complex life forms known today [70]. Therefore, most mathematical names also have biologically inspired counterparts, which will be introduced next to the mathematical terms.

Several distinctions for EAs exist. For example, if the algorithm is designed to operate in an MO problem setting, they are usually denoted as Multi-Objective Evolutionary Algorithms (MOEAs). In case an EA is designed to operate in an MO and MM setting, they are denoted as Multi-modal Multi-Objective Evolutionary Algorithms (MMOEAs).

Solutions. A solution is synonymously called an *individual* or a *genotype*. It comes from the set of all feasible solutions S, which can be described as the *search space* of the EA. An *individual* is defined as a vector or string of $\ell \geq 1$ *symbols* or *alleles*. The location of a symbol is called the *locus* and its container a *gene* or *variable*. Each of these symbols can take on a value depending on the search space S.

Scoring Individuals. In order to then solve the given problem, all types of EAs use a *population* $\mathcal{P}$ of $n \geq 2$ individuals $x^i, i \in \{0, 1, ..., n-1\}$. Each of these are scored based on the fitness function, whereby better fitness is preferred. The fittest individuals can then be used to create new individuals for the next generation. This is the so-called "Survival of the fittest" of the Darwinistic view on nature.

In a binary search space with a single objective, the EA tries to find individuals among its population with the highest objective value. On the contrary, there is a rather confusion convention in this field of research in the case of a real-valued search space. In such a search space the EA usually tries to find the solution with the smallest objective value instead of highest, i.e. minimization. Throughout this thesis, the notion of fitness in terms of the real-valued search space will be used. Thus, a solution will from now on be defined fitter if it has a lower fitness values as per the Pareto Dominance relation from Definition 3.

Finding Fit Individuals. In order to change the individuals and hopefully find better solutions, EAs run several generations of the population. In each generation, the EA performs a variation operator on or between individuals, thereby generating new individuals. These new individuals are called offspring, and together with the individuals at the start of the generation form the population at the end of a generation. Then, a selection operator is used to prune the population back to its size n. Another method is to first perform selection on the population, after which the offspring is created from the selected individuals to expand the population back to size n. Selection happens based on the previously mentioned fitness values, in order to let the 'fittest' individuals survive.

The entire process of variation and subsequent selection for the new population is repeated for several generations. It is run until an EA reaches some pre-determined termination criteria. These can be in the form of having spent a certain amount of computation time or limiting the maximum number of evaluations of the fitness function. In terms of pseudo code, the overall setup

of an evolutionary algorithm can be defined as seen in Algorithm 1.

Algorithm 1: Pseudocode for an Evolutionary Algorithm

```
1 Generation t ← 0;
2 Population P^t ← createAndEvaluateInitialIndividiuals(n);
3 while terminationCriterionNotSatisfied(P^t) do
4 |   Offspring O^t ← createAndEvaluateOffspring(P^t);
5 |   Population P^{t+1} ← selectSurvivingIndividuals(P^t, O^t);
6 |   Generation t ← t + 1;
7 end
```

Concluding all of the previous, using a population in which solutions are altered through the selection and variation operators allows the algorithm to search through the search space. It allows the method to be effective, efficient, simple and robust through mimicking the concept of natural evolution [70].

2.3 Diversity Preservation Methods

EAs need various diversity preservation methods when they are required to find multiple high-quality solutions with different properties, i.e. parameter values. These methods are required for multi-objective, multi-modal or both multi-objective and multi-modal problems [46]. Many of such tools exist, like an elitist archive that was designed for multi-objective problems andbalanced k-leader means. significantly improves most MOEAs [77]. Another tools is that of niching for multi-modal problems, whereby a population can be clustered in several ways to have these focus on separate parts of the search space. If these diversity preservation tools are not included, EAs tend to converge towards a single solution due to three distinct effects [46]:

- **Selection pressure.** The selection operator results in an expectation of disappearance of lower fitness solutions from a population.
- **Selection noise.** Random solutions have to be chosen among solutions of equal fitness in a population.
- **Operator disruption.** The application of crossover and mutation operators can result in the direct destruction of good solutions.

2.3.1 Elitist Archive

The inclusion of an elitist archive $\mathcal{E}$ in SPEA [78] suggested promising improvements over other algorithms in its capability to obtain solutions along the entire Pareto front. Such an improved capability of searching is of course desired for effective search by EAs. The notion of elitism can be described rather simply, the elitist archive is a set that holds only those solutions from the population that are non-dominated.

After each generation of an EA, the elitist archive is updated to hold the best found solutions so-far. If the population holds a solution that has improved enough to now dominate several solutions from the elitist archive, the dominated solutions from the elitist archive are removed and the improved solution from the population is added to the elitist archive.

Even though the inclusion and application of an elitist archive seems rather straightforward, a key questions arises [77]:

- Which solutions are kept in the elitist archive and for how long to preserve diversity but not let the archive grow unbounded?

In order to answer this question, several techniques have been created and researched previously, which usually entails discretization methods. Objective space discretization can come in various forms, of which adaptive grid discretization [41] and greedy Hypervolume Subset Selection (gHSS) [28] are two examples. Besides the objective space, discretization can also be performed in the decision space, like adaptive discretization techniques that take the standard deviation of solutions or the estimated probability density function of solutions into account [36].

2.3.2 Niching

Niching [37] is another of many diversity-preserving tools used by MMOEAs to effectively and simultaneously search for solutions near the (local) PS in each mode of a multi-modal optimization problem. It has been applied to MMOEAs in the form of the multi-objective particle swarm optimization using ring topology and special crowding distance (MO_Ring_PSO_SCD) algorithm [74] and Omni-optimizer [20] among others. Furthermore, niching can also be used to find distinct clusters of solutions so that each of these can focus on a different part of the approximation front, which enables MMOEAs to effectively search along the entire approximation front [56].

Since the number of modes is usually unidentified beforehand, niching methods need to be able to adapt to the specific number of modes for an MMOP, . Many such niching methods exist like the Hill-Valley Clustering [48] and Multi-Objective Hill-Valley Clustering [47] methods that cluster a population, but also crowding methods that restrict the mating pool for new solutions like in the Decision space Niching NSGA-II (DN-NSGA-II) [38]. Other algorithms like MO_Ring_PSO_SCD and Omni-optimizer employ crowding methods that look at both both objective and decision space diversity when performing selection on the population.

Fitness sharing [31] is another niching technique that is aimed at keeping the individuals of a population diverse. By degrading the fitness values of solutions if they are close to one another in decision space it tries to preserve their diversity. Herein, the difficult part is defining which individuals are to share their fitness values. For this purpose, a niche radius parameter σ_{share} can be defined as to which solutions are close to one another. A scaling factor α is also used that degrades the fitness values [27] of those solutions, by multiplying the fitness values with α. However, it has been proven hard to set correct values for these parameters [15].

2.3.3 Adaptive Steering

In (M)MOP's the size and shape of the Pareto front can vary largely depending on the problem at hand. Most (M)MOEAs try to find improvements spread over the entire Pareto front, whereas in some situations only a part of the Pareto front is interesting to the decision maker. In order to steer algorithms towards the interesting parts of the Pareto front, Adaptive Steering [2] can be used. It can be applied to both the population and elitist archive.

By constraining solutions as to what fitness values they are allowed to obtain, the algorithm can disregard certain solutions. This allows the algorithm that uses adaptive steering to disregard those parts of the Pareto front in which it is disinterested. The interesting part of the Pareto front can then be searched through more thoroughly as the elitist archive can spread its limited number of solutions in the interested region.

2.4 Linkage Modeling

In both the GBO and BBO settings, solution variables may be dependent on each other. These dependencies have to be modelled as the entire problem is not sufficiently understood in order to analytically determine these. Using such dependencies in a crossover operator, the so-called *building blocks* of interdependent variables can be preserved from disruption in the creation of new solutions [63].

Various methods for modeling these dependencies exist for EAs, where the algorithms that exploit these learned structures are called Model-Based EAs (MBEAs). As these MBEAs usually do not know the structure a priori, it has to be learned during the search for the solution. However, sometimes the structure is known a priori and a linkage model can be built beforehand. The focus of this section will be on structures that can be used and exploited by the (MO-)RV-GOMEA algorithm, since RV-GOMEA and MO-RV-GOMEA form the basis for the new algorithms that are researched in this thesis.

2.4.1 Family of Subsets

A common logical structure for modeling the dependencies and learning them during the search is the so-called Family Of Subsets (FOS). A FOS can be described as follows:
Given ℓ problem variables $x_0, x_1, ..., x_{\ell-1}$ and S the set of all variable indices $\{0, 1, ..., \ell - 1\}$, a

FOS F is a subset of the powerset of S, i.e. $F \subseteq \wp(\mathbf{S})$. Thus, $F = \{\mathbf{F}^0, \mathbf{F}^1, ..., \mathbf{F}^{|F|-1}\}$, where $\mathbf{F}^j \subseteq \{0, 1, ..., \ell - 1\}$, $j \in \{0, 1, ..., |F| - 1\}$.

Within the FOS model structure, a few different types can be distinguished. The type that is used in this thesis is the Linkage Tree (LT) FOS [63], which is a hierarchical structure of the variables. This type has a group $\mathbf{F}^i = \{0, 1, ..., \ell - 1\}$, in which all variables are present. It also has ℓ groups $\mathbf{F}^i = i$, for $i \in \{0, 1, ..., \ell - 1\}$, one for each variable. For any subset $\mathbf{F}^i$ with more than one variable, there are subsets $\mathbf{F}^j$ and $\mathbf{F}^k$ such that: $\mathbf{F}^j \cap \mathbf{F}^k = \emptyset$, $|\mathbf{F}^j| < |\mathbf{F}^i|$, $|\mathbf{F}^k| < |\mathbf{F}^i|$ and , $\mathbf{F}^j \cup \mathbf{F}^k = \mathbf{F}^i$.

2.4.2 Linkage Tree

In order to obtain a linkage model, several methods have been proposed in order to identify the unknown structure that lies under the optimization function. The problem of identifying such a structure can be seen as a form of hierarchical clustering, whereby more similar variables are hierarchically clustered. This process starts with individual variables and groups them until entire building blocks have been identified. A notion of similarity, or its opposite form distance, between all of the variables is required in order to identify the actual building blocks. During the learning of these dependencies the algorithms can, however, make mistakes. These mistakes can be divided into two types [59]:

- **Missing Linkage.** Mistakenly not identifying a dependency between two variables, thereby falsely omitting it from the LT.
- **False Linkage.** Mistakenly identifying a dependency between variables where it does not occur.

Of the mistakes mentioned above, the notion of false linkage occurs rather often. This is due to an assumption made that two or more dependencies exists. These assumptions come forth from the fact that selection is supposed to expose genotype regularities, which in turn are supposed to be captured by statistical analysis. Therefore, the actual size of a population is rather important as smaller populations tend to converge faster and have less variance, thus allowing them to more quickly give rise to false linkages.

Identifying Linkage In order to actually capture the dependencies between variables, many EA's employ so-called Dependency Structure Matrices (DSM). These matrices are square matrices in which each entry indicates the dependency between two variables. For example, a single DSM entry $d_{i,j} \in \mathbb{R}$ denotes the relationship between the i^{th} and j^{th} variables. The DSM will be used to build the linkage models.

Hierarchical Clustering To find the dependencies and construct a LT [63], the decision variables have to be hierarchically clustered. The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [55] is such a hierarchical clustering method.

Using this hierarchical clustering method the LT FOS results in $2\ell - 1$ linkage sets $\mathbf{F}^i$. The first ℓ subsets are the singleton sets like in a univariate FOS structure. The remainder of the subsets are created through the merging of the existing, unmerged subsets using the UPGMA method.

2.5 Multi Objective Real Valued Gene-pool Optimal Mixing Evolutionary Algorithm

The Multi-Objective Real-Valued Gene-pool Optimal Mixing Evolutionary Algorithm (MO-RV-GOMEA) [13] is an EA that is designed to operate in MO Real-Valued optimization problems. It is a so-called Estimation-of-Distribution Algorithm (EDA), as it tries to estimate a Gaussian distribution over the search space from which it can sample solutions. Furthermore, MO-RV-GOMEA is a continued development and partial combination of the discrete Multi-Objective Gene-pool Optimal Mixing Evolutionary Algorithm (MO-GOMEA) [45] and the Multi-Objective Adapted Maximum-Likelihood Gaussian Model Iterated Density Estimation Evolutionary Algorithm (MA-MaLGaM) [8]. It uses key concepts from both of these algorithms, which will be elaborated on in the following sections.

2.5.1 Overview of Algorithm

Before the key concepts of MO-RV-GOMEA are explained, an overview of the method is given for some guidance. MO-RV-GOMEA uses a population $\mathcal{P}$ of n individuals together with an elitist archive $\mathcal{E}$, which holds all of the non-dominated solutions. The population is initiated with n solutions sampled uniformly at random within a specified range, which can be smaller than the range of the search space.

During the optimization, several generations are performed until MO-RV-GOMEA reaches a predefined termination criterion. This can be anything in the form of a runtime budget, evaluation budget or a certain fitness value (solution quality) that is to be achieved. In each generation, including right after the initialization, selection is performed to obtain a selection set $\mathcal{S}$ of size $\lfloor \tau n \rfloor$, where $\tau \in [\frac{1}{n}, 1]$. The set $\mathcal{S}$ is chosen from the best solutions of the population $\mathcal{P}$, after each of the solutions is ranked using the well-known non-dominated sorting method [18]. Then, using the selection set $\mathcal{S}$, new values for the decision variables are sampled. These values can be used to alter the solutions in the population $\mathcal{P}$ in order to improve them towards the optimum. Furthermore, the non-dominated solutions from the elitist archive, or a subset of them, are added to the population before sampling decision variables. This is to improve the quality of the offspring created in each generation.

2.5.2 Clustering

The selection set $\mathcal{S}$ is clustered in each generation. MO-RV-GOMEA obtains q equal-sized clusters $\mathbf{C}_k$ using an efficient leader-based clustering procedure [8], which uses normalized Euclidean distances between solutions in objective space. The first m clusters are obtained from the $n_c = \frac{2}{q}|\mathcal{S}|$ best solutions ranked for each of the m objectives. These single-objective clusters are used to direct various parts of the approximation front outward to provide a better approximation of the true Pareto front. The remaining $q - m$ clusters are created using leader solutions that are heuristically selected according to a distance based heuristic [8]. The leader solutions are used as the means in a K-means-like clustering step. For each of these clusters, separate distributions are sampled to enable the efficient search of each part of the approximation front [56]. Figure 2.3 shows how this clustering step is used to improve certain parts of the approximation front.

For sake of simplicity and the fact that this clustering step is not used in one of the algorithms created in this thesis, it will be omitted from the remainder of the following subsections. Furthermore, the Real-Valued Gene-pool Optimal Mixing Evolutionary Algorithm (RV-GOMEA), which is the single-objective counterpart to MO-RV-GOMEA and is also used in one of the algorithms created in this thesis, uses the concepts described below without clustering.

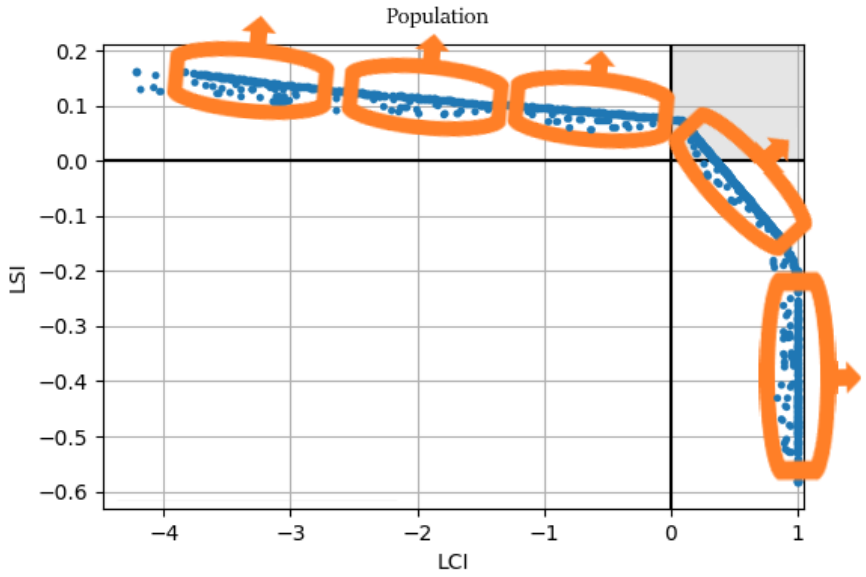


Figure 2.3: Example of the clustering step on the population that is employed in MO-RV-GOMEA in order to further optimize certain parts of the approximation front [13].

2.5.3 Partial Variation and Gene-Pool Optimal Mixing

A very important and key aspect of MO-RV-GOMEA is its ability to perform so-called partial variations, and possibly subsequent partial evaluations if it is operating in the GBO setting. So far, there is the notion of what variables can be grouped together using the LT FOS structure. As these variables are assumed to be dependent on one another, it can also be assumed that if one of the variables in a dependent group changes, all of them should be changed to not disrupt the building blocks. For this reason, the Gene-pool Optimal Mixing (GOM) operator has been introduced [64].

GOM employs the use of the LT FOS, where it models the solution variables in each linkage set $\mathbf{F}^i$ with a joint multivariate Gaussian distribution $N_{\mathbf{F}^i}$. For each of these subsets $\mathbf{F}^i$ it performs partial variation on all of the solutions in the population. More specifically, in partial variation only those decision variables for which the subset $\mathbf{F}^i$ is sampled are changed, whereas the other decision variables are left untouched.

The GOM operator can be described as a partial variation operator. By applying the operator on all solutions $\mathbf{x}$ in the population it performs these partial variations, after which it evaluates the solutions. The changes are only kept if the partially altered decision variables values of a solution $\mathbf{x}$ result in an improved fitness score. If the fitness has not improved, the decision variable values of $\mathbf{x}$ are reverted to the state before the GOM operator was applied. This means that only those changes that improve the fitness of a solution $\mathbf{x}$ are kept, from which it derives its name gene-pool optimal mixing. An pseudocode example of the GOM operator can be seen in Algorithm 2. Although the actual operator in MO-RV-GOMEA also contains parts related to the clustering of the population and other mechanism, these have been left out for ease of reading and to better grasp the optimal mixing part.

The GOM operator is used to apply partial variations for a single subset $\mathbf{F}^i$ of the LT FOS to all solutions of the population. Then, the algorithm moves on to the next subset $\mathbf{F}^j$, for which the GOM operator is again applied to all of the solutions. The order in which the subsets of the LT FOS are traversed is randomized in each generation of the EA.

Algorithm 2: $[\boldsymbol{x}_i]$ = `GenepoolOptimalMixing`$(\boldsymbol{x}_i, \boldsymbol{F}^j)$ [13]

Input: Solution $\boldsymbol{x}$, FOS subset $\boldsymbol{F}^j$
$\boldsymbol{b} \leftarrow \boldsymbol{x}$
$\boldsymbol{y}_{\boldsymbol{F}^j} \leftarrow$ SampleDistribution$(\boldsymbol{N}_{\boldsymbol{F}^j})$
$\boldsymbol{x}_{\boldsymbol{F}^j} \leftarrow \boldsymbol{y}_{\boldsymbol{F}^j}$
EvaluateSolution$(\boldsymbol{x})$
accept $\leftarrow \boldsymbol{x} \succ \boldsymbol{b}$ **or** $\neg$IsDominatedByArchive$(\mathcal{A}, \boldsymbol{x})$
if $\neg$*accept* **then**
  $\boldsymbol{x}_{\boldsymbol{F}^j} \leftarrow \boldsymbol{b}_{\boldsymbol{F}^j}$
end
return $\boldsymbol{x}$

2.5.4 Adaptation of Distribution Multipliers

As MO-RV-GOMEA uses distributions from which it samples new decision variables, it employs some ideas from MAMaLGaM in order to adapt these distributions and aid the search. These techniques [8] are called Adaptive Variance Scaling (AVS), which is triggered by the Standard Deviation Ratio (SDR). By working together, both SDR and AVS are aimed at keeping the variance in the population high enough so that the selection pressure does not diminish the variance when improvements are found far away from the mean, which could lead to premature convergence. The AVS scales the covariance of a joint multivariate Gaussian distribution of a FOS element $N_{\mathbf{F}^i}$ by multiplying it with the associated multiplier $c^{\text{mult}}_{\mathbf{F}^i}$.

The distribution multiplier is scaled up to increase the variance if the improvement of solutions happens far away from the mean of a distribution at a distance of more than one standard deviation, i.e. the SDR. The scaling happens with a multiplier $\eta^{\text{inc}} > 1$. In case solutions do tend to improve near the mean of the distribution, no scaling is applied to the distribution multiplier. Contrarily, if no improvements can be found, the covariance might be too large. This results in the necessity of decreasing the distribution multiplier, which is achieved by scaling down the multiplier with $\eta^{\text{dec}} \in (0, 1)$ until it reaches it neutral value of 1.

In case the distribution multiplier is at its neutral value of 1, further down scaling is prohibited to, once again, prevent premature convergence. However, if enough generations pass whereby the algorithm decides it is supposed to downscale the distribution multiplier, it is allowed to go below 1. The algorithm decides to down scale further if more than NIS^{MAX} (maximum no improvement stretch) generations have passed.

2.6 (Multi-Objective) Hill-Valley Evolutionary Algorithm

The Hill-Valley Evolutionary Algorithm (HillVallEA) and the Multi-Objective HillVallEA (MO-HillVallEA) are two algorithms from which the algorithms created in this thesis lend some clustering concepts. Both of these algorithms use a rather simple concept to find all modes in a multi-modal optimization problem. However, even though the clustering concept is rather simple, it has shown to be remarkably effective by virtue of the HillVallEA winning the GECCO 2019 Competition on Niching Methods for Multimodal Optimization [49]. The names of the algorithms already imply that they have been created for both the single- and multi-objective setting, whereby the multi-objective version is a continued development of the single-objective HillVallEA.

2.6.1 Hill-Valley Clustering

The method of Hill-Valley Clustering (HVC) is a two-stage niching method. It clusters and evolves the population for single objective multi-modal optimization problems, by running both stages in each generation. The first stage is used to locate each of the distinct niches. Next, the second stage initializes a core search algorithm for each of the niches. At the heart of the HVC method is the Hill-Valley Test (HVT) [67], which can be utilized to determine whether two solutions reside in the same niche.

The HVT decides if two solutions belong to the same niche by first drawing an edge between two solutions $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the search space. Along this edge, N_t evenly spread solutions are evaluated, called the test points. The value of N_t is determined by the Euclidean distance between the two

solutions divided by the expected edge length (EEL). In turn, the EEL is calculated by dividing the size of the search space by the number of solutions in the population. The formula for the EEL and N_T are as follows:

$$EEL = \left(\frac{V_X}{|\mathcal{P}|}\right)^{\frac{1}{\ell}} \tag{2.3}$$

$$N_t = 1 + \left\lfloor \frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||}{EEL} \right\rfloor \tag{2.4}$$

Here, ℓ is the problem dimension, V_X the volume of the search space and $\mathcal{P}$ the population.

If any of the N_t test points along the edge between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ have a fitness that is worse than that of either solutions at the endpoints of the edge, the method detects that there is a hill in between them. Consequently, the two solutions $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are concluded to be in separate clusters. Contrarily, if all N_t points along the edge have equal or better fitness values than both $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, both solutions are concluded to belong to the same valley. In this case they are, of course, to be clustered together. A pseudocode example of this test can be found in Algorithm 3 with a visual depiction of a HVT between two solutions in Figure 2.4.

Algorithm 3: [boolean] = `HillValleyTest`($\boldsymbol{x}_i$, $\boldsymbol{x}_j$, N_t, f) [67]

Input: Solutions $\boldsymbol{x}_i$, $\boldsymbol{x}_j$, $N_t \in \mathbb{Z}$, objective function f

for $k = 1, ..., N_t$ **do**
  $\boldsymbol{x}_{test} = \boldsymbol{x}_i + \frac{k}{N_t + 1}(\boldsymbol{x}_j - \boldsymbol{x}_i)$
  if *max*$\{f(\boldsymbol{x}_i), f(\boldsymbol{x}_j)\} < f(\boldsymbol{x}_{test})$ **then**
    return false
  end
end
return true

In order to determine in which order the solutions are to be clustered, HillVallEA uses the concept of the Nearest Better Tree [57]. First, it sorts all the solutions in the entire population in descending order of fitness. By definition, the first solution $\boldsymbol{x}_0$ automatically forms the first cluster $\mathcal{C}_0 = \{\boldsymbol{x}_0\}$. Second, the Hill-Valley Test is used on the best ($\boldsymbol{x}_0$) and second best ($\boldsymbol{x}_1$) solutions to see if they belong to the same niche and are to be clustered together. If the HVT returns false for the solutions $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$, it is concluded that $\boldsymbol{x}_1$ is in another niche, as there are no other better solutions than $\boldsymbol{x}_0$ when compared to $\boldsymbol{x}_1$.

In general, the HVT is used to test at most $d = \ell + 1$ nearest better neighbours $\boldsymbol{x}_j$ for a solution $\boldsymbol{x}_i$. These d better solutions $\boldsymbol{x}_j$ are tested in ascending order of their Euclidean distance to $\boldsymbol{x}_i$, i.e., closest first. If any of the d checks with $\boldsymbol{x}_i$ return true, it concludes that $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are in the

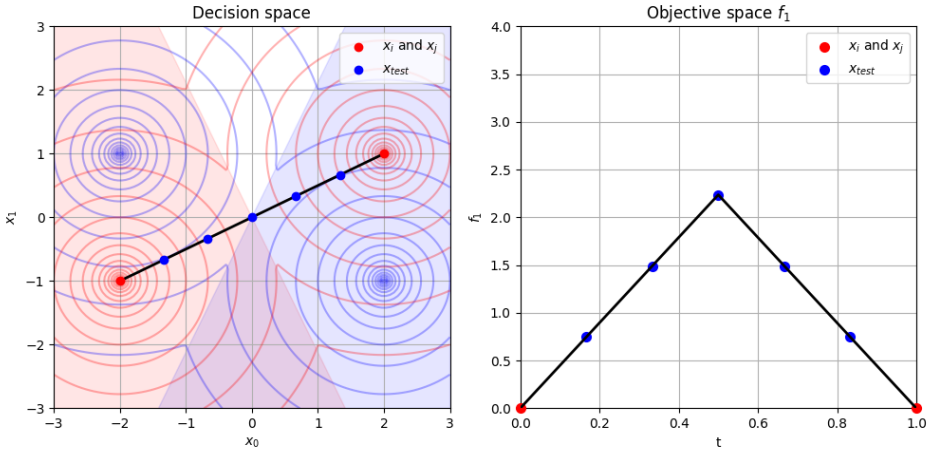


Figure 2.4: Example of the `HillValleyTest` [48] on the MinDist problem

same niche: $\mathcal{C}_k = \{\boldsymbol{x}_i, \boldsymbol{x}_j\}$. If all of the d tests return false, it concludes that the solution $\boldsymbol{x}_i$ is in a separate niche and should form a new cluster: $\mathcal{C}_l = \{\boldsymbol{x}_i\}$.

Algorithm 4: $[\mathbb{K}] =$ `HillValleyClustering`$(\mathcal{P}, f)$ [48]

Input: Population $\mathcal{P}$, Fitness functions f
Output : Set of clusters $\mathbb{K}$

$\Delta \leftarrow \left(\frac{V_X}{|\mathcal{P}|}\right)^{\frac{1}{\ell}}$;
Sort $x \in \mathcal{P}$ on fitness values for f, fittest first ;
$\mathcal{C} \leftarrow \{x_o\}$;
$\mathbb{K} \leftarrow \{\mathcal{C}\}$;
for $i \in \{1, 2, ..., |\mathcal{P}|\}$ **do**
    `// Compute all Euclidean distances to better solutions` x_p
    for $j \in \{1, 2, ..., i-1\}$ **do**
        $\delta_j \leftarrow ||x_i - x_j||$;
    end
    for $j \in \{0, ..., min(i-1, \ell+1)\}$ **do**
        $k \leftarrow$ index of j-th nearest better neighbour according to $\{\delta_p\}_{p=0}^{i-1}$;
        $N_t \leftarrow 1 + \lfloor \frac{\delta_k}{\Delta} \rfloor$;
        if *HillValleyTest*(x_k, x_i, N_t, f) **then**
            `// Add` x_i `to the cluster of` x_k
            $\mathcal{C}_{(x_k)} \leftarrow \mathcal{C}_{(x_k)} \cup \{x_i\}$ break ;
        end
    end
    if x_i *was not added to any cluster* **then**
        `// Initialize new cluster`
        $\mathcal{C} \leftarrow \{x_i\}$;
        $\mathbb{K} \leftarrow \mathbb{K}_l \cup \{\mathcal{C}\}$;
    end
end
return $\mathbb{K}$

2.6.2 Restart Scheme with Elitist Archive

Many evolutionary algorithms implement a form of a restart scheme whereby the population size can be increased over time. The interleaved multistart scheme [13], the restart-Covariance Matrix Adaptation Evolution Strategy with Increased Population (IPOP-CMA-ES) algorithm [3] and the parameter-free genetic algorithm [30] are examples of possible restart schemes. In the case of the HillVallEA method [48], an elitist archive is combined with a specific restart scheme. HillVallEA doubles the population size after each restart, similar to the scheme from IPOP-CMA-ES [3]. Together with the use the Hill-Valley Test, HillVallEA can check if a solution resides in a different niche, which is also used for the elitist archive. Therefore, HillVallEA is capable of maintaining the elites for each of the modes, despite it having been designed for single objective problems.

To prevent HillVallEA from revisiting already searched modes, it makes use of the elites that have already been found for each mode and thus stored in the elitist archive. This technique is inspired by the repelling subpopulations (RS-CMSA) algorithm [1] that defines taboo regions close to elites. HillVallEA uses a few steps to discard certain regions of the search space, for which an elite was already found in one of the earlier populations. First, it starts by adding the elites from previous runs to the population of the current restart. Second, all solutions are clustered using HVC. Lastly, it discards all clusters that have one of the elites as their best solution. The result of discarding these regions of the search space is that more attention is given to parts of the search space that are yet undiscovered in previous restarts.

In order to keep track of the different optima in each niche, HillVallEA employs a new test in order to add elites to the elitist archive $\mathcal{E}$. Since HillVallEA has been designed for single objective problems, it uses ϵ-tolerance to keep track of the best solutions found so far. If a new solution

x_i is found whereby $\exists \boldsymbol{e} \in \mathcal{E} : f(x_i) < f(\boldsymbol{e}) - \epsilon$, the elitist archive is emptied as a significantly better solution has been found. Furthermore, it employs the Hill-Valley Test in order to check if elites are from distinct niches in order to keep track of an elite for each niche. It performs the Hill-Valley Test for a new solution for $N_t = 5$ test solutions against an elite, similar to what is used in RS-CMSA [1]. Algorithm 5 shows how the test for a new elite is performed.

Algorithm 5: $[\mathcal{E}]$ = `AddEliteToArchive`($\boldsymbol{x}$, $\mathcal{E}$) [48]

Input: Candidate solution $\boldsymbol{x}$ and elitist archive $\mathcal{E}$
Output : Elitist archive $\mathcal{E} = [\mathcal{E}_0, \mathcal{E}_1, ...]$

$\boldsymbol{e} \leftarrow$ worst elite from $\mathcal{E}$;
if $f(\boldsymbol{x}_i) + TOL < f(\boldsymbol{e})$ **then**
  $\mathcal{E} \leftarrow \{\}$;
end
$\boldsymbol{e} \leftarrow$ elite closest to $\boldsymbol{x}$;
if *HillValleyTest*($\boldsymbol{x}$, $\boldsymbol{e}$, 5, f) **then**
  if $\boldsymbol{x} \succ \boldsymbol{e}$ **then**
    Remove $\boldsymbol{e}$ from $\mathcal{E}$;
    $\mathcal{E} \leftarrow \mathcal{E} \cup \{\boldsymbol{e}\}$;
  end
else
  $\mathcal{E} \leftarrow \mathcal{E} \cup \{\boldsymbol{e}\}$;
end
return $\mathcal{E}$;

2.6.3 HillValley Evolutionary Algorithm Overview

The pseudocode of Algorithm 6 shows how HillVallEA works. It initializes a new population as long as there is an optimization budget left. Starting with a small population it runs a core search algorithm $\mathcal{A}$ until termination for each of the clusters as produced by HVC. Then the best solution of each cluster is checked against the elitist archive, whereby it tries to store all global optima with fitness values within a certain tolerance of each other. If in the current run no new solutions were added to the elitist archive $\mathcal{E}$, it restarts the algorithm with a larger population and larger cluster sizes with the hope of finding new global optima.

By using a larger population in each restart, smaller niches can be found during initialization. The use of the restart scheme, as explained in the previous section, prevents the algorithm from optimizing modes for which an elite has already been found. Therefore, more attention can be given to the smaller niches.

Algorithm 6: $[\mathcal{E}]$ = `HillVallEA`($\mathcal{A}$) [48]

Input: Core search algorithm $\mathcal{A}$
Output : Elitist archive $\mathcal{E} = [\mathcal{E}_0, \mathcal{E}_1, ...]$

$n \leftarrow 16\ell$; `// initial population size`
$\mathcal{E} \leftarrow \{\}$; `// elitist archive`
while *budget remaining* **do**
  $\mathcal{P} \leftarrow$ `UniformSampling`(n) ;
  $\mathcal{P} \leftarrow \mathcal{P} \cup \mathcal{E}$;
  $\mathcal{S} \leftarrow$ Selection$(\mathcal{P})$;
  $\mathbb{K} =$ `HillValleyClustering`$(\mathcal{S}, f)$;
  for $\mathcal{C}_i \in \mathbb{K}$ **do**
    if *best solution in* $\mathcal{C}_i$ *is an elite* **then**
      continue ; `// skip already optimized niches`
    end
    $\boldsymbol{x}_i \leftarrow \mathcal{A}\left(\mu(\mathcal{C}_i), \Sigma(\mathcal{C}_i)\right)$;
    `AddEliteToArchive`$(\boldsymbol{x}_i, \mathcal{E})$;
  end
  if *no new solution was added to* $\mathcal{E}$ **then**
    $n \leftarrow 2 * n$;
  end
end

2.6.4 Multi-Objective Hill-Valley Clustering

MO-HillVallEA [47] is a more elaborate version as compared to the HillVallEA algorithm. It extends the HVC method in the form of Multi-Objective HVC (MO-HVC) by using the same concept of the Hill-Valley Test, but now clustering the entire population for each of the m objective functions separately. This results in m separate clustering sets of the population. To reduce all of these m sets these to a single clustering set, the intersection of all of them is taken, which entails that only those solutions that are in the same cluster for each objective are clustered together for the final clustering set. The pseudocode of Algorithm 7 gives an example of how MO-HVC works.

Algorithm 7: $[\mathbb{C}]$ = `MultiObjectiveHillValleyClustering`$(\mathcal{P}, f)$ [47]

Input: Population $\mathcal{P}$, Fitness functions $\{f_0, f_1, ..., f_m\}$
Output : Set of clusters $\mathbb{C}$

for $l \in \{0, 1, ..., m\}$ **do**
  $\mathbb{K}_l \leftarrow$ `HillValleyClustering`$(\mathcal{P}, f_l)$
end
`// Take intersection of all clusters in` $\mathbb{K}_l$
$\mathbb{C} \cap \{\mathbb{K}_l\}_{l=0}^{m-1}$
return $\mathbb{C}$

2.6.5 Cluster Registration

MO-HillVallEA is, like its predecessor HillVallEA, a two-staged approach whereby it first clusters the population and then runs a core search algorithm for each of the resulting clusters. The core search algorithm in the case of MO-HillVallEA is MAMaLGaM. As this algorithm uses certain parameters for each of the distinct clusters it needs to transfer these parameters between the clusters of each generation. Therefore, the notion of cluster registration [8] is used to transfer these parameters between each of the generations.

In cluster registration, the clusters are registered to their closest relative from the previous generation. A closest relative to a cluster is defined in terms of smallest Euclidean distance on their cluster means in decision space. Because no one-on-one relationship can be expected, as the number of clusters produced by MO-HVC can differ each generation, certain cluster parameters from the previous generation can be shared among multiple clusters in the next generation. Transferring these parameters allows MAMaLGaM to effectively search within each mode as it does not

lose valuable information. An example of the parameters that are transferred are the distribution multipliers as mentioned in Section 2.5.4.

2.7 Bézier Evolutionary Algorithm

A key feature of the algorithms proposed in this thesis is that of Bézier parameterizations as approximation sets for bi-objective optimization [72]. The idea of Bézier curves is not a new one in computer science, as it has widely been used in animation and computer graphics in order to represent smooth trajectories and curves [26]. Such a parameterization allows the algorithms to model an approximation set as a smooth curve through decision space. In order to optimize the approximation sets parameterized by Bézier curves, the algorithm uses the Uncrowded HyperVolume indicator (UHV) [66], which is an adaptation of the HyperVolume (HV) indicator [78]. This adaptation allowed the only known Pareto-compliant indicator [25, 79] to take dominated solutions into account and push optimization towards regions where these solutions were no longer dominated. Furthermore, the use of the HV indicator allows EAs to converge closer to the PS as compared to the Pareto dominance-based algorithms [6].

2.7.1 Uncrowded Hypervolume Indicator

The use of Bézier parameterizations means that all of the solutions in a population $\mathcal{P}$ are entire approximation sets. This entails that these solutions can no longer be scored using the regular objective functions, as these are only defined for single solutions, not solution sets. Therefore, another concept of ranking these solution sets is required in the form of the hypervolume indicator [78].

By employing the hypervolume measure, an EA is enabled in its search for optimal Pareto fronts, as an approximation set with an optimal hypervolume has been proven to be a subset of the Pareto set [25]. However, the measure does seem to have a certain bias towards convex inner portions of the objective space [76]. More research has also been performed as to which Pareto sets are preferred; how the solutions in such a set are spread over the Pareto front; and what a good choice of the reference point would be [4].

The first usages of the original hypervolume indicator can be traced back quite far, when it was only used to partially guide the search. For example, using the HV indicator as selection criteria in SMS-EMOA [7, 23], or as a secondary domination-based selection criteria in CMA-ES [33]. It can also be used in a hybrid EA approach, whereby multiple core search algorithms are run, similar to (MO-)HillVallEA, after which the HV measure can be used to select the best outcome of the results of these core search algorithms [53].

The original hypervolume measure did seem to suffer from a drawback, since dominated solutions would have no selection pressure towards the non-dominated region in the search space [24]. Recently, an adaptation to the hypervolume indicator has been created in the form of the Uncrowded HyperVolume (UHV) measure [66]. The UHV has been adapted from the HV measure by adding a penalty term in the form of uncrowded distance for dominated solutions. This penalty term deducts the Euclidean distances as measured from the objective values of a dominated solution to the approximation front from the hypervolume value. The added penalty results in increased fitness if dominated solutions move towards the undominated regions, e.g. towards the approximation front, as the Euclidean distance and subsequent penalty term diminishes. The UHV measure has also been implemented in GOMEA in the form of a new algorithm called the Uncrowded HyperVolume Gene-pool Optimal Mixing Evolutionary Algorithm (UHV-GOMEA) [69].

In order to mathematically explain the uncrowded hypervolume indicator, let $f : \mathcal{X} \to \mathbb{R}^m$ be a fitness function that is to be minimized. $\mathcal{X} \subseteq \mathbb{R}^\ell$ is the, possibly constraint, ℓ-dimensional search space, i.e., decision space. Then, let $\mathcal{S} \subseteq \mathcal{X}$ be a solution set with possibly dominated solutions according to the Pareto dominance relation of Definition 3.

Definition 7. *The function of the approximation set $\mathcal{A}$ given by the solution set $\mathcal{S}$, $\mathcal{A} : \wp(\mathcal{X}) \to \wp(\mathcal{X})$, where $\wp(\mathcal{X})$ is the powerset of $\mathcal{X}$, is defined as:*

$$\mathcal{A}(\mathcal{S}) = \{\boldsymbol{x} \in \mathcal{S} | \nexists \boldsymbol{y} \in \mathcal{S} : \boldsymbol{y} \succ \boldsymbol{x}\}$$

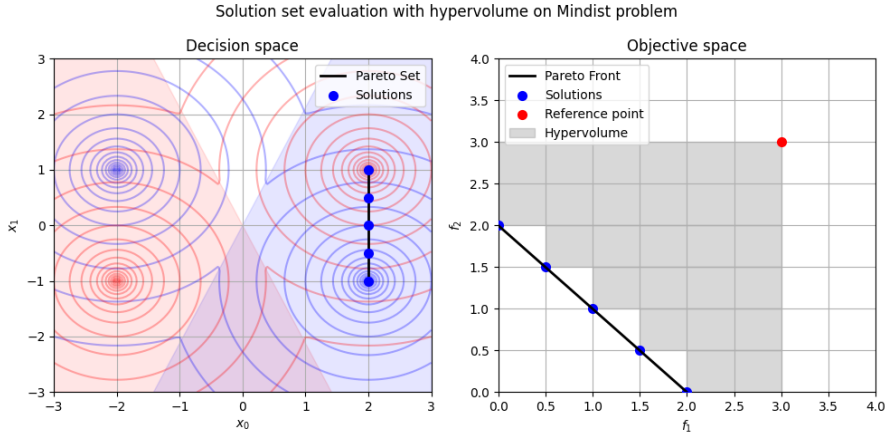


Figure 2.5: Example of hypervolume calculation for a solution set that lies on top of the Pareto set/front.

The hypervolume measure $HV : \wp(\mathcal{X}) \to \mathbb{R}$ is a function that measures the area, or volume in higher dimensional problems, dominated by all solutions in the approximation set $\mathcal{A}$. As this area or volume can not be unbounded, which would make it incalculable, it is bounded by a reference point $\boldsymbol{r} \in \mathbb{R}^m$. The hypervolume is then the area between $\boldsymbol{r}$ and $\mathcal{A}$. Figure 2.5 shows a depiction of the hypervolume calculation for a solution set $\mathcal{S}$, which is a subset of one of the PSs, on the MinDist problem using a reference point $\boldsymbol{r} = [3, 3]$.

From this definition, it seems that the choice of reference point $\boldsymbol{r}$ can be rather important. This has also been shown to be true, since the choice of reference point can influence whether or not the approximation set can cover the endpoints of the Pareto front [4]. Even in the case where the reference point would be set to infinity, by virtue of which the hypervolume can no longer be used for selection as all approximation sets dominate an infinite area, the endpoints can be missed.

In order to take the dominated solutions into account and rank entire solutions sets $\mathcal{S}$ instead of just the approximation set $\mathcal{A}$, the uncrowded hypervolume [66] can be used. This measure adapts the hypervolume measure in order to add a penalty term to dominated solutions. By doing so, the uncrowded hypervolume adds a form of selection pressure in order to push all dominated solutions to the undominated region, which is useful when used as an optimization function.

The uncrowded hypervolume uses the hypervolume measure for the undominated solutions in $\mathcal{S}$, i.e., $\mathcal{A}(\mathcal{S})$, and deducts the uncrowded distances $ud_f(\boldsymbol{x}_i, \mathcal{A})$ for each of the dominated solutions in $\mathcal{S}$. The uncrowded distance $ud_f(\boldsymbol{x}_i, \mathcal{A})$ calculates the shortest Euclidean distance in objective space between solution $\boldsymbol{x}_i$ and the approximation boundary defined by $\delta f(\mathcal{A})$ and $\boldsymbol{r}$. If a solution $\boldsymbol{x}_i$ is not dominated by either $\mathcal{A}$ or $\boldsymbol{r}$, then $ud_f(\boldsymbol{x}_i, \mathcal{A}) = 0$.

Definition 8. *The combination of the hypervolume measure and the uncrowded distance results in the Uncrowded HyperVolume (UHV) measure [66] for solution sets $\mathcal{S}$:*

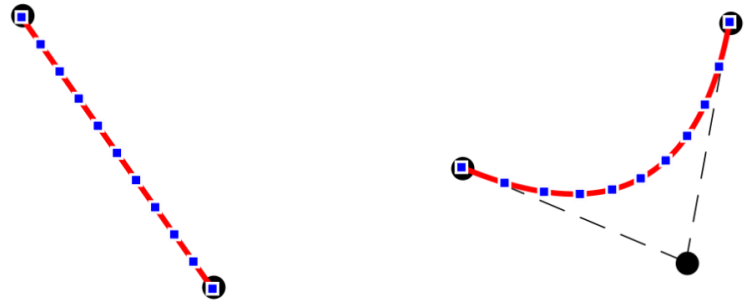
$$UHV_f(\mathcal{S}) = HV_f\left(\mathcal{A}(\mathcal{S})\right) - \frac{1}{|\mathcal{S}|} \sum_{\boldsymbol{x} \in \mathcal{S}} ud_f\left(\boldsymbol{x}, \mathcal{A}(\mathcal{S})\right)^m$$

Let $\phi \in \mathbb{R}^n$ be a parameterization consisting of n decision variables, then $\mathcal{S}(\phi) = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ...\}$ is an operator that transforms the parameterization ϕ into a solution set.

Definition 9. *With the use of the UHV measure and a parameterization as $\mathcal{S}(\phi)$, an UHV-based optimization problem can now be defined as:*

$$\textit{maximize} \quad UHV_{f,\mathcal{S}(\phi)} = HV_f\left(\mathcal{A}(\mathcal{S}(\phi))\right) - \frac{1}{|\mathcal{S}(\phi)|} \sum_{\boldsymbol{x} \in \mathcal{S}(\phi)} ud_f\left(\boldsymbol{x}, \mathcal{A}(\mathcal{S}(\phi))\right)^m$$

$$\textit{with} \quad f : \mathcal{X} \subseteq \mathbb{R}^\ell \to \mathbb{R}^m, \mathcal{S} : \mathbb{R}^n \to \wp(\mathcal{X}), \phi \in \mathbb{R}^n$$



$$\mathbf{B}(t;\mathcal{C}_2) = (1-t)\mathbf{c}_1 + t\mathbf{c}_2 \quad \mathbf{B}(t;\mathcal{C}_3) = (1-t)^2\mathbf{c}_1 + 2(1-t)t\mathbf{c}_2 + t^2\mathbf{c}_3$$

Figure 2.6: Example of two Bézier curves in red for $q \in \{2,3\}$ control points C_q in black, respectively left and right. In blue there are the $p = 11$ test points denoting the set of solutions $S_{p,q}(C_q)$, which are spread evenly in the domain of t [72].

Using the UHV-based optimization problem above together with the RV-GOMEA algorithm gave rise to the algorithm called Uncrowded HyperVolume Genepool Optimal Mixing Evolutionary Algorithm (UHV-GOMEA) [69].

2.7.2 Bézier Parameterization

As mentioned previously, a Bézier curve $\mathbf{B}(t;C_q)$ can be used in order to define a solution set that can be evaluated with the UHV measure from Definition 8. This gives rise to an optimization problem whereby Bézier parameterizations can be optimized, similar to the UHV optimization problem from Definition 9. In order to define a solution set $\mathcal{S}$ as given by the Bézier curve $\mathbf{B}(t;C_q)$, the definition of a Bézier curve $\mathbf{B}(t;C_q)$ is required.

Definition 10. *An ℓ-dimensional Bézier curve $\mathbf{B}(t;C_q)$, using $q \geq 2$ control points $\boldsymbol{c}_i$ in an ordered set $C_q = \{\boldsymbol{c}_1, ..., \boldsymbol{c}_q\}$, is defined as:*

$$\mathbf{B}(t;C_q) = \sum_{i=1}^{q} \binom{q-1}{i-1} (1-t)^{q-i-1} t^{i-1} \boldsymbol{c}_i \text{ for } 0 \leq t \leq 1$$

$$\text{with } \boldsymbol{c}_i \in \mathbb{R}^\ell, \ell \text{ the problem dimensionality}$$

The endpoints of the Bézier curve are always defined by its first and last control points, $\boldsymbol{c}_1$ and $\boldsymbol{c}_q$. The intermediary control points are usually not located on the curve, but are used to determine the trajectory of the curve. With the parameterization of a Bézier curve, an evenly spread set of p points can be selected to form a solution set of fixed size, $\mathcal{S}_p = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_p\}$ with $\boldsymbol{x}_i \in \mathbb{R}^\ell$.

Definition 11. *The set of solutions $S_{p,q}(C_q) = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_p\}$ is defined as:*

$$S_{p,q}(C_q) = \left\{ \mathbf{B}\left(\frac{0}{p-1}; C_q\right), \mathbf{B}\left(\frac{1}{p-1}; C_q\right), ..., \mathbf{B}\left(\frac{p-1}{p-1}; C_q\right) \right\}$$

Figure 2.6 gives an example of how a Bézier curve can be used to denote the solution set $S_{p,q}(C_q)$. In red it shows the Bézier curve if it was drawn infinitely precise, with the control points C_q denoted in black. The $p = 11$ blue points are the solutions denoted by $S_{p,q}(C_q)$ once they have been evenly spread along the curve in the domain of $t \in [0,1]$.

In order to have evolutionary algorithms directly optimize solution sets $S_{p,q}(C_q)$ defined by Bézier curves, a parameterization is required to form the decision variables. A solution set $S_{p,q}(C_q)$ is parameterized by concatenating the decision variables in the set C_q and encoding them as one solution [7, 72]. The result is solution being of the form $[\boldsymbol{c}_1, ..., \boldsymbol{c}_q] \in \mathbb{R}^{q*\ell}$. Finally, the order of the control points is reversed, at any point during optimization, to $[\boldsymbol{c}_q, ..., \boldsymbol{c}_1] \in \mathbb{R}^{q*\ell}$ in case the evaluation of the control points indicates that $f_1(\boldsymbol{c}_q) < f_1(\boldsymbol{c}_1)$.

2.7.3 Evaluation of Bézier Solution Sets

In order to evaluate a solution set $S_{p,q}(C_q)$, a number of functions were previously introduced [72] to further promote the correct optimization towards the Pareto front. These function ensure that a certain navigational order is preserved and enforced and the solution set $S_{p,q}(C_q)$ does not fold over on itself in objective space as in Figure 2.7. The definitions as given below only apply in the bi-objective problem setting, for multi-objective problems where $m > 2$ the definitions below no longer hold. Figure 2.7 illustrates these functions to give the reader a more graphical indication.

The navigational Bézier (nb) order o_{nb} ensures that the order of solutions, as intrinsically defined with Definition 11, will take dominated solutions into account. A new function $A^{nb}(S_{p,q})$ has been introduced that calculates o_{nb}. The order is defined as starting from the best solution for objective f_1 traversing all of the solutions in the set $S_{p,q}(C_q)$ up until the best solution in f_2 has been reached. All solutions that are dominated by others, and that lie on the curve, are omitted from the subset that defines the navigational order. Using only the solution indices as specified in o_{nb}, an approximation set $\mathcal{A}_{p,q,o_{nb}}$ is a subset of $S_{p,q}(C_q)$. The pseudocode of Algorithm 8 shows how the $\mathcal{A}_{p,q,o_{nb}}$ can be obtained from $S_{p,q}(C_q)$.

Algorithm 8: $[\mathcal{A}_{p,q,o_{nb}}, o_{nb}] =$ `NavigationalOrder`$(S_{p,q}(C_q), o^{bez})$ [72]

Input: Bézier solution set $S_{p,q}(C_q)$, with intrinsic order o^{bez}
Output : Approximation set $\mathcal{A}_{p,q,o_{nb}}$

$\eta \leftarrow \arg\min_{i \in \{1,\dots,p\}} f_1\left(\boldsymbol{x}_{o_i^{bez}}\right)$
$o_{nb} \leftarrow o_\eta^{bez}$
$\mathcal{A}_{p,q,o_{nb}} \leftarrow \{\boldsymbol{x}_{o_\eta^{bez}}\}$
for $j \in \{\eta, \dots, p\}$ **do**
  if $\boldsymbol{x}_{o_j^{bez}} \in \mathcal{A}(\mathcal{S}_{p,q})$ *and* $f_2\left(\boldsymbol{x}_{o_j^{bez}}\right) < f_2\left(\boldsymbol{x}_{o_{end}^{bez}}\right)$ **then**
    $o_{nb} \leftarrow [o_{nb}; o_j^{bez}]$
    $\mathcal{A}_{p,q,o_{nb}} \leftarrow \mathcal{A}_{p,q,o_{nb}} \cup \{\boldsymbol{x}_{o_j^{bez}}\}$
  end
end
return $\mathcal{A}_{p,q,o_{nb}}, o_{nb}$

Besides the enforcement of a smooth navigational order, the authors of BezEA also introduced a new constraint function $C\left(\mathcal{S}_{p,q}, o^{bez}\right)$. It is employed in order to not only push all of the dominated solutions towards the undominated region of the search space, but also spread the solutions evenly over the front and prevent it from intersecting itself. The constraint function employs the uncrowded distance metric $ud_f(\boldsymbol{x}_i, \mathcal{A})$ [66] in order to let the dominated solutions become no longer dominated. Furthermore, all dominated solutions are pulled towards their neighbouring solutions on the Bézier curve in order to further increase pressure towards the unfolding of Bézier curves in objective space during optimization. This is achieved by taking the Euclidean distance in objective space between the dominated solution and its neighbours on the Bézier curve as an additional constraint value. All dominated solutions from $S_{p,q}(C_q)$ now have their uncrowded distance to the approximation front and Euclidean distance to their neighbours in objective space summed up as a constraint for the total solution set. Combining the use of the constraint function together with constraint domination [19], this constraint pushes all solutions along the curve towards the undominated region and unfolds the Bézier curve. The pseudocode of Algorithm 9 shows the constraint function as described.

Figure 2.7 shows how all of the above is combined in order to evaluate the Bézier solution sets $S_{p,q}(C_q)$. It shows how the hypervolume value is calculated based on those solutions in the approximation set $\mathcal{A}_{p,q,o_{nb}}$. Furthermore, it shows the solutions that violate the navigational order, due to which these are omitted from the approximation set. The uncrowded distance values $ud_f(\boldsymbol{x}_i, \mathcal{A})$ are also shown graphically in order to give an insight as to how these are calculated.

Algorithm 9: $[\mathcal{C}]$ = `BézierConstraint`$(S_{p,q}(C_q),\, o^{bez})$ [72]

Input: Bézier solution set $S_{p,q}(C_q)$, with intrinsic order o^{bez}
Output : Constraint value $\mathcal{C} \geq 0$

$\mathcal{A}_{p,q,o_{nb}},\, o_{nb} \leftarrow$ `NavigationalOrder`$(S_{p,q}(C_q),\, o^{bez})$ `// See Algorithm 8`
$\mathcal{C} \leftarrow \frac{1}{|S_{p,q}(C_q)|} \sum_{\boldsymbol{x} \in S_{p,q}(C_q)} ud_f(\boldsymbol{x}, \mathcal{A}_{p,q,o_{nb}})$
for $j \in \{1, ..., |S_{p,q}(C_q)| - 1\}$ **do**
  if $\boldsymbol{x}_{o_j^{bez}} \notin o_{nb}$ *or* $\boldsymbol{x}_{o_{j+1}^{bez}} \notin o_{nb}$ **then**
    $\mathcal{C} \leftarrow \mathcal{C} + ||f(\boldsymbol{x}_{o_j^{bez}}) - f(\boldsymbol{x}_{o_{j+1}^{bez}})||$ `// Euclidean distance in` $\mathbb{R}^m$
  end
end
return $\mathcal{C}$

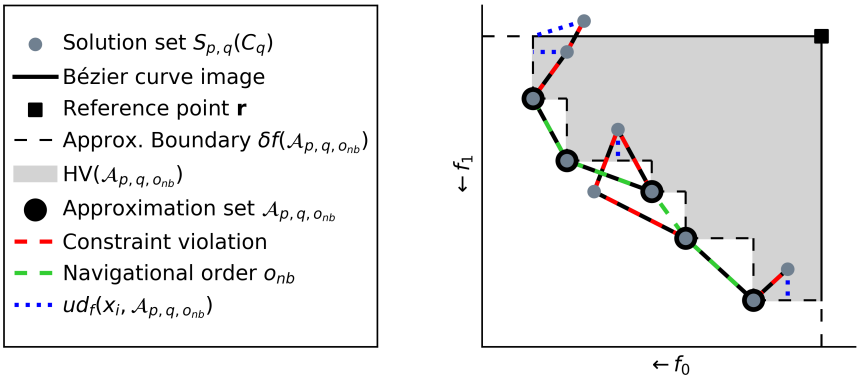


Figure 2.7: Example of the evaluation of Bézier solution sets [72].

2.7.4 Bézier Curve Optimization Problem

The previous sections explained all of the required steps in order to now formulate an optimization problem for Bézier curve parameterizations. The problem formulation combines the navigational order function of Algorithm 8, the constraint function of Algorithm 9 and the hypervolume definition. It is an adaptation of the UHV optimization problem of Definition 9, although it uses all of the concepts behind it through the mentioned algorithms.

Definition 12. *A Bézier curve optimization problem is defined by:*

$$\begin{aligned} \textit{maximize} \quad & HV_{f,\mathcal{S}_{p,q}(\phi)} = HV_f\left(\mathcal{A}_{p,q,o_{nb}}\right)) \\ \textit{with} \quad & \texttt{BézierConstraint}(S_{p,q}(\phi), o^{nb}) = 0 \\ \textit{where} \quad & f : \mathcal{X} \subseteq \mathbb{R}^\ell \rightarrow \mathbb{R}^m, \mathcal{S}_{p,q} : \mathbb{R}^{q*\ell} \rightarrow \wp(\mathcal{X}), \phi \in \mathbb{R}^{q*\ell}, \\ & [\mathcal{A}_{p,q,o_{nb}}, o^{nb}] = \texttt{NavigationalOrder}(S_{p,q}(\phi), o^{bez}) \end{aligned}$$

2.8 High-Dose-Rate Brachytherapy

High-Dose-Rate (HDR) brachytherapy is a form of internal radiation treatment. Through the use of catheters implanted within the patients body, radiation can be administered closely to the cancerous tumour. The radiation is produced by a radioactive source that is made up of, for instance, Iridium-192 atoms. The source is held at specific locations, called dwell positions, within the patient for a predefined amount of time, called the dwell time. The source gives off a high dosage of radiation due to which the treatment time can be kept relatively short, i.e., under half an hour.

The purpose of HDR brachytherapy is to administer a prescribed radiation dose to the target volume, which is measured in Gray (Gy). Simultaneously, for the healthy organs surrounding the

tumour, called the Organs At Risk (OARs), the administered radiation dose is to be kept under a prescribed maximum. In order to achieve this, a treatment plan is to be created. The plan describes the time during which the radioactive source is kept at each of the predefined locations in the catheters, thereby administering a dose from within the body. As there is a trade-off between irradiating the tumour enough and preserving the OARs as best as possible, a bi-objective problem formulation has been introduced, which has been optimized by EAs [43]. Thus, the optimization concerns finding the best possible dwell times for each dwell position in order to administer a radioactive dose that spares the healthy OARs as much as possible, but also irradiates the tumour enough.

2.8.1 Evaluation of Treatment Plan

HDR brachytherapy treatment plans describe a specific radiation dose distribution which is to be administered to the tissues around the implanted catheters. In order to determine if the OARs remain under a certain radiation dose and the target volume is irradiated enough, a specific protocol has been determined by various hospitals. Two of such protocols are those of the Amsterdam UMC, location AMC, where a planned radiation dose of 13 Gy, in the old protocol used until 2020, or 15 Gy, in the new protocol used since 2020, is used for HDR brachytherapy treatment plans for prostate cancer. However, for some patients, it is not possible to satisfy all criteria in these protocols, which can for example be attributed to the sub-optimal placement of catheters.

Protocols. Clinical protocols for HDR brachytherapy consist of so-called Dose-Volume (DV) indices, each of which has its own corresponding criterion. Often there are two types of criterion described in a clinical protocol:

- V_d^o criteria define the desirable cumulative volume of an organ o that receives at least a radiation dose level d. The dose d is relative to the planning aim dose, i.e., 13 or 15 Gy.
- D_v^0 criteria define the maximum radiation dose level for the highest irradiated cumulative volume v of an organ o.

As an example for the criteria, $\mathbf{V}_{100}^{prostate} > 95\%$ implies that the percentage of prostate volume with a planned dose of at least 100% of the planning aim dose should be at least 95% [32]. Not all criteria concern a minimum volume that is to receive a certain prescribed radiation dose. For example, the $\mathbf{V}_{200}^{prostate} < 20\%$ indicates that the cumulative prostate volume that is covered by more than 200% of the prescribed radiation dose cannot be larger than 20% of the entire prostate volume. There are also criteria that prevent excessive radiation exposure, e.g., the D_v^0 criteria. For example, the $D_{1cm^3}^{bladder} < 86\%$ criteria defines that the most-irradiated 1 cubic centimetre receives at most 86% of the prescribed radiation dose. Table 2.1 shows all of the DV indices and corresponding criteria used at the Amsterdam UMC up until 2020, whereas Table 2.2 shows the new protocol in effect since 2020.

Dose Calculation Points. As it is key for the treatment plans to adhere to the clinical protocol, an evaluation method is required to compare treatment plans with the prescribed aspiration values of Tables 2.1 or 2.2. This is done through the use of the so-called dosimetry engine, which has been defined based on recommendations by the American Association of Physicists in the Medicine (AAPM) Task Group 43 formula (TG-43) [52, 60]. Currently, the body of a patient is modelled as a uniform body of water with uniform mass. Furthermore, the dosimetry engine incorporates the strength and shape of the employed radiation source.

In order to determine the administered radiation dose to all of the OARs and target volume, the dosimetry engine uses a Monte Carlo method. A large number of so-called dose calculation

Table 2.1: Old HDR brachytherapy planning dose-volume criteria at Amsterdam UMC with a planning-aim dose of 13Gy, used until 2020.

Prostate	Bladder	Rectum	Urethra	Vesicles
$V_{100\%} > 95\%$	$D_{1cm^3} < 86\%$	$D_{1cm^3} < 78\%$	$D_{0.1cm^3} < 110\%$	$V_{80\%} > 95\%$
$V_{150\%} < 50\%$	$D_{2cm^3} < 74\%$	$D_{2cm^3} < 74\%$		
$V_{200\%} < 20\%$				

Table 2.2: New HDR brachytherapy planning dose-volume criteria at Amsterdam UMC with a planning-aim dose of 15Gy, since 2020.

Prostate	Bladder	Rectum	Urethra	Vesicles
$V_{100\%} > 95\%$	$D_{1cm^3} < 86.66\%$	$D_{1cm^3} < 73.33\%$	$D_{0.1cm^3} < 120\%$	$V_{73\%} > 95\%$
$V_{150\%} < 40\%$	$D_{2cm^3} < 80\%$	$D_{2cm^3} < 63.33\%$	$D_{30\%} < 110\%$	
$V_{200\%} < 15\%$				
$D_{90\%} > 100\%$				

points are used to approximate the actual radiation dose for the patient. For each of these dose calculation points, the radiation dose is the total sum of the planned dose contributions from the dwell positions. The dose contributions themselves are the product of the dwell time with the dose rate, i.e., the time a radioactive source is kept at a dwell position and the amount of radiation it emits, and the distance of the dwell position to the dose calculation point.

Let D denote the set of dose calculation points of size n_D and T the set of dwell positions of size n_T. A dose rate matrix $\mathbf{R}$ of size $n_D \times n_T$ is computed for a specific radiation source. Each element $\mathbf{R}_{ij}$ specifies the dose rate between a dose calculation point i and a dwell position j. A treatment plan can be represented as a vector of dwell times $\boldsymbol{t}$ of size n_T for each of the activated dwell positions. Then, the dose distribution represented by a vector $\boldsymbol{d}$ of length n_D for a treatment plan $\boldsymbol{t}$ can be computed through the multiplication of the dose rate matrix $\mathbf{R}$ and the dwell times $\boldsymbol{t}$:

$$\boldsymbol{d} = \mathbf{R}\boldsymbol{t}$$

Dose Volume Index Evaluation. The DV indices can be approximated using the calculated dose distribution $\boldsymbol{d}$. To compute these values, a set D_0 denotes the dose calculation points within an organ o, whereby $D_0 \subset D$. Depending on whether a V_d^o or D_d^o index is to be approximated, two distinct equations are used. For V_d^o indices an indicator function $\chi(d_i, d)$ is required, where d_i is the total radiation at dose calculation point i and d the target dose:

$$\chi(d_i, d) = \begin{cases} 1 & d_i \geq d \\ 0 & d_i < d \end{cases}$$

Then the DV index V_d^o is approximated as:

$$V_d^o = \frac{1}{n_D^0} \sum_{i \in D_0} \chi(d_i, d)$$

For D_v^0 indices all dose calculation points in the organ o, denoted by the vector $\boldsymbol{d}^o$, the vector of dose calculation points sorted in descending order $\boldsymbol{d}^{s,o}$ is used. Then, with v_0 denoting the volume associated with a dose calculation point within organ o, D_v^0 is computed as:

$$D_d^o = \boldsymbol{d}^{s,o}_{\left\lfloor \frac{v}{v^0} \right\rfloor}$$

Evaluation and Optimization Speed. In order to correctly approximate the dose distribution, a large number of dose calculation points is required. Using smaller amounts of dose calculation points speeds up the calculation, but then there can be a discrepancy between the approximated radiation dose and the actual radiation dose due to the Monte Carlo simulation. Using 20,000 dose calculation points in total was shown to be an effective amount in order to approximate the dose distribution well when optimizing with single core processor systems. At the end of the optimization run, 500,000 points are used in order to re-evaluate the entire set of optimized treatment plans for a more precise measurement [42].

More recent work introduced the use of a Graphics Processing Unit (GPU) to further speed up the dose calculations and allow for more accurate results through the use of more dose calculation points. 100,000 dose calculation points in total are used during optimization, whereas the re-evaluation with 500,000 dose calculation points at the end remained the same [11]. Using the GPU accelerated version, the compute time required to obtain high-quality treatment plans was reduced from one hour in clinical practice to just 3 minutes, even though more time is required for actual

selection of the preferred treatment plans. Moreover, the GPU version uses a problem formulation with normalized objectives, which will be explained in the sections below. This normalized problem formulation is different from the one that was found to produce treatment plans that are preferable to those produced through manual optimization with the aid of optimization software by clinicians in 53/54 cases [50].

2.8.2 Bi-Objective Problem Formulation

A bi-objective formulation has been created that represents the inherent trade-off between coverage of the target volumes and sparing of the surrounding healthy tissue [43]. If each of the DV indices of Tables 2.1 and 2.2 is a separate objective, it would be difficult to efficiently solve an optimization problem as it is challenging to find improvements when optimizing close to the Pareto set. Furthermore, using a bi-objective formulation has the added benefit that it can be easily visualized, in contrast to a 9 or 11-objective formulation for each of the DV indices. By visualizing the bi-objective formulation in the form of a graph showing the trade-off between both objectives, the interpretability of the results for human experts is drastically improved, which could aid them in decision making. The bi-objective problem formulation that is discussed in this section has been employed for the old protocol used until 2020. The next section, i.e., Section 2.8.3, will discuss the adapted bi-objective problem formulation for the new protocol.

Coverage and Sparing. The bi-objective formulation divides the DV indices of the old protocol in Table 2.1 into two groups depending on whether they have a maximization or minimization aim [44]. The $V_{200\%}^{prostate}$, $V_{150\%}^{prostate}$ can be omitted from these groups as it was found that they were never violated [50]. However, since the current practice in the clinics is to optimize with all DV indices, the $V_{200\%}^{prostate}$ and $V_{150\%}^{prostate}$ will be included in the problem formulation as given here. The DV indices are grouped as follows:

1. Target Coverage: $V_{100\%}^{prostate}$ and $V_{80\%}^{vesicles}$
2. Organ sparing: $V_{200\%}^{prostate}$, $V_{150\%}^{prostate}$, $D_{1cm^3}^{rectum}$, $D_{2cm^3}^{rectum}$, $D_{1cm^3}^{bladder}$, $D_{2cm^3}^{bladder}$ and $D_{0.1cm^3}^{urethra}$

A maximin approach is employed to combine the Target Coverage group into the so-called Least Coverage Index (LCI), whereas the Organ Sparing group is combined into the Least Sparing Index (LSI). Here, the LCI objective value corresponds to the worst performing DV index of the Target Coverage group and the LSI objective corresponds to the worst performing DV index of the Organ Sparing group. The LCI group contains only volume indices, whereas the LSI group contains both dose and volume indices.

LCI and LSI Evaluation. For both the volume indices V_d^o and dose indices D_d^o distance functions δ_v are employed to determine how close a DVI is to its respective aim, as described in the clinical protocol. These resulting values for the distance functions are synonymously called delta DVI. For the target coverage group, the volume indices V_d^o have a lower bound threshold $V_d^{o,min}$ that is used to determine how well a target o is covered by the required dose d. The distance function δ_c for a V_d^o index is then:

$$\delta_c(V_d^o) = V_d^o - V_d^{o,min}$$

For the sparing group another distance function δ_d is employed, which uses a $D_d^{o,max}$ or $V_d^{o,max}$ upper bound threshold for each DV index. The further D_d^o is below $D_d^{o,max}$, or V_d^o below $V_d^{o,max}$, the better the organ is spared. The distance function δ_{sd} for a D_d^o index, or δ_{sv} for a V_d^o index, is then:

$$\delta_{sd}(D_d^o) = D_d^{o,max} - D_d^o$$
$$\delta_{sv}(V_d^o) = V_d^{o,max} - V_d^o$$

Using the distance functions defined above, the LCI and LSI objectives can now be defined as follows:

$$\begin{aligned} LCI(R\boldsymbol{t}) =& \min\{\delta_c\left(V_{100\%}^{prostate}\right), \delta_c\left(V_{80\%}^{vesicles}\right)\} \\ LSI(R\boldsymbol{t}) =& \min\{\delta_{sv}\left(V_{200\%}^{prostate}\right), \delta_{sv}\left(V_{150\%}^{prostate}\right), \delta_{sd}\left(D_{1cm^3}^{rectum}\right), \delta_{sd}\left(D_{2cm^3}^{rectum}\right), \\ & \delta_{sd}\left(D_{1cm^3}^{bladder}\right), \delta_{sd}\left(D_{2cm^3}^{bladder}\right), \delta_{sd}\left(D_{0.1cm^3}^{urethra}\right)\} \end{aligned}$$

Here, the LCI value has the range $[-0.95, 0.05]$ and the LSI value has the range $(-\infty, 0.74]$. If the LCI has value 0.05, it indicates that the entire volume of each target volume is covered by the recommended dose level. Similarly, if the LSI has value 0.74, this implies that there is no radiation to the OARs. For the old protocol, the objectives of LCI and LSI together stated the bi-objective optimization problem. Combining these two, having both the LCI and LSI values above 0 indicates that, for a specific treatment plan, all clinical criteria are satisfied. However, sometimes plans that do not satisfy all DVIs can still be clinically acceptable, but that is for clinicians to decide.

2.8.3 Normalized Bi-Objective Problem Formulation

For the new protocol in use since 2020, another bi-objective optimization problem has been defined, which has been built upon the weighted LCI and LCI optimization problem [11]. It is not formulated as the trade-off between LCI and LSI, but between weighted and normalized coverage LCI_{wn} and sparing LSI_{wn}. The weighted and normalized versions take into account all underlying indices instead of only the most exceeded or least satisfied index. This is to ensure that if an index is not the least satisfied or most exceeded during optimization, but an improvement is still found for that index, the improvement translates to a better objective value. The old LCI and LSI objectives would only show an improvement for a treatment plan if the least satisfied or most exceeded index would improve, thus ignoring other possible improvements that could occur during optimization.

The DV indices of Table 2.2 are once again divided into two groups for coverage and sparing. Here, the $V_{200\%}^{prostate}$ and $V_{150\%}^{prostate}$ can be omitted with similar reasons to the previous protocol and the $D_{90\%}^{prostate}$ can be omitted as it is always satisfied when the $V_{100\%}^{prostate}$ is satisfied [11]. However, all DV indices are shown here since that is the current clinical practice. The DV indices are grouped as follows:

1. Target Coverage: $V_{100\%}^{prostate}$, $D_{90\%}^{prostate}$ and $V_{73\%}^{vesicles}$
2. Organ sparing: $V_{200\%}^{prostate}$, $V_{150\%}^{prostate}$, $D_{1cm^3}^{rectum}$, $D_{2cm^3}^{rectum}$, $D_{1cm^3}^{bladder}$, $D_{2cm^3}^{bladder}$, $D_{0.1cm^3}^{urethra}$ and $D_{30\%}^{urethra}$

The normalized deltas δ_{cdn} for D indices and δ_{cvn} V indices in the coverage group, and δ_{sn} for both D and V indices in the sparing group must be calculated in order to calculate the normalized objective values. These are defined as:

$$\delta_{cdn}(D_d^o) = \frac{D_d^{o,max} - D_d^o}{0.3 * D_d^{o,max}} \tag{2.5}$$

$$\delta_{cvn}(V_d^o) = \frac{V_d^{o,max} - V_d^o}{1.0 - V_d^{o,max}} \tag{2.6}$$

$$\delta_{sn}(V_d^o) = \frac{V_d^o - V_d^{o,max}}{V_d^{o,max}} \tag{2.7}$$

Similar to the previous LCI ans LSI objective, a normalized LCI and LSI are also defined. These are not used as objectives during optimization, but are used in order to show the final approximation fronts. The normalized LCI and LSI are defined as:

$$\begin{aligned} LCI_n(R\boldsymbol{t}) =& \min\{\delta_{cdn}\left(D_{90\%}^{prostate}\right), \delta_{cvn}\left(V_{100\%\%}^{prostate}\right), \delta_{cvn}\left(V_{73\%}^{vesicles}\right)\} \\ LSI_n(R\boldsymbol{t}) =& \min\{\delta_{sn}\left(V_{200\%}^{prostate}\right), \delta_{sn}\left(V_{150\%}^{prostate}\right), \delta_{sn}\left(D_{1cm^3}^{rectum}\right), \delta_{sn}\left(D_{2cm^3}^{rectum}\right), \\ & \delta_{sn}\left(D_{1cm^3}^{bladder}\right), \delta_{sn}\left(D_{2cm^3}^{bladder}\right), \delta_{sn}\left(D_{0.1cm^3}^{urethra}\right), \delta_{sn}\left(D_{30\%}^{urethra}\right)\} \end{aligned} \tag{2.8}$$

For the normalized and weighted LCI and LSI, a sorted list of the deltas Δ_c^s is created for the LCI and Δ_s^s for LSI after the normalization of the indices:

$$\begin{aligned} \Delta_c^s =& sort^{\text{DESC}}\left([\delta_{cdn}(D_{90\%}^{prostate}), \delta_{cvn}(V_{100\%}^{prostate}), \delta_{cvn}(V_{73\%}^{vesicles})]\right) \\ \Delta_s^s =& sort^{\text{DESC}}([\delta_{sn}\left(V_{200\%}^{prostate}\right), \delta_{sn}\left(V_{150\%}^{prostate}\right), \delta_{sn}\left(D_{1cm^3}^{rectum}\right), \delta_{sn}\left(D_{2cm^3}^{rectum}\right), \\ & \delta_{sn}\left(D_{1cm^3}^{bladder}\right), \delta_{sn}\left(D_{2cm^3}^{bladder}\right), \delta_{sn}\left(D_{0.1cm^3}^{urethra}\right), \delta_{sn}\left(D_{30\%}^{urethra}\right)]) \end{aligned}$$

With the sorted list of DV deltas, the LCI_{wn} and LSI_{wn} can be calculated with a weighted sum of the deltas. The DV delta with the highest value, i.e., the least violated index, gets a weight of 1. The weights then increase exponentially with a factor of $\lambda = 10$ [11]. Afterwards, the weights are normalized to be within the range $[0, 1]$. The LCI_{wn} and LSI_{wn} then become:

$$\begin{aligned} LCI_{wn}(R\boldsymbol{t}) =& w_c\left(\delta_{cdn}\left(D_{90\%}^{prostate}\right)\right) + w_c\left(\delta_{cvn}\left(V_{100\%}^{prostate}\right)\right) + w_c\left(\delta_{cvn}\left(V_{73\%}^{vesicles}\right)\right) \\ LSI_{wn}(R\boldsymbol{t}) =& w_s\left(\delta_{sn}\left(V_{200\%}^{prostate}\right)\right) + w_s\left(\delta_{sn}\left(V_{150\%}^{prostate}\right)\right) + w_s\left(\delta_{sn}\left(D_{1cm^3}^{rectum}\right)\right) + \\ & w_s\left(\delta_{sn}\left(D_{2cm^3}^{rectum}\right)\right) + w_s\left(\delta_{sn}\left(D_{1cm^3}^{bladder}\right)\right) + w_s\left(\delta_{sn}\left(D_{2cm^3}^{bladder}\right)\right) + \\ & w_s\left(\delta_{sn}\left(D_{0.1cm^3}^{urethra}\right)\right) + w_s\left(\delta_{sn}\left(D_{30\%}^{urethra}\right)\right) \end{aligned}$$

Where $w_c(\delta)$ and $w_s(\delta)$: (2.9)

$$\begin{aligned} w_c\left(\delta\right) =& \frac{\lambda^{r-1}}{\sum_{i=1}^{|\Delta_c^s|}\lambda^{i-1}}\delta \text{ with } r \text{ s.t. } \Delta_c^s[r] = \delta \\ w_s\left(\delta\right) =& \frac{\lambda^{r-1}}{\sum_{i=1}^{|\Delta_s^s|}\lambda^{i-1}}\delta \text{ with } r \text{ s.t. } \Delta_s^s[r] = \delta \end{aligned}$$

Here r is a one-based index that is used to indicate the position of a DV index in the sorted lists Δ_c^s and Δ_s^s. The LCI_{wn} and LSI_{wn} objectives now have upper bounds of 1.0, meaning that the plans that satisfy all DVIs are within the range $[0, 1]$ for both objectives.

2.8.4 Bézier Curve Specific Exploitable Properties

In order to adapt the bi-objective problem formulation for the use of Bézier curve parameterizations, a few specific adaptations are used. These state a reference point for the hypervolume calculations, a separate linkage model and exploitable properties specific to the Bézier curves [72]. All of the adaptations were researched for the old Amsterdam UMC protocol defined in the previous section.

Reference point. The bi-objective formulation for HDR brachytherapy is formulated as a trade-off between the LCI and LSI. These two objectives are calculated based on the dose distribution, whereby the aim is to obtain a positive value for both. However, not in all cases can both the LCI and LSI obtain positive values in order to obtain clinically acceptable treatment plans [11]. Therefore, a reference point is set such that some margin is taken into account for slightly worse LCI and LSI values. The reference point for the hypervolume calculation is set to $r = (-0.04, -0.2)$ [72].

With the reference point and the use of the hypervolume problem formulation employed in BezEA, no further constraint functions are required to ensure that the resulting treatment plans always satisfy $LCI \geq -0.04$ and $LSI \geq -0.2$. The use of the reference point itself is sufficient to guide the search due to the uncrowded distance constraint used in BezEA. Furthermore, the built in margins for LCI and LSI are smaller than that in the original GPU accelerated research [11] as plans with LCI < -0.04 were indicated to not be of clinical interest [50].

Test solution interpolation. As the RV-GOMEA algorithm, which is at the basis of BezEA, is capable of partial evaluations, it can evaluate only the part of a solution that has changed [11]. A linkage model is used to guide the algorithm in choosing which decision variables are to be changed at the same time. This ensures that the algorithm is capable of solving the HDR brachytherapy problem with a smaller population size and subsequently less time [43].

It is inefficient to ignore the relation that if one of the decision variable values of the control points is changed, the corresponding dwell time for all p plans changes. Thus, first, each of the q control points has their i-th decision variable grouped together, resulting in ℓ clusters of q decision variables. Subsequently, the UPGMA clustering method can be used to construct a linkage model. UPGMA iteratively merges the ℓ clusters of q decision variables until a full FOS element is constructed. Furthermore, a lower bound is set to the minimum FOS element size as changing just one dwell time out of the hundreds of dwell times used in HDR brachytherapy leads to only marginal improvements, while still requiring a relatively large amount of computational

effort for the partial evaluation. The lower bound is set to $5q$ for all of the clusters, similar to MO-RV-GOMEA with a lower bound of 5 [11].

Last of the exploitable properties is that of an improved evaluation method that exploits the linearity of the dose distribution. Since the dose distributions $\boldsymbol{t}_i$ of each of the p test solutions on a Bézier curves is a linear interpolation between the dose distributions of the q control points $\boldsymbol{c}_j$, a speedup can be obtained in terms of computing the dose calculation points. Because p is generally larger than q, the following formula can be used:

$$\boldsymbol{t}_i = B\left(\frac{i-1}{p-1}; C_q\right) = \sum_{j=1}^{q} \binom{q-1}{j-1} (1-t)^{q-j-1} t^{j-1} \frac{i-1}{p-1} \boldsymbol{c}_j = \sum_{j=1}^{q} b_{j,q}^{i,p} \boldsymbol{c}_j$$

The formula for the dose calculation points then becomes [72]:

$$\boldsymbol{d}_i = R\boldsymbol{t}_i = R\left(\sum_{j=1}^{q} b_{j,q}^{i,p} \boldsymbol{c}_j\right) = \sum_{j=1}^{q} b_{j,q}^{i,p} \left(R\boldsymbol{c}_j\right) \tag{2.10}$$

With the use of the formula above, the product $R\boldsymbol{c}_j$ is independent of the test solution i. Thus, the number of matrix-vector multiplications is hereby reduced from p to q, even though the LCI and LSI values still have to be computed p times. The original BezEA algorithm actually optimized the dwell times $\boldsymbol{x} = \sqrt{\boldsymbol{t}}$. This results in more freedom to fit a Bézier curve close to $\boldsymbol{t}_i = 0$, as not all dwell positions are used in HDR brachytherapy.

Chapter 3

Bézier Evolutionary Algorithm Revision

The original Bézier Evolutionary Algorithm has a few very desirable properties for decision makers. It also showed promising results when compared to other algorithms, even for the difficult HDR Brachytherapy problem. However, during the course of this thesis, several issues arose. These are explained in this chapter together with their proposed solutions.

First of the issues that have been research is the need for an adjusted navigational order, due to which some solutions that lie on the approximation front were erroneously rejected. Second, an adapted constraint function that includes scaling to even further prevent the collapse of approximation fronts. Lastly, a revised evaluation for HDR Brachytherapy as an accidental error was made in the derivation of the evaluation function for test solutions of Bézier paremeterized solution sets.

3.1 Adjusted Navigational Order

One of the initial ideas in order to cluster Bézier solutions was to cluster them based on their hypervolume values. As these fitness values are single-objective, the original HVC method seemed applicable. However, initial results were not promising. The results of the clustering step were ambiguous, with Bézier solutions from both modes being grouped together. Thus, the resulting clusters seemed to completely ignore mode boundaries for which the clustering is needed. The usage of the Hausdorff or Fréchet distance, i.e., distance measures for subsets in a space, did not solve the issue.

In order to gain insight into why the clustering step did not show the results that were expected, a landscape analysis was performed. As an unlimited amount of combinations of two control points for a Bézier curve could be made in a 2D landscape, one of the control points is fixed to a specific position, whereas the other control point is continuously adapted. This landscape analysis is performed for the Mindist problem, described as the multi-objective multi-modal optimization problem in the background of Section 2.1.3.

Bézier solution sets $S_{p,q}(C_q)$ always have p test solutions as per Definition 11. However, depending on the way p test solutions are distributed in the search space and their associated objective values, Algorithm 8 returns the approximation set $\mathcal{A}_{p,q,o_{nb}}$ that is a subset of $S_{p,q}(C_q)$. For reference, Figure 3.1 shows the evaluation of a Bézier solution set, if it would be equal to the Pareto set, in order to show how the hypervolume is calculated in Figure 3.2. Figure 3.2 itself shows the landscape analysis for Bézier solutions set. The left figure shows the hypervolume values, whereas the right figure shows the number of test solutions on the approximation front, i.e., in the approximation set $\mathcal{A}_{p,q,o_{nb}}$, for the hypervolume calculation of a Bézier solution. Combined these figures indicate a correspondence between the number of test solutions on the approximation front of a Bézier solution set and the hypervolume value associated to said Bézier solution set, which is displayed by the contour lines in the hypervolume landscape corresponding with the number of solutions on the approximation front. The correspondence of these two graphs instigated the need to check two Bézier solution sets and determine how the hypervolume calculation of the approximation front of a Bézier solution set exactly occurs.

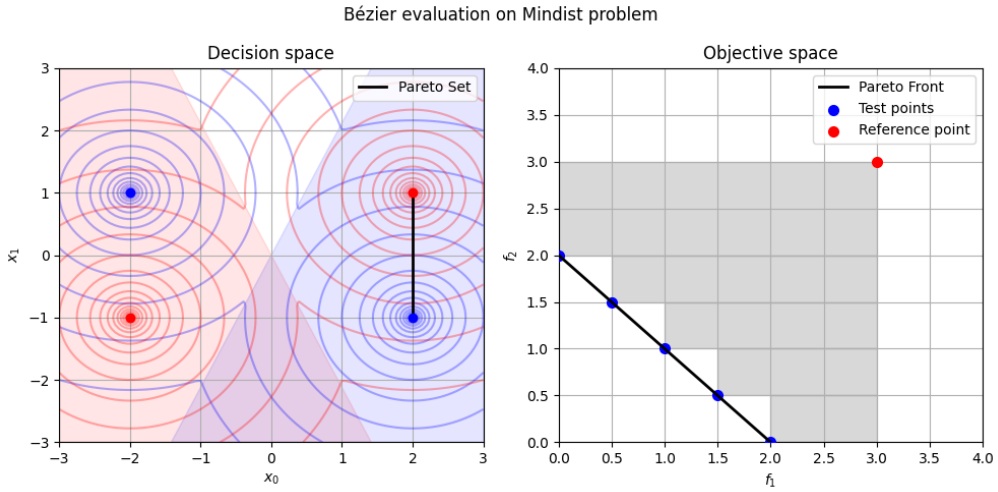


Figure 3.1: Multi-modal multi-objective optimization problem Mindist. The left figure shows mode boundaries as shaded areas for corresponding objectives, which are depicted through the contour lines, and the Pareto set as the black line for the mode shaded in blue. Right figure shows the evaluation of the Pareto set for 5 test points if it were a Bézier solution. The volume of the grey shaded region is the hypervolume for the Pareto set.

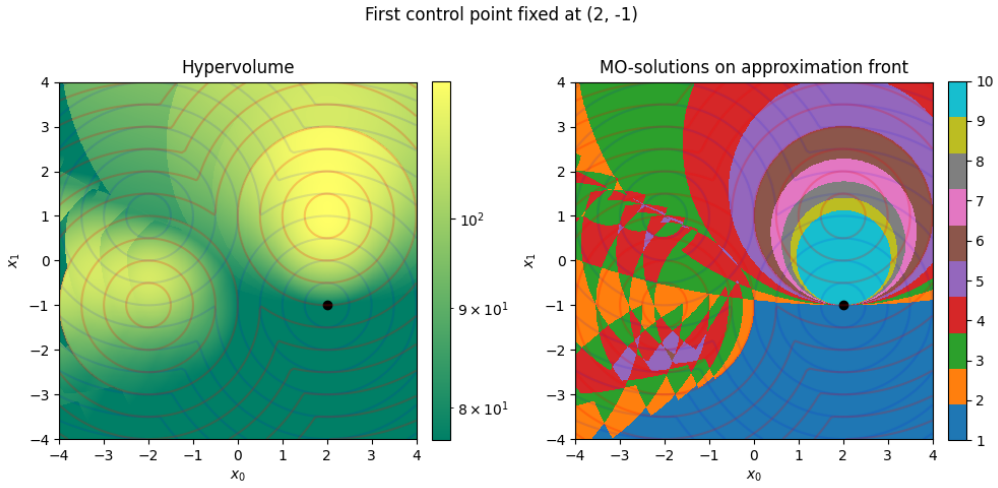


Figure 3.2: Figure showing the hypervolume fitness landscape for Bézier solutions and the number of MO solutions on the approximation front. The black dot is one of the control points that has been fixed at that specific location. The color at a location in the hypervolume figure represents the hypervolume value if a Bézier solution were drawn from the black point to the point where the value is read. The figure clearly indicates the presence of sharp differences between adjacent positions in a landscape that is expected to be continuous. The right figure is read in a similar manner, but shows the number of solutions on the approximation front. The figures indicate a correspondence between the number of solutions on the approximation front and the hypervolume fitness value associated to it.

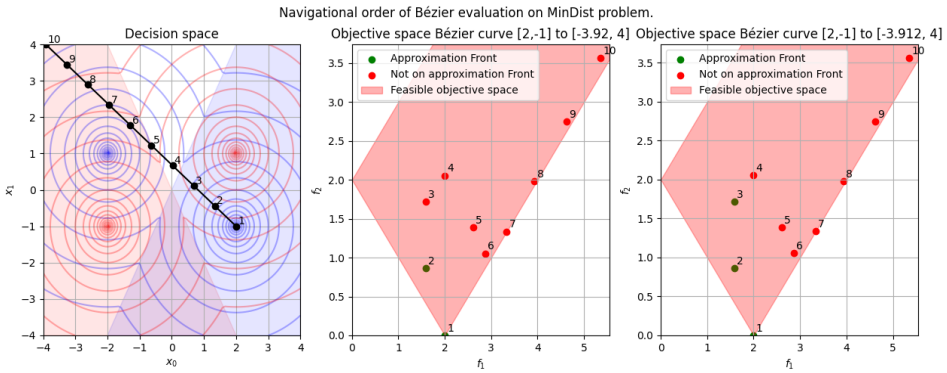


Figure 3.3: Figure shows the evaluation of two Bézier solutions at both sides of one of the clear boundaries in Figure 3.2. The left-most figure shows where the Bézier solutions are in decision space and the two right-most figures show the evaluation of the test solutions, annotated with their index, along the Bézier curves in objective space.

Manual inspection of the results complied with the sharp difference as seen in Figure 3.2. Figure 3.3 shows the evaluation of two Bézier solutions set at both sides of one of the aforementioned boundaries in the hypervolume fitness landscape of Figure 3.2. Even though no noticeable issue arises from the objective space plots of Figure 3.3, the usage of Algorithm 8 is problematic. This algorithm gives the navigational order for the test solutions on the Bézier curve. At line 4 of that algorithm, it loops over the indices of the test solutions starting from the one with best objective value for f_1 and only increasing in index to those with worse objective values for f_1. However, Figure 3.3 shows a decrease in index when traversing the solutions on the front from the best objective values for f_1 to the worst.

The source code of the author of BezEA partly ignored the code of Algorithm 8 and started by adding the test solution with index 1 in all cases instead of the test solution fittest in f_1. Next, the indices of solutions were traversed in sorted order based on decreasing f_1 values. Due to the defined order of control points whereby $f_1(c_1) < f_1(c_q)$, the author presumable assumed that the first control point, equal to test solution with index 1, is always the fittest in terms of f_1 of all test solutions. However, in both of the cases as seen in Figure 3.3, a solution with a higher index is the fittest in terms of f_1. Another difference can be found in line 4 of Algorithm 8, since the authors implementation did not loop over the solutions in order $\{\eta, ..., p\}$. On the contrary, it looped over the solutions in order of f_1 fitness values, even though it enforced increasing indices in this for loop.

Both the differences in assuming what solution is to be put in the approximation set first and the adjusted ordering of indices meant that in the middle plot of Figure 3.3 only test solution with index 1 was used for the hypervolume calculation. Thus, it erroneously ignored the test solution with index 2. In the case of right most plot of the same figure, only solutions with index 1 and 2 were included for the hypervolume calculation, whereas the solution with index 3 is ignored. The omission of these solutions results in a wrong hypervolume calculation as the objective space that is dominated by one of those solutions is not taken into account.

To prevent the issue explained in the previous paragraphs a new navigational order function is proposed. This function requires either monotonically increasing or decreasing indices of test solutions on the approximation front. Furthermore, all indices are enforced to be between the indices of the fittest and least fit test solutions on the approximation front in terms of objective function f_1. These requirements still comply with the originally proposed idea behind the navigation order, whereby it starts with the best solution for objective f_1 and traverses all of the solutions in the approximation set $\mathcal{A}(S_{p,q}(C_q))$ up until the best solution in f_2 has been reached. Algorithm 10 shows the newly proposed navigational order function for Bézier curves. Figure 3.4 shows the hypervolume fitness landscape in a similar setup to Figure 3.3, but with the use of the new navigational order of Algorithm 10. This figure indicates that the newly proposed algorithm correctly identifies the solutions on the approximation front, thereby creating a continuous hypervolume fitness landscape as expected.

Algorithm 10: $[\mathcal{A}_{p,q,o_{nb}}, o_{nb}]$ = `NavigationalOrder`$(S_{p,q}(C_q), o^{bez})$

Input: Bézier solution set $S_{p,q}(C_q)$, with intrinsic order o^{bez}
Output : Approximation set $\mathcal{A}_{p,q,o_{nb}}$

```
// Sort order of indices on approximation front of Bézier solution set.
```
$o^{temp} \leftarrow$ indices of $o^{bez} \in \mathcal{A}(\mathcal{S}_{p,q})$ sorted on f_1
$\eta \leftarrow o_1^{temp}$
$\theta \leftarrow o_{|o^{temp}|}^{temp}$
$o_{nb} \leftarrow o_\eta^{bez}$
$\mathcal{A}_{p,q,o_{nb}} \leftarrow \{\boldsymbol{x}_{o_\eta^{bez}}\}$
$\gamma \leftarrow \eta$
for $i \in \{2, ..., |o^{temp}| - 1\}$ **do**
  $j \leftarrow o_i^{temp}$
  `// Check if front should be monotonically increasing or decreasing in indices.`
  if $(\eta < \theta)$ **then**
    `// Check if new index is between front endpoints and monotonically increasing.`
    if $(\eta < j < \theta)$ *and* $j > \gamma$ **then**
      $\gamma \leftarrow j$
      $o_{nb} \leftarrow [o_{nb}; o_j^{temp}]$
      $\mathcal{A}_{p,q,o_{nb}} \leftarrow \mathcal{A}_{p,q,o_{nb}} \cup \{\boldsymbol{x}_{o_j^{bez}}\}$
    end
  end
  if $(\theta < \eta)$ **then**
    `// Check if new index is between front endpoints and monotonically decreasing.`
    if $(\theta < j < \eta)$ *and* $j < \gamma$ **then**
      $\gamma \leftarrow j$
      $o_{nb} \leftarrow [o_{nb}; o_j^{temp}]$
      $\mathcal{A}_{p,q,o_{nb}} \leftarrow \mathcal{A}_{p,q,o_{nb}} \cup \{\boldsymbol{x}_{o_j^{bez}}\}$
    end
  end
end
$o_{nb} \leftarrow o_\theta^{bez}$
$\mathcal{A}_{p,q,o_{nb}} \leftarrow \{\boldsymbol{x}_{o_\theta^{bez}}\}$
return $\mathcal{A}_{p,q,o_{nb}}$, o_{nb}

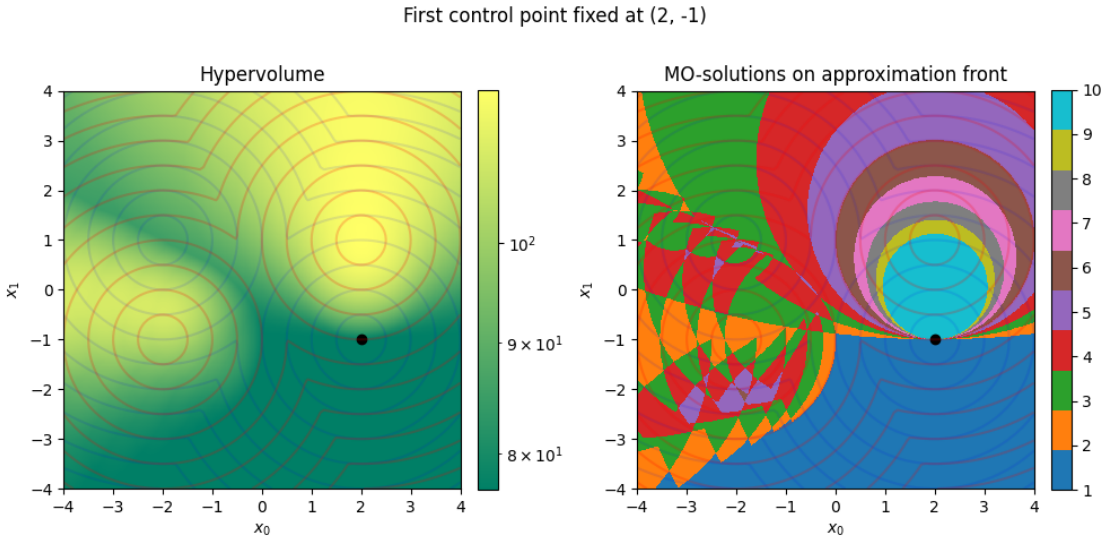


Figure 3.4: Figure showing the hypervolume fitness landscape for Bézier solutions and the number of MO solutions on the approximation front. The black dot is one of the control points that has been fixed at that specific location. The color at a location in the hypervolume figure represents the hypervolume value if a Bézier solution were drawn from the black point to the point where the value is read. With the use of the new `NavigationalOrder` function of Algorithm 10, sharp differences are no longer visible and a continuous hypervolume fitness landscape is observed.

3.2 High-Dose-Rate Brachytherapy Evaluation Revision

As part of the preparation of the creation of the new algorithms proposed in this thesis, a trial run of the original BezEA algorithm was performed on the HDR brachytherapy problem. The results were imported into a visualization tool whereby the approximation set for a specific patient could be inspected. However, the results that originated from the BezEA algorithm showed promising results in literature [72], which no longer held true when they were imported into the visualization tool.

Figure 3.5 shows the results of using Equation 2.10 for the Bézier specific interpolation of HDR brachytherapy from section 2.8.4 on the left. The grey region is the region in which all DVI criteria are met, which indicates that a plan could be clinically acceptable. The right plot shows the actual results when each of the p test solutions along the curve has been re-evaluated separately. The sharp contrast between the plots clearly shows that actually no DVI criteria are met, thus showing that the used interpolation formula is not valid.

As BezEA optimizes on the square root of the dwell times, the interpolation is no longer linear in shape, but quadratic. Therefore, the interpolation formula has to be adjusted. The derivation of the new interpolation has been performed for the simplest cases of $q \in \{2, 3\}$, where one of the p test solutions $\boldsymbol{x}_i$ is given by:

$$\boldsymbol{x}_i = \begin{cases} (1-t)\boldsymbol{c}_1 + t\boldsymbol{c}_2 & q = 2, t = \dfrac{i-1}{p-1} \\ (1-t)^2\boldsymbol{c}_1 + 2(t-1)t\boldsymbol{c}_2 + t^2\boldsymbol{c}_3 & q = 3, t = \dfrac{i-1}{p-1} \end{cases}$$

The derivation for the interpolation of the dose calculation points with $q = 2$ is:

$$\begin{aligned} \boldsymbol{d}_i &= R\left(\boldsymbol{x}_i\right)^2 \\ &= R\left((1-t)\boldsymbol{c}_1 + t\boldsymbol{c}_2\right)^2 \\ &= (1-t)^2 R(\boldsymbol{c}_1)^2 + 2(1-t)tR\boldsymbol{c}_1\boldsymbol{c}_2 + t^2 R(\boldsymbol{c}_2)^2 \\ &= (1-t)^2 d_{c_1} + 2(1-t)tR(\boldsymbol{c}_1 \cdot \boldsymbol{c}_2) + t^2 d_{c_2} \\ &\neq (1-t)d_{c_1} + td_{c_2} (\text{According to Equation 2.10.}) \end{aligned} \tag{3.1}$$

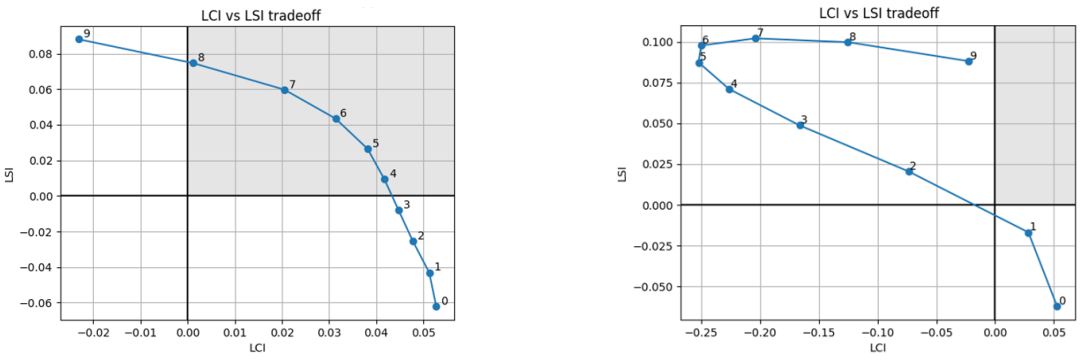


Figure 3.5: Figures showing the result of Bézier specific interpolation using the square root of the dwell times on the left versus the actual approximation front if all test solutions are re-evaluated separately on the right. Grey area entails having surpassed all clinical aims.

And the derivation for the interpolation of the dose calculation points with $q = 3$ is:

$$\begin{aligned}
\boldsymbol{d}_i =& R\left(\boldsymbol{x}_i\right)^2 \\
=& R\left((1-t)^2\boldsymbol{c}_1 + 2(t-1)t\boldsymbol{c}_2 + t^2\boldsymbol{c}_3\right)^2 \\
=& (1-t)^4 R(\boldsymbol{c}_1)^2 + 4t(1-t)^3 R(\boldsymbol{c}_1 \cdot \boldsymbol{c}_2) + 2t^2(1-t)^2 R(\boldsymbol{c}_1 \cdot \boldsymbol{c}_3) + \\
& 4t^2(1-t)^2 R(\boldsymbol{c}_2)^2 + 4t^3(1-t) R(\boldsymbol{c}_2 \cdot \boldsymbol{c}_3) + t^4 R(\boldsymbol{c}_3)^2 \\
=& (1-t)^4 d_{\boldsymbol{c}_1} + 4t(1-t)^3 R(\boldsymbol{c}_1 \cdot \boldsymbol{c}_2) + 2t^2(1-t)^2 R(\boldsymbol{c}_1 \cdot \boldsymbol{c}_3) + \\
& 4t^2(1-t)^2 d_{\boldsymbol{c}_2} + 4t^3(1-t) R(\boldsymbol{c}_2 \cdot \boldsymbol{c}_3) + t^4 d_{\boldsymbol{c}_3}
\end{aligned}$$

Using the new derived interpolation formula, the number of matrix-vector multiplication can still be reduced in case a similar setup to that of BezEA is used for the HDR brachytherapy problem. There, $p = 10$ test solutions and $q = 2$ control points were used to optimize for [72]. Using the new interpolation, this implies that the number of matrix-vector multiplications is reduced from 10 to 3, with the inclusion of a new intermediate point created by the matrix vector multiplication $R(\boldsymbol{c}_1 \cdot \boldsymbol{c}_2)$. In general, the number of matrix-vector multiplications is reduced from p to $\frac{q(q+1)}{2}$. This means that with 4 control points, a reduction is no longer possible with $p = 10$ test solutions as then the number of evaluations for the interpolation equals the number of test solutions.

Using the new derived interpolation for Bézier solutions from Equation 3.1, the BezEA algorithm was again tested on the HDR brachytherapy problem. Figure 3.6 shows the results of using the new interpolation formula. The left figure shows the results according to the new formula, whereas the right figure shows the results if all of the separate test solutions would be evaluated separately. The correspondence between the figures shows that the newly derived interpolation equation functions correctly.

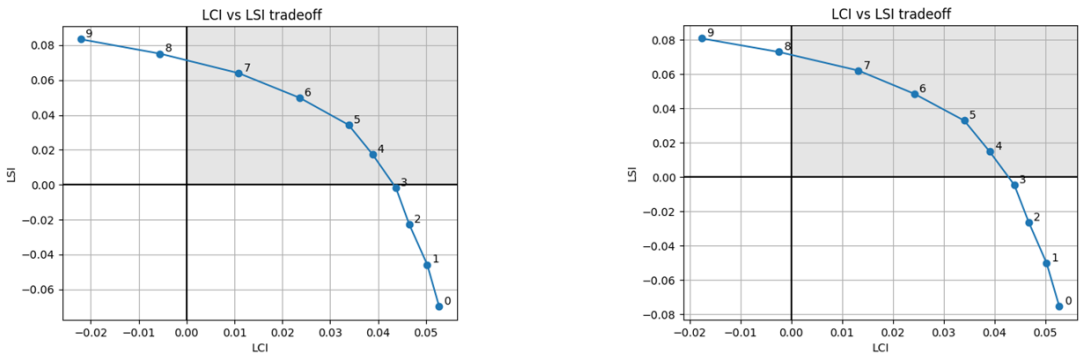


Figure 3.6: Figures showing the result of the newly defined Bézier specific interpolation equation when using the square root of the dwell times on the left versus the actual approximation front if all test solutions are re-evaluated separately on the right. Grey area entails having surpassed all clinical aims.

3.3 Constraint Scaling

With the use of the new interpolation formula from the previous section, the need arose to test BezEA on a larger set of patients in order to re-evaluate its effectiveness. However, initial results on one of the patients brought to light a further issue. In that particular case, the BezEA algorithm seemed incapable of reaching the reference point as set, due to which it achieved a hypervolume of 0 and subsequently no acceptable treatment plans. Furthermore, the patient where this issue occurred was a difficult case in the sense that other algorithms were also incapable of obtaining treatment plans that satisfy all clinical aims. Not only were further tests performed with the square root encoding of the dwell times, as the newly derived interpolation formula has an influence on the evaluation of a Bézier curve, but also the problem formulation whereby the dwell times of a treatment plan are encoded directly. In both situations similar results were found, leading to the conclusion that although the new interpolation seemed to be working, further issues were still present. Figure 3.7 shows the resulting approximation front for the mentioned patient in the left plot.

Upon further investigation of the issue, it became clear that the constraint values never reached zero as can be seen in the right plot of Figure 3.7. Over a run of up to two hours of compute time, the Bézier constraint of Algorithm 9, that tries to unfold the Bézier curve, overtook the uncrowded distance, that is used in the same constraint function to force all solutions to the undominated region of the search space. This issue required further insight as to what exactly was happening during the optimization. To do so, the approximation fronts at each generation of BezEA were plotted for several patients.

Figure 3.8 shows the approximation fronts that are parameterized by Bézier solution sets for two different patients at the 5th generation of the BezEA algorithm. The left figure shows the correct unfolding of the approximation fronts of a patient for which acceptable plans could be obtained, whereas the right figure indicates that the approximation fronts have all collapsed for the patient where no acceptable plans could be obtained. These plots, combined with the right plot of Figure 3.7 that shows the course of constraint values, led to the conclusion that the Bézier constraint seems to be too powerful as it seems to pull all Bézier solutions close to each other and fold over on each other in objective space. This constraint function is defined by:

$$c = \sum_{j=1}^{p} ||f(\boldsymbol{x}_{o_j^{bez}}) - f(\boldsymbol{x}_{o_{j+1}^{bez}})|| \text{ if } o_j^{bez} \notin o_n b \text{ or } o_{j+1}^{bez} \notin o_n b$$

In order to overcome this limitation, the Bézier constraint algorithm was adapted to reduce the influence of the part of the constraint that tries to unfold the curves, as it seemed to be doing the opposite. By reducing its influence, the uncrowded distance is a more prevalent part of the constraint function, which was hypothesized to lead to reaching the reference point earlier in optimization. After the reference point was reached, the unfolding can still take place as constraint domination forces lower constraint values over better objective values.

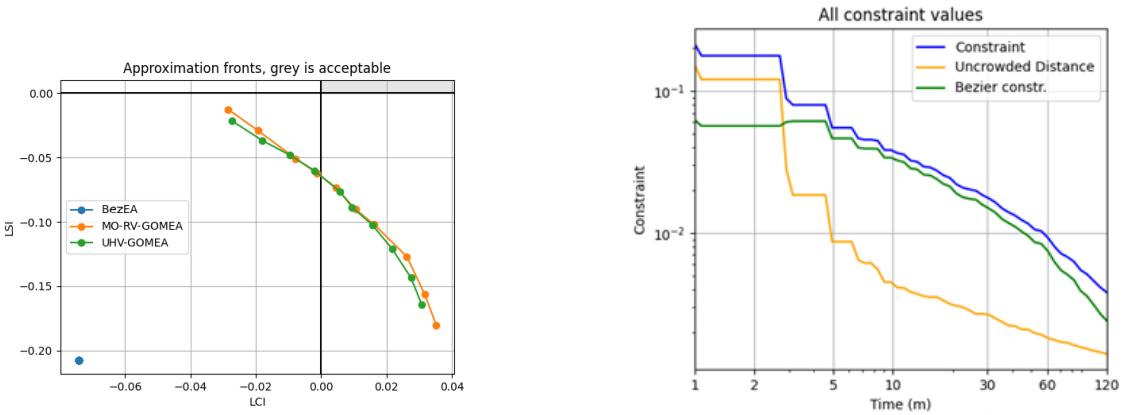


Figure 3.7: Left figure showing the resulting approximation front for MO-RV-GOMEA and UHV-GOMEA and collapsed approximation front of BezEA for the HDR brachytherapy problem with direct encoding of dwell times. Results were similar with square root encoding of dwell times. Right figure shows the average constraint values of the entire BezEA population over an entire run of the algorithm on the same patient as the left figure.

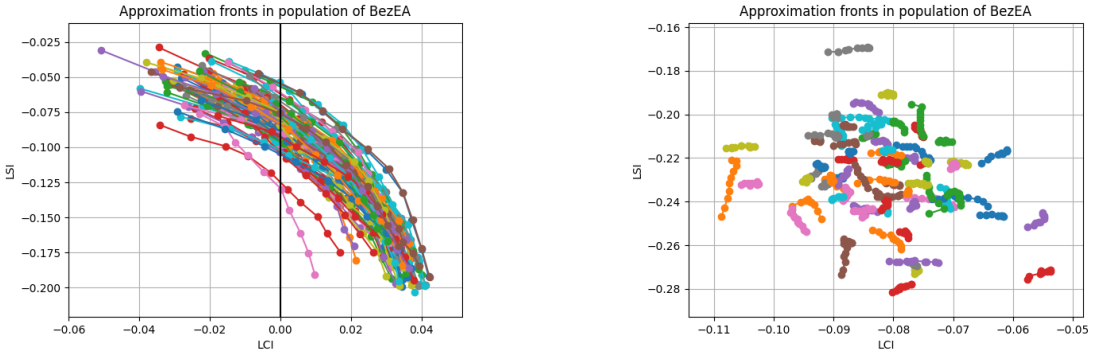


Figure 3.8: Two figures showing the approximation fronts parameterized by Bézier solution set for two different patients. Both figures show the entire population at the 5th generation. Left figure displays intended behaviour, whereas the right figure shows the collapse of the approximation fronts in objective space.

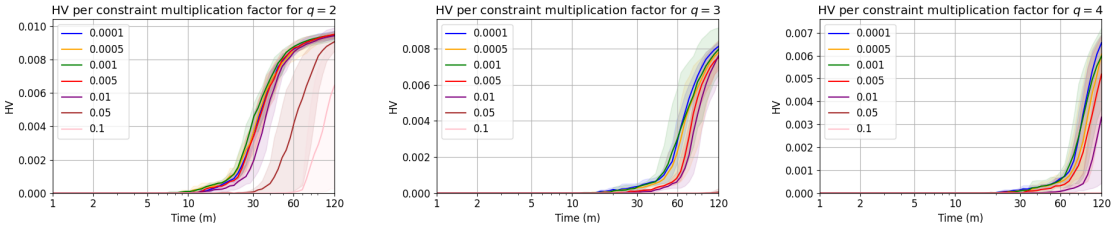


Figure 3.9: Three plots showing the obtained hypervolume for Bézier solutions sets with $q \in \{2, 3, 4\}$ (from left to right) for different values with which the Bézier constraint is multiplied on a patient for which no clinically acceptable plans can be obtained. The figures indicate that multiplying the Bézier constraint with 10^{-q} will lead to obtaining acceptable hypervolume values before the end of the run.

As the reduction of the Bézier constraint could lead to the introduction of a new variable, a test was performed in order to determine a rule of thumb as to what value seems appropriate. Several values with which the Bézier constraint would be reduced were tested for $q \in \{2, 3, 4\}$ control points on the patient for which approximation sets could not be found. The results of these tests can be seen in Figure 3.9.

Figure 3.9 indicates that in order to prevent the inclusion of another parameter, the unfolding constraint can be multiplied by 10^{-q}. Thus, reducing its influence further with the use of more control points, where the rest of the constraint is kept similar. The new constraint function, as incorporated into Algorithm 9 from the background section, can be seen in line 5 of Algorithm 11 below.

Algorithm 11: $[\mathcal{C}] =$ `BézierConstraint`$(S_{p,q}(C_q),\, o^{bez})$

Input: Bézier solution set $S_{p,q}(C_q)$, with intrinsic order o^{bez}
Output : Constraint value $\mathcal{C} \geq 0$

1 $\mathcal{A}_{p,q,o_{nb}},\, o_{nb} \leftarrow$ NavigationalOrder$(S_{p,q}(C_q),\, o^{bez})$ `// See Algorithm 10`
2 $\mathcal{C} \leftarrow \frac{1}{|S_{p,q}(C_q)|} \sum_{\boldsymbol{x} \in S_{p,q}(C_q)} ud_f(\boldsymbol{x}, \mathcal{A}_{p,q,o_{nb}})$
3 **for** $j \in \{1, ..., |S_{p,q}(C_q)| - 1\}$ **do**
4 **if** $\boldsymbol{x}_{o_j^{bez}} \notin o_{nb}$ *or* $\boldsymbol{x}_{o_{j+1}^{bez}} \notin o_{nb}$ **then**
5 $\mathcal{C} \leftarrow \mathcal{C} + ||f(\boldsymbol{x}_{o_j^{bez}}) - f(\boldsymbol{x}_{o_{j+1}^{bez}})|| * 10^{-q}$ `// Eucledian distance in` $\mathbb{R}^m$
6 **end**
7 **end**
8 **return** $\mathcal{C}$

Chapter 4

Multi Modal-Bézier Evolutionary Algorithm

In this section the Multi-Modal Bézier Evolutionary Algorithm, abbreviated as MM-BezEA, is described. It has been created through the combination and modification of techniques that have been explained in the previous chapters. The sections below try to answer the first two sub questions of research question 1. Namely 1.1 and 1.2, whereby Bézier curve parameterizations can be clustered and kept in an elitist archive. The overarching research question 1 itself will be attempted to be answered in Chapter 6, where empirical results for MM-BezEA will be shown. The most notable contributions of this chapter are, in order of appearance:

1. The adjustments implemented in Hill Valley Clustering in order to apply it to Bézier curve parameterizations.
2. The initialization of approximation sets within niches, to prevent ambiguous clustering.
3. The use of a newly created Bézier Hill Valley Test for the inclusion of Bézier solution sets in the elitist archive.

4.1 Clustering Approximation Sets

The Bézier solution sets $S_{p,q}(C_q)$ are evaluated according to the Bézier curve optimization problem of Definition 12. Since the objective is a scalar that denotes the hypervolume value, the HVC approach is essentially applicable in a rather straightforward manner because it intuitively allows the clustering of single-objective problems. However, the approximation set $\mathcal{A}_{p,q,o_{nb}}$ that is used in the HV calculation only considers the undominated indices of the Bézier solution set $S_{p,q}(C_q)$ as defined in the navigational order o_{nb}. Therefore, the objective value given to the solution set $S_{p,q}(C_q)$ seems highly dependant on how many dominated solutions there are on the Bézier curve, due to its orientation and length through the decision space. Employing either the Hausdorff or Fréchet distance measures for the distances between sets instead of Euclidean distance in the concept of the nearest better tree did not change the results. Figure 4.1 shows the resulting clusters when clustering two populations of Bézier solution sets based on their hypervolume values. The figure indicates that Bézier solution sets from both modes are clustered together, which is not the intended behaviour.

To enable the clustering of Bézier solution sets $S_{p,q}(C_q)$, the idea behind MO-HVC can be used on the set of control points C_q, because each of these is a single solution as normally defined in MO optimization. Also, since a solution set is defined to be deteriorating in f_1 and improving in f_1 according to o_{nb}, the order of the control points is inverted if $f_0(\boldsymbol{c}_1) < f_0(\boldsymbol{c}_q)$ does not hold [72]. Accordingly, the i-th Bézier solution can be designated to be in the same niche as the j-th Bézier solution if their control points $\boldsymbol{c}_l^i \in C_q^i$ and $\boldsymbol{c}_l^j \in C_q^j$ for $l = \{1, ..., q\}$ are in the same niche. In a general sense, the same HVC approach as used in HillVallEA is employed. However, inspiration has been taken from the MO-HVC approach to produce a new test for Bézier solution sets. Algorithm 12 displays how the Hill-Valley Test of Algorithm 3 is used in the new Bézier Hill-Valley Test.

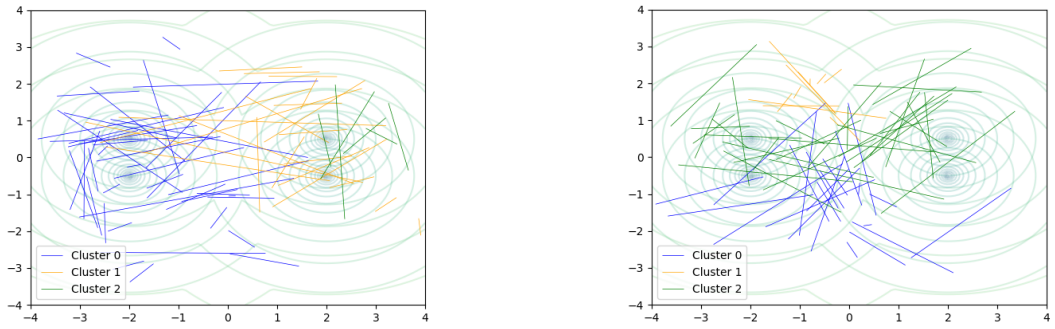


Figure 4.1: Two figures showing the clustering based on the hypervolume value of Bézier solution sets on the MinDist problem. Both figures show different, randomly initialized, populations. The figures indicate that Bézier solution sets from both modes are clustered together.

Algorithm 12: $[B]$ = `Bezier-HillValleyTest`($\mathbf{S_i}$, $\mathbf{S_j}$, N_t, f)

Input: Solutions sets $\mathbf{S_i}$, $\mathbf{S_j}$, int N_t, objective functions $f_0, ..., f_{m-1}$
Output : Whether $\mathbf{S_i}$ and $\mathbf{S_j}$ belong to the same niche

```
for l = 1, ..., q do
    c_{i,l} = control point l of S_i
    c_{j,l} = control point l of S_j
    // Check if c_{i,l} and c_{j,l} are in same niche for all m objectives
    for k = 0, ..., m − 1 do
        // see Algorithm 3
        if HillValleyTest(c_{i,l}, c_{j,l}, N_t, f_k) then
            return false
        end
    end
end
return true
```

4.2 Initialization Within Niches

The original BezEA algorithm initializes all solution sets by sampling from a uniform distribution over the search space. As it is an MOEA that was not designed for multi-modal optimization, the uniform initialization allows solution sets to be initialized within or in between any niche(s). Clustering these solutions with the newly introduced Bézier HVT will result in finding a large number of separate niches as each control point has to be in the same mode.

To prevent this, a new initialization method for Bezier solution sets is proposed to enforce their initialization within a niche. First, an iteration of MO-HVC is run on a set of $q * N$ solutions, N being the population size, that is sampled from a uniform distribution over the search space. The resulting clusters from the MO-HVC step include all of the $\boldsymbol{x}_{test}$ solutions that come from the application of Algorithm 3. Using the outcome of MO-HVC, Bézier solution sets can be initialized within niches. A visualization of the result can be seen in Figure 4.2.

As there are several small clusters on the boundary of the niches in Figure 4.2, selection is added as the second step of the initialization procedure to focus on the larger clusters that have fitter solutions. Selection is performed on the clusters proportional to their size in order to reduce their combined size, including test solutions, down to $q * N$. Only those clusters that are above a minimum size of $10\sqrt{\ell}$ are considered. The minimum size is based on the recommended population size for univariate factorization used in the Adapted Maximum Likelihood Gaussian Model Iterated Density-estimation Evolutionary Algorithm (AMaLGaM-IDEA or abbreviated as AMaLGaM) [9]. This is due to the fact that RV-GOMEA, which is used as the core search algorithm in MM-BezEA, is based on AMaLGaM and the recommended population size is a low estimate required for a good approximation for univariate factorizations. Thus, the smallest clusters below the minimum size of $10\sqrt{\ell}$ are discarded and others will have applied a selection operator to only retain the fittest solutions.

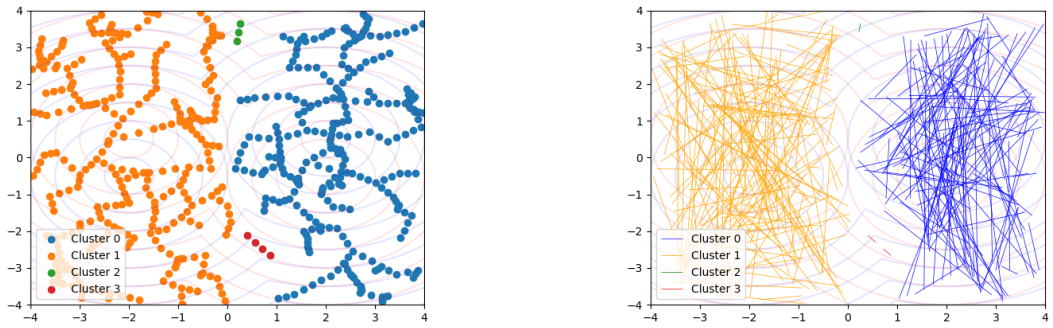


Figure 4.2: Initialization of Bézier solutions set ($q = 2$) for MinDist. Left figure shows initial MO-HVC of separate MO-solutions. Right figure shows the initialized Bézier solution sets using the clustering result of the left figure.

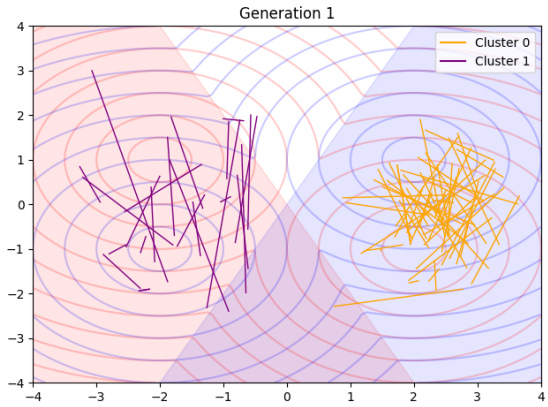


Figure 4.3: Initialization of Bézier solutions ($q = 2$) for MinDist problem.

In the last step of the initialization, the Bézier solution sets $S_{p,q}(C_q)$ are initialized by randomly choosing q solutions as the control points. These solutions come from one single cluster $\mathcal{C}$ that originates from the second step. The final result of the entire initialization procedure can be seen in Figure 4.3. This figure shows that the clusters are situated somewhat closer around the optima and both of the niches are clearly separated.

Answer to Research Question 1.1

Figure 4.3 also shows that research question 1.1 has been answered. Since, the combination of Algorithm 12 combined with the initialization procedure described in this section is a solution for the question of how to apply Hill Valley Clustering to cluster Bézier curve parameterizations.

4.3 Elitist Archive

As a Bézier solution set $S_{p,q}(C_q)$ is scored using a scalar that denotes the hypervolume, the elitist archive would traditionally only hold a single Bézier solution set with the best hypervolume value. The goal of MM-BezEA is to find multiple, and possibly all, optima, one for each niche. Therefore, it needs a way of keeping track of the distinct approximation sets for each niche.

Inspiration from the HillVallEA method can be taken in the form of keeping track of all global optima. However, the goal is to find all optima for each niche. Therefore, Algorithm 5 has been adapted in order to enable it for Bézier solution sets. The adaptations also allow for their use to check and hold all optima instead of only global optima, which will be explained in the paragraphs below.

During optimization, the use of the `Bezier-HillValleyTest` for the clustering of Bézier solution sets resulted in finding sets that were in between niches in later restarts. Figure 4.4 shows the resulting elitist archive if the `Bezier-HillValleyTest` is employed without any changes in order to keep track of elites. The elites denoted in red and purple in the figure are of no interest, whilst

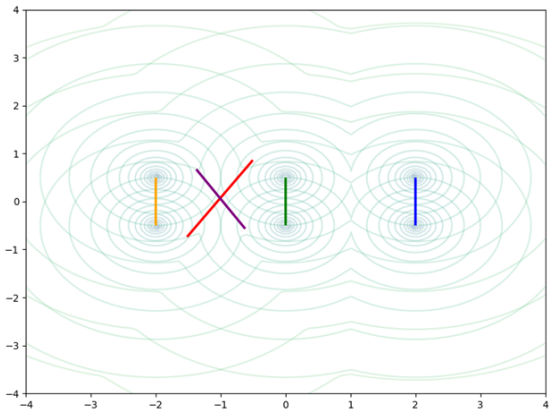


Figure 4.4: Elitist archive of Bézier solution sets after several restarts on the MinDist problem with 3 centres per objective function. The elites denoted in red and purple are of no interest, whilst the others are.

the others that correspond with the optima are of interest. These are of no interest since they do not comply with the niches in which the Pareto sets are located. The red and purple elites actually lie in between two niches. However, due to the requirement in the `Bezier-HillValleyTest` that each i'th control point of two Bézier solution sets is in the same niche, the uninteresting elites are not pruned from the archive.

The `Bezier-HillValleyTest` can be adapted to return true if only one control point is in the same niche instead of all. Since, the two elites that are not desired in Figure 4.4 share at least one control point with an elite that is close to, or on top of, the optima. The `Bezier-HillValleyTest` of Algorithm 12 can be adapted to check if just one control point is in the same niche. This adapted version can be found in Algorithm 13 that denotes the `Bezier-Elite-HillValleyTest`.

Using the new test, a new function can be created to update the elitist archive and hold elites for each niche. Inspiration is taken from HillVallEA [48], where the `HillValleyTest` is used to preserve a solution for each niche in the elitist archive for an SO EA. The new function to update the Elitist archive can be seen in Algorithm 14. If only global Pareto sets are of interest, this algorithm can be adapted further by enabling lines 1 to 6.

Algorithm 13: $[B]$ = `Bezier-Elite-HillValleyTest`($\mathbf{S_i}$, $\mathbf{S_j}$, N_t, f)

Input: Solutions sets $\mathbf{S_i}$, $\mathbf{S_j}$, int N_t, objective functions $f_0, ..., f_{m-1}$
Output : Whether $\mathbf{S_i}$ and $\mathbf{S_j}$ belong to the same niche

```
1  for l = 1, ..., q do
2      c_{i,l} = control point l of S_i
3      c_{j,l} = control point l of S_j
       // Check if c_{i,l} and c_{j,l} are in same niche for all m objectives
4      for k = 0, ..., m − 1 do
           // see Algorithm 3
5          if not HillValleyTest(c_{i,l}, c_{j,l}, N_t, f_k) then
6              return true
7          end
8      end
9  end
10 return false
```

Algorithm 14: $[\mathcal{E}]$ = `AddBezierEliteToArchive`($\boldsymbol{x}$, $\mathcal{E}$) [48]

Input: Candidate Bézier solution set $\boldsymbol{x}$ and elitist archive $\mathcal{E}$
Output : Elitist archive $\mathcal{E} = [\mathcal{E}_0, \mathcal{E}_1, ...]$

if *Only looking for global elites* **then**
  $\boldsymbol{e} \leftarrow$ best elite from $\mathcal{E}$;
  if $(\boldsymbol{x} + TOL) \succ \boldsymbol{e}$ **then**
    $\mathcal{E} \leftarrow \{\}$;
  end
end
$\boldsymbol{e} \leftarrow$ elite closest to $\boldsymbol{x}$;
if *Bezier-Elite-HillValleyTest*($\boldsymbol{x}$, $\boldsymbol{e}$, 5, f) **then**
  if $\boldsymbol{x} \succ \boldsymbol{e}$ **then**
    Remove $\boldsymbol{e}$ from $\mathcal{E}$;
    $\mathcal{E} \leftarrow \mathcal{E} \cup \{\boldsymbol{e}\}$;
  end
else
  $\mathcal{E} \leftarrow \mathcal{E} \cup \{\boldsymbol{e}\}$;
end
return $\mathcal{E}$;

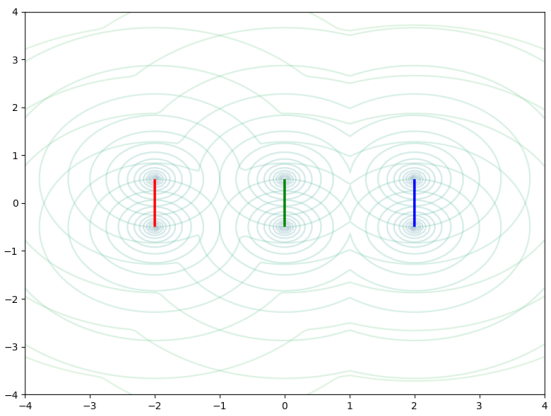


Figure 4.5: Elitist archive of Bézier solution sets after several restarts on the MinDist problem with 3 centres per objective function using new test from Algorithm 14

Answer to Research Question 1.2

Figure 4.5 shows the found elites when using the new `AddBezierEliteToArchive` function denoted in Algorithm 14. This plot clearly shows that the Pareto sets have been approximated correctly and no approximation sets have been stored in the elitist archive that do not correspond to a Pareto set. Thus, it indicates that research question 1.2 has been answered in the form of Algorithm 14.

4.4 Algorithm Overview

MM-BezEA has a similar structure as the restart scheme in HillVallEA [48] that is described in section 2.6.2. Every iteration, the combination of taboo regions and subsequent initialization of Bézier curves takes place. For each of the resulting niches, a core search algorithm is run for one generation, which in the case of MM-BezEA is the RV-GOMEA algorithm [12] that is also used in BezEA.

At the end of each generation, the Bézier HVT of Algorithm 12 is used in the HVC step. This step takes all solutions originating from all clusters and clusters them again for the next generation. The best solution from each cluster is then checked against the elitist archive using the new `AddBezierEliteToArchive` test of Algorithm 14. In between generations, the notion

of cluster registration [8] is used on the cluster mean closest in decision space to transfer the parameters for RV-GOMEA between the clusters of each generation.

If the employed core search algorithm denotes that a cluster has converged, e.g., the variance of the fitness values or decision variable values approaches a small predetermined constant, the cluster is stopped and removed. This happens for all clusters, after which a new generation is run for each of the remaining clusters separately. Once all clusters have been stopped due to convergence, the next iteration is started with twice the population size. Doubling the population size reduces the need for parameter tuning [1], as smaller niches that can include Pareto sets can then be captured in subsequent restarts. During the initialization of a new iteration, the clusters that have an elite as their best solutions are omitted as they have previously been optimized for. This is similar to the idea of the restart scheme employed in HillVallEA [48].

An overview of the algorithm in the form of pseudocode is given in Algorithm 15.

Algorithm 15: $[\mathbb{E}]$ = `MM-BezEA`(f, N, p, q)

Input: MO function f, population size N, number of test points p, number of control points q, evaluation or runtime budget
Output : Elitist archive $\mathbb{E} = [\mathcal{E}_0, \mathcal{E}_1, ...]$

$\mathbb{E} = \{\}$
while *budget remaining* **do**
  $\mathcal{P}_{mo}$ = UniformSampling($q \cdot N$, f)
  $\mathbb{C}_{mo}$ = MO-HillValleyClustering($\mathcal{P}_{mo}$, f)
  $\mathbb{C}$ = InitializeBezierSolutions($\mathbb{C}_{mo}$, q, p, f)
  for $\mathcal{C}_i \in \mathbb{C}$ **do**
    $\boldsymbol{x} \leftarrow$ best from cluster $\mathcal{C}_i$
    $\mathbb{E}$ = AddBezierEliteToArchive($\boldsymbol{x}, \mathbb{E}$)
  end
  ρ = InitModelParameters($\mathbb{C}$)
  `// Run the core_search_algorithm for each cluster until the cluster converges.`
  while $\mathbb{C} \neq \emptyset$ **do**
    $\mathcal{P} = \mathbb{E}$
    for $(\mathcal{C}_i, \rho_i) \in (\mathbb{C}, \rho)$ **do**
      $(\mathcal{O}_i, \rho_i)$ = core_search_algorithm$(\mathcal{C}_i, \rho_i)$
      $\mathcal{P} = \mathcal{P} \cup \mathcal{O}_i$
    end
    $\mathbb{C}_{prev} = \mathbb{C}$
    $\mathbb{C}$ = BezierHillValleyClustering($\mathcal{P}, f$)
    for $\mathcal{C}_i \in \mathbb{C}$ **do**
      $\boldsymbol{x} \leftarrow$ best from cluster $\mathcal{C}_i$
      $\mathbb{E}$ = AddBezierEliteToArchive($\boldsymbol{x}, \mathbb{E}$)
    end
    $\mathbb{C}$ = RemoveElitesFrom($\mathbb{C}$)
    $(\mathbb{C}, \rho)$ = ClusterRegistration$(\mathbb{C}, \mathbb{C}_{prev}, \rho)$
    $\mathbb{C}$ = CheckConvergedClusters($\mathbb{C}$)
  end
  $N \leftarrow 2N$
end

Chapter 5

Set Bézier Evolutionary Algorithm

As MM-BezEA is a niching method that divides its attention over all of the niches found by HVC, it is intractable to use MM-BezEA in highly multi-modal landscapes. In order to overcome this limitation, a second algorithm is introduced that uses MO-RV-GOMEA [13] with an adapted fitness function definition and encoding. Rather than encoding each solution to be one Bézier curve, and clustering the population to find different niches, this algorithm, named Set-BezEA, encodes and searches for a set of b Bézier curves. It does so without further niching the population, which prevents this approach from falling apart in highly multi-modal landscapes.

In the original BezEA, solutions represent Bézier curves that consist of the q control points, which are all paremeterized as regular solutions, concatenated into the set of control points $C_q = \{c_1, ..., c_q\}$. Given a problem dimensionality ℓ, a Bézier solution thus consists of $q \cdot \ell$ parameters, also denoted as $\phi \in R^{q \cdot \ell}$. In Set-BezEA, a solution represents b Bézier curves that consists of b sets of q control points C_q concatenated into a set of control points denoted as $C_{q,b}$, totalling $q \cdot b$ control points. The control points for the i-th Bézier solution are denoted as C_q^i. Hence, solutions in Set-BezEA take the form $[C_q^1, ..., C_q^b] \in \mathbb{R}^{b \cdot q \cdot \ell}$, also denoted as $\phi \in \mathbb{R}^{b \cdot q \cdot \ell}$. The notation $S_{p,q}^i$ is used to denote the i-th Bézier solution of the set $S_{p,q,b}$ with b Bézier solutions. The definition of $S_{p,q}^i$ is equal to the definition of $S_{p,q}$ in Equation 11.

5.1 Fitness Function Formulation

In order to evaluate the newly formulated solutions, a new fitness function is proposed that can be seen in mathematical form in Equation 13. The original fitness function formulation from BezEA is used in order to determine the constraint $L(S_{p,q,b}(\phi))$ (from Algorithm 11) and $\text{HV}_{f,S_{p,q,b}(\phi)}$ values for all of the b Bézier solutions combined. Fitness sharing [31] is incorporated with a niching radius based on the expected edge length [48] (EEL), which is checked between each of the b Bézier solutions $S_{p,q}^i$. If any of these solution sets are within the EEL of each other, they are to share their HV values amongst all solutions that are close to one another by dividing their fitness values by the number of solutions nearby.

However, fitness sharing only describes a minimal separation criterion. Typically, in practice, decision makers would like to be able to find Bézier solutions that are diverse, yet all of high quality, while limited in number. For this reason, the problem is defined as an MO formulation. Herein, the second objective is a distance objective $\Delta_{S_{p,q,b}}$ that is to be maximized. It calculates the sum of cosine distances between all pairs of the b Bézier solutions to move them apart in decision space and promote finding diverse approximation sets.

Definition 13. *The set of Bézier curves optimization problem is defined by:*

$$\begin{aligned}
&\textit{maximize } f_0\textit{:} \quad HV_{f,S_{p,q,b}(\phi)} = \sum_{i=0}^{b-1} \frac{HV_f\left(A^{nb}\left(S^i_{p,q}(\phi)\right)\right)}{n_i} \\
&\textit{maximize } f_1\textit{:} \quad \Delta_{S_{p,q,b}(\phi)} = \sum_{i=0}^{b-1}\sum_{j=i+1}^{b-1} \frac{S^i_{p,q}(\phi)\cdot S^j_{p,q}(\phi)}{||S^i_{p,q}(\phi)||\,||S^j_{p,q}(\phi)||} \\
&\textit{with constraint: } L\left(S_{p,q,b}(\phi)\right) = \sum_{i=0}^{b-1} L_b\left(S^i_{p,q}(\phi), o^i_{nb}(\phi)\right)
\end{aligned}$$

where $f : \mathbb{R}^\ell \to \mathbb{R}^m, \phi \in \mathbb{R}^{b*q*\ell}$,

A^{nb} : *solution set* $S^i_{p,q}(\phi) \to$ *approximation set* $\mathcal{A}^i_{p,q,o_{nb}}$

n_i *number of Bézier curves in same niche as* $S^i_{p,q}$

$o^i_{nb}(\phi)$ *the navigational order of the i-th Bézier solution of Algorithm 10*

$L_b(S^i_{p,q}(\phi), o^i_{nb}(\phi))$ *the Bézier constraint function of Algorithm 11*

5.2 Adaptive Steering

The MO formulation of the previous section creates a trade-off between finding far apart solutions at the edges of the search space and those closer to the PSs. However, especially assuming the global optimum cannot be obtained, the primary goal is to obtain distinct solutions that are of high quality. Most (M)MOEAs try to find improvements spread over the entire PF, whereas in some situations only a part of the PF is interesting. In order to steer Set-BezEA towards the interesting parts of the PF, the notion of adaptive steering [2] can be used.

The constraining of solutions has a delayed start and is done in an iterative fashion. If solutions are constrained right from the start, not enough solutions may have been found to produce a well spread approximation front. Thus, adaptive steering has a delayed start to make sure premature convergence is prevented. The iterative nature ensures that after an improvement has been found, the constraints are updated to keep pushing the approximation front towards the interested region.

Since low-quality and far apart solution sets are not of interest, adaptive steering is applied on the first objective value. This is to enforce high-quality approximation sets, at the cost of possibly reduced dissimilarity. Starting at allowing Bézier solution sets $S_{p,q,b}$ to be within 50% of the best achieved cumulative HV it scales linearly in 50 generations to 99%. This will force all Bézier solution sets $S_{p,q,b}$ towards higher HV values when combined with constraint domination [19].

5.3 Linkage Structure and Partial Evaluations

To optimize for the new fitness formulation, Set-BezEA uses MO-RV-GOMEA, which uses a linkage model [63] to determine which variables are to be changed at the same time. It is often taken to be a linkage tree that is constructed using UPGMA [55] to iteratively cluster decision variables. A new linkage structure is designed for Set-BezEA to guide MO-RV-GOMEA and allow it to exploit the underlying structure of a problem. For this new linkage structure, the ideas behind the linkage structures defined for BezEA [72] and UHV-GOMEA [69] were combined.

When the j-th decision variable of a control point in Bézier curve changes, the j-th decision variable of all test solutions of the approximation set $S^i_{p,q}(C_q)$ changes, which makes it beneficial to change the same decision variable for all q control points of a Bézier curve at once. This is the first layer of the new linkage structure, where, for each of the b Bézier solutions, their j-th decision variables of all q control points are grouped. Subsequently, the ℓ clusters of q decision variables for one Bézier solution are hierarchically grouped using UPGMA until all decision variables of one Bézier solution are clustered. The second layer combines the b resulting groups of the decision variables of all b Bézier solutions by concatenating them into one large linkage tree. Figure 5.1 gives an overview of how the linkage tree is structured and how the decision variables are grouped in the different layers. This linkage structure would for example be used if Set-BezEA optimizes two approximation sets for the MinDist problem.

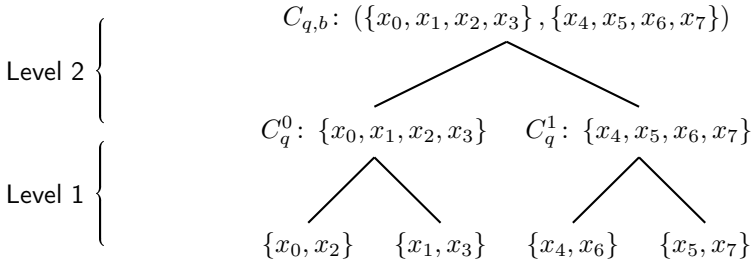


Figure 5.1: Set-BezEA linkage structure ($b = q = \ell = 2$)

The newly defined linkage structure enables MO-RV-GOMEA to perform partial evaluations in all situations. Even if the underlying problem setting is black-box in nature, partial evaluations are enabled by the second layer, as each of these Bézier curves can be evaluated separately for its HV contribution to the first objective. The second distance objective always has to be recalculated, but it does not require any evaluations of the original fitness function.

The use of the new linkage tree had a profound impact on the ability of Set-BezEA to find optima. With the new linkage tree Set-BezEA could find the distinct optima on the MinDist problem in less than 10 generations. Without the information embedded in this structure, it took up to hundreds of generations when random variables were changed. Figure 5.2 shows the result of Set-BezEA on the MinDist problem when using the new LT.

Answer to Research Question 2.1

Figure 5.2 indicates that Set-BezEA has successfully approximated the two Pareto sets that are furthest away from each other in decision space in the MinDist problem. This indicates that Set-BezEA is not only capable of obtaining distinct, but also high-quality approximation sets. Therefore, research question 2.1 has been succesfully answered in the form of Set-BezEA that uses the newly proposed fitness function formulation.

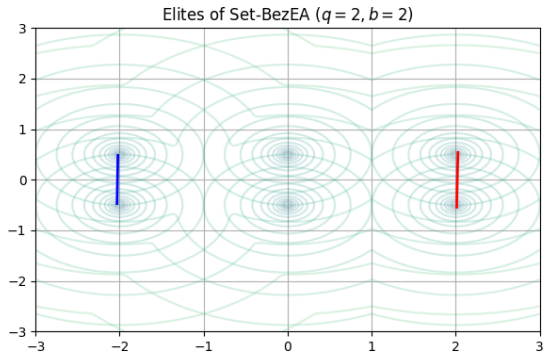


Figure 5.2: Best approximation sets after run of Set-BezEA on MinDist problem.

5.4 MO-RV-GOMEA Adaptations

The approximation sets $S^i_{p,q}(C_q)$, defined by the Bézier curves, are evaluated with the HV indicator that makes use of a reference point. The control points $\boldsymbol{c}^i_1$ and $\boldsymbol{c}^i_q$ are likely to come as close as possible to the constraint region, defined by being dominated by the reference point, to maximize the HV. However, if one of these two control points falls just slightly in the constraint region, the constraint function $C(S_{p,q,b}(\phi))$ combined with constraint domination will make it one of the worst performing solutions. Furthermore, MO-RV-GOMEA makes use of the Gene-pool Optimal Mixing (GOM) variation operator, thereby disallowing deterioration of fitness. These two effects combined, quickly reduce the variance in the population, since very slight adjustments can significantly deteriorate the individual.

The reduced variance as a result of the aforementioned effects was most profound in higher dimensional and more difficult optimization problems like HDR brachytherapy. In order to prevent

the reduced variation in the population, the GOM part of MO-RV-GOMEA has been adapted to allow deterioration of the objective values, i.e., accepting a modification that would otherwise have been rejected, with a probability of 5%. This is similar as to the adaptation of the GOM operator as previously done in RV-GOMEA [12]. The adaptation enables MO-RV-GOMEA to effectively search the outer regions of the search space close to the constraint space by preserving more variance.

Another adaptation to MO-RV-GOMEA is that of the disablement of the clustering step which is used to improve certain parts of the approximation front. Since the control points of a Bézier solution set model the approximation set, the test solutions themselves are no longer sampled by MO-RV-GOMEA. Therefore, it is of no use to cluster various parts of the approximation front. Furthermore, in the case of $q = 2$ control points, only the decision variables of the control points at the extremes of the approximation front are sampled. This further nullifies the clustering step as MO-RV-GOMEA cannot sample any decision variables in the middle of the approximation front.

Chapter 6

Experiments

MM-BezEA and Set-BezEA are empirically benchmarked on several test problems in order to examine their strengths and weaknessses. The results are compared to MO-HillVallEA [47], MO-RV-GOMEA [13], and MO_Ring_PSO_SCD [74]. MO_Ring_PSO_SCD is implemented through the PlatEMO framework [65] in Matlab, where all of the used metrics and problems are manually implemented as well. For the other algorithms, original C++ implementations are used.

6.1 Test Problems

Several test problems are employed on which the (M)MOEAs are benchmarked. First of these is the MinDist problem [47] that was described in Section 2.1.3 of the introduction. This problem has linear PSs that are to be found. The other employed functions are frequently used in literature, namely OmniTest [20], Two on One [58], and Sympart {1,2,3} [61]. Lastly, several problems are taken from the Multi-modal Multi-objective test Function (MMF) benchmark suite [39] in the form of MMF {1,2,12,14,15}, of which the first two are relabeled versions of SSUF {1,3} [38].

A mix of PS and PF shapes have been chosen to determine the capabilities of the two new algorithms on different problem types. Table 6.1 shows some of the important characteristics for each of the problems. For all problems with a configurable number of PSs, it is set to 2, likewise the problem dimension ℓ is fixed to 2. In order to determine the values of the performance indicators, the reference PSs will be made using 5000 points that adhere to the analytical formulas describing the PSs. In the case of Two on One, a very close approximation is used, as no analytical formula is available [58]. The objective functions and Pareto set definitions of the problems defined in Table 6.1 can be found in Appendix A.

Table 6.1: Bi-Objective Problem Instances and Characteristics.

Problem	ℓ	#PS	PS Shape	PF Shape
MinDist	$[2, \ldots, \infty) \in \mathbb{Z}$	n	Linear	Convex
Omni Test	$[2, \ldots, \infty) \in \mathbb{Z}$	3^ℓ	Linear	Convex
Two on One	2	2	Linear	Convex
Sympart 1 & 2	2	9	Linear	Convex
Sympart 3	2	9	Non-linear	Convex
MMF 1 & 2	2	2	Non-linear	Convex
MMF 12	$[2, \ldots, \infty) \in \mathbb{Z}$	n	Linear	Disconnected
MMF 14 & 15	$[2, \ldots, \infty) \in \mathbb{Z}$	n	Linear	Concave

6.2 Experimental Setup

In order to get a fair comparison, each of the algorithms will be given an equally sized budget of 200,000 function evaluations for each of the problems. This removes the influence of the used programming languages, i.e., Matlab vs C++, as the computation time is not limited. The parameters of MO-RV-GOMEA, MO_Ring_PSO_SCD, and MO-HillVallEA are set to the values reported in relevant literature. Furthermore, for each problem and metric, the average and standard deviation over 31 runs will be taken to obtain a fair comparison.

The elitist archives sizes $N_{\mathbb{E}}$ are set to be 1250 for MO-RV-GOMEA and MO-HillVallEA. The population size N is set to 96 for MO-RV-GOMEA and 250 for MO-HillVallEA. MO-RV-GOMEA uses a linkage tree as its linkage model, with a total of 5 clusters. For MO_Ring_PSO_SCD the population size is 800 [74]. The C_1, C_2, and W parameters, which are used for the velocity update of the particle swarm optimizer, are set to 2.05, 2.05, and 0.7289 respectively. For the MM-BezEA and Set-BezEA algorithms, the number of control points q for each approximation set is set to 2 so that they are searching for linear approximation sets. Just like for the original BezEA algorithm, these two algorithms are both given population sizes of 76 [72]. The number of test points p is set to 7. Lastly, Set-BezEA is always configured to produce $b = 2$ Bézier solutions as approximation sets. Otherwise, the parameter b has to be reduced for certain problems since some only have two PSs.

6.3 Performance Indicators

Hypervolume. Various performance indicators are used to depict the characteristics of the newly proposed algorithms and compare them to the established algorithms. The HV indicator [78] is used to see how well the algorithms perform in terms of getting close the the PF. As a result of the use of test points in the MM-BezEA and Set-BezEA algorithms, these algorithms have a limited amount of points in the approximation set that can be used to calculate the HV values. Therefore, a subset of the approximation set will be taken for MO-RV-GOMEA, MO-HillVallEA and MO_Ring_PSO_SCD in order to allow for a fair comparison based on the HV indicator. Specifically, the same number of test points, i.e., 7, is selected for a fair comparison by means of greedy Hypervolume Subset Selection (gHSS) [28].

Pareto Set Proximity. A relatively new performance indicator for multi-modal multi objective optimization is also employed, named Pareto Set Proximity (PSP) [74]. It is an indicator that determines how well all PSs are approximated by taking the Cover Rate (CR), that shows how well the extremes of all PSs are captured by the approximation set $\mathcal{S}$, divided by the Inverted Generational Distance in decision space (IGDX) [75]. The IGDX can be used to determine how close the approximation sets $\mathcal{S}$ are to the PSs P^*. For the IGDX measure, the approximation sets as produced by MM-BezEA and Set-BezEA are interpolated by taking 1000 intermediate points before determining the IGDX value. This can be performed relatively easily as interpolating these parameterizations does not require any extra fitness evaluations.

$$PSP = \frac{CR}{IGDX} \tag{6.1}$$

$$IGDX(P^*, \mathcal{S}) = \frac{\sum_{\boldsymbol{v} \in P^*} \min\left(||\boldsymbol{x} - \boldsymbol{v}|| \text{ for } \boldsymbol{x} \in \mathcal{S}\right)}{|P^*|} CR \qquad = \left(\prod_{i=1}^{n} \delta_i\right)$$

$$\delta_i = \begin{cases} 1 & V_i^{max} = V_i^{min} \\ 0 & v_i^{min} \geq V_i^{max} || v_i^{max} \leq V_i^{min} \\ \left(\dfrac{\min\left(v_i^{max}, V_i^{max}\right) - \max\left(\left(v_i^{min}, V_i^{min}\right)\right)}{V_i^{max} - V_i^{min}}\right) & \text{otherwise} \end{cases}$$

Where V_i^{min} and V_i^{max} are the minimum and maximum value for the $i'th$ decision variable of the PS P^* and v_i^{min} and v_i^{max} are the minimum and maximum value for the $i'th$ decision variable of the approximation set $\mathcal{S}$.

Smoothness. The last employed performance indicator is that of smoothness, for which the definition as introduced in the work on BezEA is followed [72], which was introduced by the same authors that worked on the Bézier parameterizations for approximation sets. It captures how smooth an approximation set can be navigated in terms of decision variables values. The smoothness is calculated by measuring the detour length in decision space when traversing the approximation set from solution $\boldsymbol{x}_{i-1}$ to solution $\boldsymbol{x}_{i+1}$ via an intermediate solution $\boldsymbol{x}_i$, as compared to directly going from $\boldsymbol{x}_{i-1}$ to $\boldsymbol{x}_{i+1}$. The smoothness approaches its maximum value of 1 if all solutions would be colinear in decision space, where the lowest possible value is 0. In cases where the approximation set is clustered, like in MO-HillVallEA, MM-BezEA and Set-BezEA, the average smoothness over all clusters will be taken. In the other cases the smoothness over the entire approximation set $\mathcal{S}$ is taken. $\mathcal{S}$ is sorted on increasing objective values for f_1, with o denoting the navigational order when sorted on f_1. For Set-BezEA and MM-BezEA o is equal to the navigational order o_{nb}.

$$Sm(\mathcal{S}, o) = \frac{1}{|\mathcal{S}| - 2} \sum_{i=2}^{|\mathcal{S}|-1} \frac{||\boldsymbol{x}_{o_{i-1}} - \boldsymbol{x}_{o_{i+1}}||}{||\boldsymbol{x}_{o_{i-1}} - \boldsymbol{x}_{o_i}|| + ||\boldsymbol{x}_{o_i} - \boldsymbol{x}_{o_{i+1}}||} \tag{6.2}$$

6.4 Results

The results for all problems and algorithms can be seen in Tables 6.2, 6.3, and 6.4. These tables show the HV, PSP, and Smoothness indicators respectively. The results are in scientific notation showing the mean and standard deviation within brackets taken over 31 runs. In bold is the best performing method per problem for a specific indicator if it has a significant difference to other methods in a pairwise comparison according to the Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction.

Hypervolume The HV results clearly show that all algorithms are capable of performing nearly equally in obtaining a good approximation front. However, The MM-BezEA and Set-BezEA algorithms do deteriorate in performance on the MMF1 and 2 problems that have non-linear PSs. The deterioration is inherently caused by the chosen parameterizations that create approximation sets which are linear in shape. Another problem instance where a smaller HV for the new algorithms is obtained, is that of MMF12. Here, despite MM-BezEA obtaining the best PSP values, the approximation sets did not fully approximate the actual PSs and did not cover the endpoints.

Pareto Set Proximity The PSP indicator shows similar results, except that MO-RV-GOMEA performs worse as it is not an MMOEA and therefore does not explicitly search for multiple niches. Again promising results for MM-BezEA and Set-BezEA are shown in cases where linear PSs can be found, with MM-BezEA obtaining better or equal performance to all other algorithms. In the problems with non-linear PSs, MO-HillVallEA and MO_Ring_PSO_SCD manage to find better approximations.

Smoothness The smoothness results show, as intended, that the chosen parameterizations inherently cause smooth approximation sets. Therefore, MM-BezEA and Set-BezEA always obtain a perfect smoothness of 1.0. Other algorithms do not obtain this, except for MO-RV-GOMEA on 2 of the 11 problems.

A visualization of the results of MM-BezEA on the MinDist problem is given in Figure 6.1, and if Set-BezEA would be plotted, the figure would show the same results. This figure depicts the smooth course of the decision variables in the parallel coordinates plot for the rightmost approximation set in decision space. It contrasts sharply to the parallel coordinates plot of Figure 1.3. Thus, indicating that both Set-BezEA and MM-BezEA obtained the preferred results of smoothly navigable approximation sets.

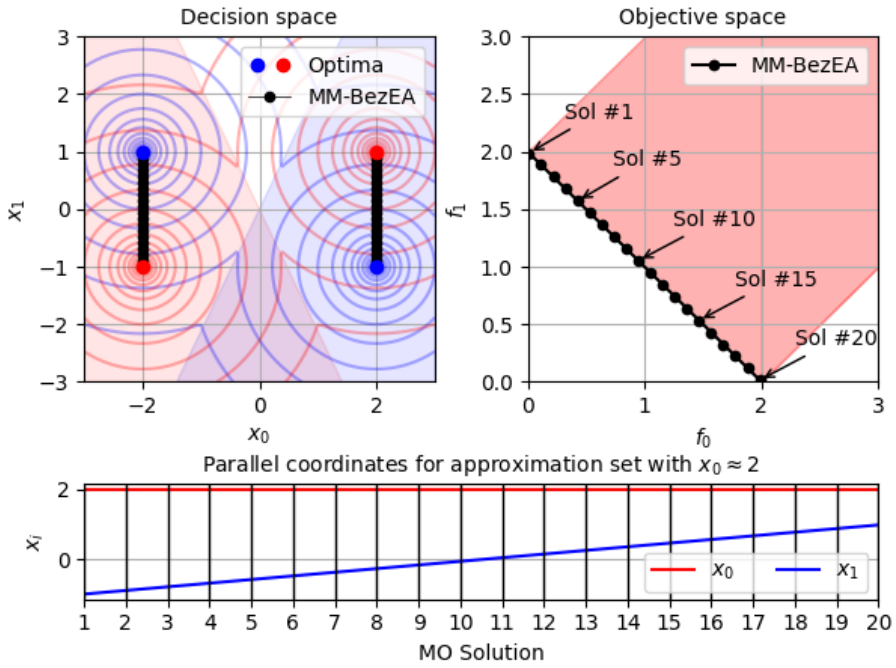


Figure 6.1: Approximation sets and front with parallel coordinates plot for one of the approximation sets as produced by MM-BezEA on MinDist problem [47].

Table 6.2: Hypervolume results (avg. (± st.dev.)) per problem and algorithm over 31 runs, bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction).

Problem	MM-BezEA	Set-BezEA	MO-HillVallEA	MO_Ring_PSO_SCD	MO-RV-GOMEA
MinDist	**1.17e+2 (±9.43e-5)**	**1.17e+2 (±3.02e-4)**	1.17e+2 (±1.62e-2)	1.17e+2 (±5.95e-3)	1.17e+2 (±1.36e-3)
OmniTest	**8.48e+0 (±2.18e-6)**	8.43e+0 (±2.34e-1)	8.47e+0 (±1.96e-3)	8.47e+0 (±7.34e-4)	8.47e+0 (±4.82e-4)
Sympart 1	**1.17e+2 (±2.16e-5)**	1.16e+2 (±1.67e+0)	1.17e+2 (±1.42e-2)	1.17e+2 (±7.64e-3)	1.17e+2 (±1.30e-3)
Sympart 2	**1.17e+2 (±2.91e-5)**	1.15e+2 (±1.97e+0)	1.17e+2 (±7.77e-3)	1.17e+2 (±8.75e-3)	1.17e+2 (±4.47e-3)
Sympart 3	**1.17e+2 (±9.61e-5)**	1.16e+2 (±1.76e+0)	1.17e+2 (±1.61e-2)	1.17e+2 (±9.21e-3)	1.17e+2 (±4.91e-3)
TwoOnOne	1.13e+2 (±1.33e-4)	1.13e+2 (±9.27e-4)	1.13e+2 (±2.39e-4)	**1.13e+2 (±1.82e-4)**	1.13e+2 (±1.10e-4)
MMF 1	6.04e-1 (±3.86e-2)	6.29e-1 (±5.63e-2)	**8.05e-1 (±2.37e-4)**	**8.05e-1 (±8.69e-5)**	**8.05e-1 (±6.60e-5)**
MMF 2	6.34e-1 (±2.01e-4)	5.88e-1 (±9.42e-2)	8.04e-1 (±6.90e-4)	8.04e-1 (±9.59e-4)	**8.05e-1 (±1.75e-4)**
MMF 12	1.78e+0 (±2.02e-6)	1.38e+0 (±2.02e-1)	2.06e+0 (±2.57e-3)	2.06e+0 (±2.05e-3)	**2.06e+0 (±1.49e-4)**
MMF 14	5.63e+0 (±1.33e-5)	5.63e+0 (±6.40e-3)	**5.63e+0 (±7.26e-4)**	5.63e+0 (±1.92e-3)	5.63e+0 (±2.23e-4)
MMF 15	5.56e+0 (±2.03e-2)	5.56e+0 (±4.47e-3)	**5.57e+0 (±6.52e-4)**	5.56e+0 (±1.54e-3)	**5.57e+0 (±1.79e-4)**

Table 6.3: Pareto set proximity results (avg. (± st.dev.)) per problem and algorithm over 31 runs, bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction).

Problem	MM-BezEA	Set-BezEA	MO-HillVallEA	MO_Ring_PSO_SCD	MO-RV-GOMEA
MinDist	**3.26e+2 (±7.31e+1)**	**3.77e+2 (±9.41e+1)**	5.02e+1 (±2.74e+0)	6.97e+1 (±6.73e+0)	1.21e+0 (±1.65e+0)
OmniTest	**2.36e+2 (±8.81e+1)**	1.71e-2 (±3.34e-2)	7.13e+1 (±2.76e+0)	6.90e+1 (±9.56e+0)	1.36e-1 (±2.43e-1)
Sympart 1	**2.67e+2 (±1.21e+2)**	1.22e-3 (±2.89e-3)	3.56e+1 (±1.57e+0)	2.79e+1 (±3.69e+0)	1.16e-2 (±2.70e-2)
Sympart 2	**3.09e+2 (±8.46e+1)**	2.76e-4 (±3.54e-4)	3.60e+1 (±7.42e-1)	2.38e+1 (±2.46e+0)	1.02e-2 (±1.71e-2)
Sympart 3	6.41e+1 (±7.77e+1)	1.27e-3 (±2.17e-3)	**4.33e+1 (±2.10e+0)**	2.64e+1 (±5.62e+0)	8.09e-3 (±1.33e-2)
TwoOnOne	**3.04e+2 (±2.18e+2)**	**2.69e+2 (±1.15e+2)**	4.50e+1 (±7.21e-1)	2.45e+1 (±1.03e+1)	2.68e+0 (±1.11e+0)
MMF 1	7.22e+0 (±2.41e+0)	4.56e-1 (±3.55e-1)	3.17e+1 (±6.84e-1)	**3.80e+1 (±6.79e+0)**	1.02e+0 (±2.89e-1)
MMF 2	4.00e+0 (±2.49e+0)	4.29e-2 (±5.23e-2)	**1.17e+2 (±1.21e+1)**	5.04e+1 (±1.51e+1)	2.18e+0 (±1.41e+0)
MMF 12	**2.42e+1 (±7.72e+0)**	4.83e-1 (±6.81e-1)	1.94e+1 (±6.46e+0)	1.53e+1 (±1.46e-1)	8.67e+0 (±1.36e+0)
MMF 14	**2.68e+3 (±4.82e+2)**	1.72e+3 (±1.60e+3)	3.70e+2 (±1.03e+1)	2.31e+2 (±2.16e+1)	1.08e+0 (±2.84e+0)
MMF 15	**2.73e+2 (±4.17e+1)**	2.96e+2 (±1.41e+2)	2.65e+2 (±4.58e+0)	2.44e+2 (±7.25e+0)	2.24e+1 (±1.40e-2)

Table 6.4: Smoothness results (avg. (± st.dev.)) per problem and algorithm over 31 runs, bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction).

Problem	MM-BezEA	Set-BezEA	MO-HillVallEA	MO_Ring_PSO_SCD	MO-RV-GOMEA
MinDist	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	8.09e-1 (±3.70e-2)	7.63e-1 (±5.91e-3)	8.94e-1 (±1.81e-1)
OmniTest	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	9.23e-1 (±5.65e-3)	7.28e-1 (±9.76e-3)	7.16e-1 (±1.96e-1)
Sympart 1	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	8.76e-1 (±2.52e-2)	6.84e-1 (±9.76e-3)	7.01e-1 (±2.09e-1)
Sympart 2	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	8.78e-1 (±2.14e-2)	5.73e-1 (±9.30e-3)	7.87e-1 (±1.68e-1)
Sympart 3	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	8.70e-1 (±3.10e-2)	5.29e-1 (±1.18e-2)	8.30e-1 (±1.92e-1)
TwoOnOne	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	7.77e-1 (±1.21e-2)	7.47e-1 (±9.36e-3)	7.51e-1 (±1.59e-1)
MMF 1	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	9.01e-1 (±3.98e-2)	7.82e-1 (±8.49e-3)	5.86e-1 (±9.13e-2)
MMF 2	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	9.38e-1 (±1.26e-2)	5.01e-1 (±8.92e-3)	8.46e-1 (±1.60e-1)
MMF 12	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	8.35e-1 (±2.73e-2)	6.35e-1 (±9.86e-3)	**1.00e+0 (±0.00e+0)**
MMF 14	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	9.31e-1 (±1.17e-2)	8.71e-1 (±1.11e-2)	9.62e-1 (±1.05e-1)
MMF 15	**1.00e+0 (±0.00e+0)**	**1.00e+0 (±0.00e+0)**	9.19e-1 (±1.43e-2)	8.63e-1 (±8.54e-3)	**1.00e+0 (±0.00e+0)**

6.5 Scalability

Set-BezEA has been created to overcome scalability issues with MM-BezEA in highly multi-modal problems. To evaluate this, the OmniTest problem is used, that is scalable in its problem dimensions whilst simultaneously scaling the number of PSs. MM-BezEA and MO-RV-GOMEA will be compared to Set-BezEA with $b \in 2, 3, 4$, to show the impact of requesting more approximation sets. Problem dimensions $\ell \in \{2, 3, 5, 10, 20, 30, 50, 100\}$ have been considered to see how the algorithms behave. The evaluations budget is set to $100000 * \ell$, with the rest of the parameters equal to before.

The PSP metric is omitted, as OmniTest scales exponentially in the number of PSs if the problem dimensionality is increased. In case $\ell \geq 8$, there are over 5000 PSs ($3^8 = 6561$), whereas only 5000 points of the PSs are analytically determined for the indicators. Thus, only the smoothness and HV indicators are used as these do not depend on the number of solutions drawn per PS. Furthermore, as MM-BezEA and Set-BezEA produce approximation sets that are each within a single niche, the approximation set of MO-RV-GOMEA will be post-processed with MO-HVC to compare the average HV per niche.

Figures 6.2 and 6.3 respectively show the the HV and smoothness indicators results for the dimensionality comparison. The results show that on average MO-RV-GOMEA seems to obtain reasonable approximation sets within each niche. However, many of these niches have only one or a few solutions, which is also depicted by the lower smoothness values of the entire approximation set on the right. Furthermore, the worst performance in terms of HV per niche can be nearly two orders of magnitude lower than Set-BezEA when $\ell = 100$. MM-BezEA is, like MO-RV-GOMEA, capable of finding a good approximation set for at least one niche as seen by the fact that the best observed performance of all algorithms coincide, but MM-BezEA also has the worst observed performance. For this specific problem, it is caused by the fact that there are so many niches (3^ℓ), that the algorithm cannot keep track of all of them.

In the case of Set-BezEA, the results indicate that it is capable of finding decent and distinct approximation sets. When more approximation sets are optimized for, the average HV can decline slightly. However, even HV per niche remains high in comparison to the other algorithms, whilst the smoothness stays at its maximum of 1, showing that Set-BezEA exhibits the intended behaviour.

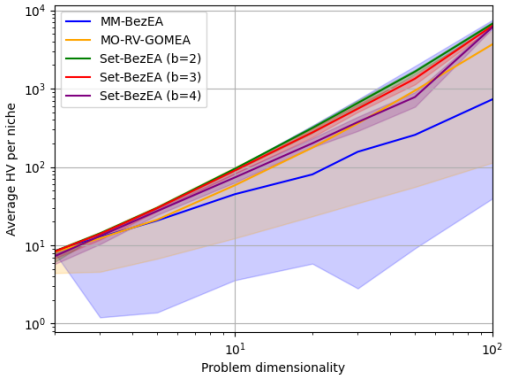


Figure 6.2: Hypervolume results of dimensionality comparison. The averages are depicted as a line and the min. and max. values as shaded areas over 31 runs.

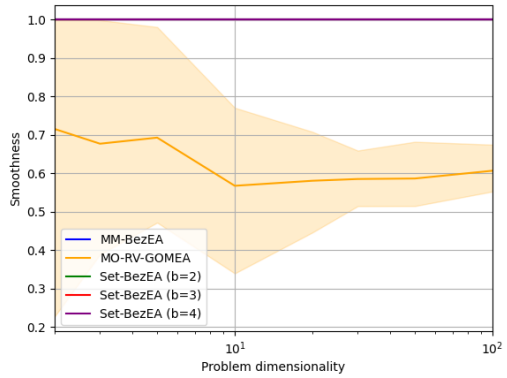


Figure 6.3: Approximation set smoothness results of dimensionality comparison. The averages are depicted as a line and the min. and max. values as shaded areas over 31 runs. MM-BezEA and Set-BezEA smoothness results overlap.

6.6 Obtaining any distinct or all Pareto sets

Here the behaviour of MM-BezEA is further investigated on the SymPart 1 problem. This problem has 9 Pareto sets. Therefore, the behaviour of the algorithm, including restarts whereby other parts of the search space are focused on, can be investigated rather well.

Figure 6.4 shows a generation from each of three (re-)starts of MM-BezEA. The first run will have a lower population size than the subsequent runs and has complete freedom as to which mode is focused on. This results in a few clusters that have to be merged in order to obtain a large enough population over which a distribution can reliably be estimated. Therefore, the fifth generation is shown for the first run of the algorithm, as MM-BezEA is still trying to find its way into the two modes shown in the figure in the first generation.

When looking at the first and second restarts in Figure 6.4, they clearly show that other parts of the search space are focused on in the first generation of these restarts. When the left most and middle plot of Figure 6.5 are overlayed on top of the first and second restart, i.e. middle and right most plot of Figure 6.4, they indicate that the discarding of already optimized niches works as expected. Thus, previously optimized niches in both the first start and the first restart are not optimized for in the final right most plot. This is exactly the intended behaviour in order to find an approximation set for each of the nine Pareto sets.

Figure 6.5 shows the elites at the end of each (re-)start of MM-BezEA. The right most plot indicates that the algorithm has successfully found an approximation set for all 9 of the PSs. This is furthermore emphasized by Table 6.3, which shows that MM-BezEA obtained the highest PSP value of all tested algorithms. Therefore, it can be concluded that MM-BezEA is indeed able to

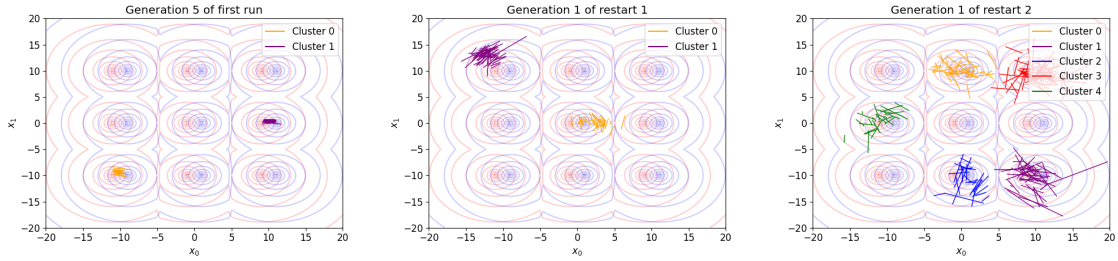


Figure 6.4: Plots showing the population at specific generation of several (re-)starts of the MM-BezEA algorithm. Taken together, the plots indicate that the taboo regions of the restart scheme seem to work as intended on the SymPart 1 problem.

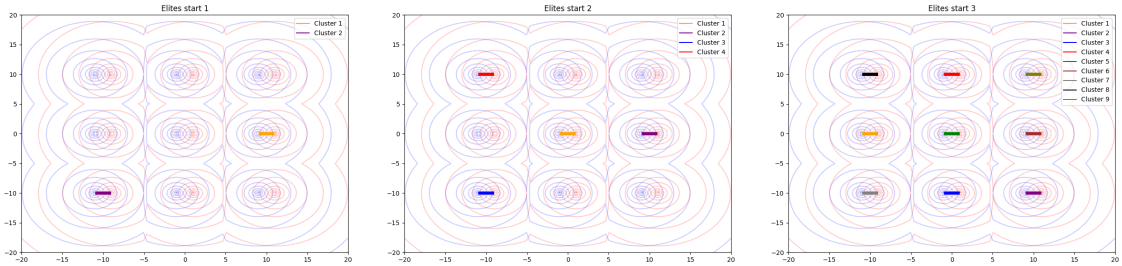


Figure 6.5: Three plots showing the elitist archive of the MM-BezEA algorithm on the SymPart 1 problem. From left to right, each plot shows the elites after a new start of the algorithm.

obtain approximation sets for all Pareto sets in the case of linear Pareto sets.

Figure 6.6 shows the elites found by the Set-BezEA algorithm on the Sympart 1 problem. It shows that two distinct approximation sets have been found that are dissimilar, even though they are not the most far apart as possible. This can probably be ascribed to the use of the cosine similarity in the second objective of Set-BezEA instead of, for example, using the Euclidean distance. However, Table 6.2 shows that in all cases with linear PSs, Set-BezEA is capable of obtaining high quality approximation sets.

Answer of Research Questions 1 and 2

To conclude, research question 1 has been answered in the form of the MM-BezEA algorithm that is capable of finding all smoothly navigable approximation sets, even though it is only in the case of linear Pareto sets. In cases of non-linear Pareto sets, Table 6.3 shows that MM-BezEA is currently not capable of approximating all Pareto sets. Thus, further research would be required for non-linear Pareto sets as MM-BezEA is then not an answer to research question 1. A similar conclusion can be drawn from the results for research question 2. Set-BezEA is found to be capable of finding multiple high quality approximation sets that are dissimilar in the case of linear Pareto sets, although they are not as far apart as possible.

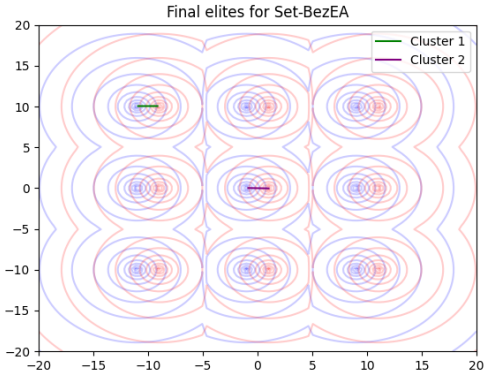


Figure 6.6: Plot showing the obtained elites for Set-BezEA on the SymPart 1 problem.

Chapter 7

Application to High-Dose-Rate Brachytherapy

The purpose of the algorithms researched in this thesis was for them to be applicable to HDR brachytherapy. As stated in the introduction, currently a non-intuitive course of the dwell times is found when navigating the approximation front from plans with good coverage to those with good sparing. The non-intuitive change can be depicted as drastic changes in the dose distributions when considering two adjacent treatment plans on the approximation front, similar to the drastic changes in Figure 1.3 of the introduction. The use of Bézier curve parameterizations is a solution to these drastic changes as adjacent treatment plans on the approximation front now vary smoothly in dwell times. Figure 6.1 gave a good depiction of how the decision variable values change smoothly when navigating the approximation front.

In this chapter Set-BezEA will be evaluated on the HDR brachytherapy problem. It will be compared to BRIGHT that is currently used in the clinical setting, which is defined as MO-RV-GOMEA applied to a specific bi-objective problem formulation. The purpose of this comparison is to obtain answers to research question 3 and its sub questions 3.1 and 3.2. MM-BezEA will be omitted from the experiments conducted in this chapter as Figure 6.2 of the results section clearly shows the reduced quality of approximation fronts when optimizing for highly multi modal problems.

First the empirical benchmark setup, indicators and results will be discussed. The last part of this chapter will dive into a clinical evaluation of a few treatment plans by clinicians in order to determine the possible usability of Set-BezEA in the clinic. It is also of paramount importance to obtain feedback on the optimized treatment plans from medical experts as they have the knowledge on what is truly acceptable to be used in the clinical practice.

7.1 Experimental Setup

The experimental setup will be chosen to mimic the setting used in clinical practice. This entails that the problem formulation is chosen based on the new protocol used since 2020 as described in Table 2.2. The objectives are the weighted and normalized LSI_n and LCI_n objectives as described in Equation 2.9.

A total of 7 anonymized patients are used in the comparison. All of them have previously been treated at the Amsterdam UMC hospital before 2018 using the protocol of Table 2.1 that was employed until 2020 and the microSelectron-HDRv2 Ir-192 source. The specific information per patient in terms of number of catheters, total number of dwell positions and range of dwell positions per catheter can be found in Table 7.1. In total, the patients have between 14 and 19 catheters inserted, with the dwell positions ranging from 161 up to 272 per patient.

Each of the patients has different characteristics due to their differences in anatomy. This means that for some patients a treatment plan that satisfies all aims in the clinical protocol can hardly be found. On the other hand, for another patient it might be easy to obtain a treatment plan that satisfies all aims.

Table 7.1: Table stating number of catheters, total number of dwell positions and number of dwell positions per catheter per patient.

Patient	Catheters	Total dwell positions	Dwell positions per catheter
1	19	272	6-19
2	16	226	11-19
3	15	253	10-21
4	16	161	4-15
5	14	236	9-27
6	17	270	9-22
7	17	211	7-18

Linkage Tree. Both BRIGHT and Set-BezEA use MO-RV-GOMEA at their basis, due to which both algorithms are capable of partial evaluations. Therefore, a static linkage tree will be used that is computed a priori based on the distances between the dwell positions, which is equal to the current setup used in BRIGHT. With these distances stored in a dependency structure matrix, the FOS elements for both algorithms are created using the UPGMA method as described in Section 2.4.2. Here, UPGMA first clusters the decision variables corresponding to the closest dwell positions. A lower bound of 5 decision variables (dwell times) per FOS element is used, as changing just 1 decision variable value hardly changes the overall outcome of a treatment plan in cases with up to 300 decision variable values, and a partial evaluation of a single dwell time has a relatively high computational cost. In Set-BezEA the lower bound on the FOS element size is increased by the number of control points used, similar to the structure described in Section 5.3. The lower bound is enforced after the hierarchical clustering step. This entails the pruning of the linkage trees at their leaves, until all leaves of the linkage tree have the required minimum size.

Optimization and Reevaluation. A comparison based purely on compute time would be unfair as Set-BezEA currently has no GPU parallelization built in, whereas BRIGHT has been optimized to make use of parallelization with a GPU [10]. Furthermore, the number of dose calculation points is highly influential on the outcome. They can significantly increase or decrease the required computation time, since these are used directly in the calculation of the DV indices. In order to keep the computation time manageable for Set-BezEA, a total of 20,000 dose calculation points are used during optimization for both algorithms, i.e., 4,000 per target volume and OAR. After optimization, all of the solutions are reevaluated with 500,000 dose calculation points to obtain more accurate results. These settings comply with the research performed with HDR brachytherapy on single core processor systems [42].

Since Set-BezEA optimizes parameterized approximation sets, the reevaluated approximation sets will be post-processed after optimization to obtain 100 solutions per approximation set. These larger approximation sets allow for more precise plots to be produced instead of showing only the p test points that are scattered apart.

Budget. In clinical practice, BRIGHT is given 300 seconds of compute time in order to optimize treatment plans for a patient, due to which roughly 10^6 evaluations are performed per patients in the case of 20,000 dose calculation points. When more dose calculation points are used, the amount of evaluations within the 5 minutes naturally reduces due to the increased computational cost of a single evaluation. Since the purpose of the experiments conducted here is to obtain a comparison to BRIGHT in the clinical setting, BRIGHT will also be given 300 seconds of compute time in these experiments. However, depending on the number of control points, the problem dimensionality increases for Set-BezEA when optimizing Bézier curve parameterizations, and hence makes it a more difficult optimization problem. This problem worsens even more when asking Set-BezEA for more curves (approximation sets), which entails a further increase of the problem dimensionality. For example, with 2 control points and 2 Bézier curves, Set-BezEA has 4 times as many decision variables to optimize as compared to BRIGHT. Thus, in the experiments the amount of evaluations given to Set-BezEA will be increased depending on the number of curves it is asked to produce to give the algorithm enough of a budget to obtain decent results and determine the hypervolume convergence. Set-BezEA is also given a high limit on the computation time in order for it to reach its evaluations budget first.

Parameter Overview. The rest of the decision variables are kept at their advised values from relevant literature. BRIGHT is given a population size of 96 [10], whereas Set-BezEA is given a population size of 76 based on 2 control points [72]. Other parameter values used in clinical practice, like the initialization range and elitist archive size, were obtained directly from one of the authors of BRIGHT. Table 7.2 shows an overview of the parameter values used for both BRIGHT and Set-BezEA.

Table 7.2: Experimental setup for BRIGHT and Set-BezEA on the HDR brachytherapy problem with AMC2020 protocol

Parameter	**Algorithm**	
	BRIGHT	**Set-BezEA** ($p = 10, q = 2, b \in \{2, 3, 4\}$)
Dose calculation points	20.000 (4.000 per volume)	
Initialization range	$[0, 1]$	
Population size	96	76
Elitist archive size	1000	50 (equals 1000-2000 test-solutions: 50 * 10 * b)
Maximum evaluations	10^6	$b * 7.5 * 10^5$
Maximum runtime	300s	86400s (24h)
Linkage tree	Static (based on dwell position distance)	
FOS element lower bound	5	10 ($5 * q$)

7.2 Performance Indicators

Hypervolume. In order to evaluate the empirical results, the hypervolume indicator will be used to assign a quality to the approximation fronts. The reference point is chosen to be -0.2 for LSI_n as treatment plans on the approximation fronts below this value have very small differences in LCI, see for example Figure 1.2. For LCI_n, the reference point is chose to be -0.8, which is a direct translation from the reference point -0.04 [72] used in the non-normalized LCI to the normalized LCI_n, which is a multiplication by a factor of 20 based on the range of satisfying the clinical aims being $[0, 0.05]$ for LCI and $[0, 1]$ for LCI_n.

For Set-BezEA the Hypervolume value is computed based on the $p = 10$ test points. Thus, similar to the employed procedure from the previous chapter, the approximation set of BRIGHT will be post processed with gHSS [28] to reduce its size to 10. The results will be averaged over a total of 10 runs. Even though an average over more runs would be preferential, this was currently not possible due to time limitations as a result of the runtime of Set-BezEA.

In the clinic there is no time to optimize multiple times for a specific patient before radiation treatment. Thus, it is important for the algorithm to produce decent results in terms of hypervolume, as a higher hypervolume is equivalent to obtaining an approximation front with plans that have lesser violations of the clinical aims or better satisfied all clinical aims. Even though hypervolume alone does not at all implicitly define clinically acceptable results, it does allow for an empirical comparison between the two algorithms.

Since Set-BezEA can be asked to produce several approximation sets it is interesting to see what the minimum and maximum obtained HV values are for those approximation sets. To see the extremes in terms of obtained HV, Box and Whisker plots will created using the results of all 10 runs. These plots will furthermore show the median and its 95% confidence interval through notches in the boxplot by bootstrapping the results 10,000 times.

Many other performance indicators, like the Pareto Set Proximity indicator, require the computation of the Pareto set. Since the HDR brachytherapy problem is a real-world problem for which limited domain knowledge is available, no optimum can currently be computed. Therefore, the Pareto set is unknown, thus making many other indicators non applicable.

Satisfying Clinical Aims. Overall, the hypervolume plots indicate the quality of the approximation front. However, those numbers do not translate to having satisfied all clinical aims de-

scribed in the protocol. In order to concretize whether the algorithms have achieved an approximation front that has satisfied all aims, there need to be plans on the front of which the normalized LCI_n and normalized LSI_n values must both be larger than 0. For example, the approximation fronts with plans that have satisfied all clinical aims run through the first quadrant of a plot, i.e., $LCI_n > 0$ and $LSI_n > 0$. The first quadrant is also denoted as the "golden corner".

The best and worst optimized approximation fronts in terms of obtained hypervolume will be plotted. These are shown in order to further investigate if all clinical aims are satisfied. All of the plots will be added to the appendix.

The LCI_n and LSI_n display the least satisfied or most exceeded DVI values, as seen in Equation 2.8. Even though the LCI_n and LSI_n are related to having satisfied the clinical aims in general, it is still of interest to observe the behaviour of the underlying DVIs. Therefore, the so-called delta DVI values will be plotted against the LCI_n value to see their course when navigating the approximation front. The delta DVI values are computed accoring to Equations 2.5, 2.6, and 2.7, respectively for the dose and volume indices of the coverage group and all indices of the sparing group. The plots will show the delta DVIs for the runs in which the best and worst approximation fronts were obtained and are also added to the appendix for the reader to further investigate them.

Smoothness and Dose Distribution Dissimilarity. Results in terms of the smoothness indicator will be omitted. It is expected that BRIGHT will produce less smooth results, as was the case for MO-RV-GOMEA in Table 6.4 of Chapter 6. Manual inspection of the statistics output of the experiments also pointed out that Set-BezEA always produced results that are perfectly smooth. This is due to the setup of the experiments as Set-BezEA works with 2 control points, which inherently produces perfectly smooth approximation sets. Therefore, it is more of interest to interpret parallel coordinate plots for the decision variable values of certain approximation sets. These could indicate that Set-BezEA can have larger or smaller variances, even though the approximation set is perfectly smooth.

Another type of variance that is to be inspected is that of dose distribution dissimilarity. The second objective of Set-BezEA tries to force all approximation sets to be as dissimilar as possible through the use of the cosine similarity measure. The Euclidean distance measure was omitted from the second objective. This was done because preliminary testing showed that the use of the Euclidean distance as a second objective would lead to certain dwell positions getting the absolute maximum dwell time. The maximum dwell time at a single location is not desirable due to the extreme amount of radiation that would produce locally.

The Euclidean distance measure will, however, be used to see the dissimilarity between Set-BezEA and BRIGHT. For each of the three settings of Set-BezEA ($b \in \{2, 3, 4\}$), the average Euclidean distance between all of the plans in each of the b approximation sets will be compared to the average distance of the b approximation sets to the approximation set of BRIGHT. This is to quantify whether there is a larger variance of approximation sets in Set-BezEA, i.e., what the second objective should enforce, than when compared to BRIGHT. A Welch T-Test with $p < 0.0001$ will be used to determine whether, for each of $b \in \{2, 3, 4\}$ separately, the difference amongst approximation sets of Set-BezEA is larger than the distance between Set-BezEA and BRIGHT.

For comparison, the average Euclidean distance between all plans in all approximation sets by BRIGHT will also be determined. This is to further quantify and substantiate claims whether or not Set-BezEA is capable of actually obtaining different treatment plans.

Reevaluation Impact. All sets of optimized treatment plans are reevaluated at the end of a run. For BRIGHT, this entails the reevaluation of both the population and elitist archive in order to obtain a accurate approximation set after reevaluation. Since the reevaluation significantly increases the accuracy of the dosimetry engine, e.g. 20,000 vs 500,000 total dose calculation points, it could very well be that BRIGHT might overfit on the smaller set of dose calculation points. This can happen due to the fact that BRIGHT has more freedom in choosing plans from all modes to put in its approximation set, whereas Set-BezEA has to model complete approximation sets within a single niche.

The reevaluation of the approximation sets with 500,000 dose calculation points is highly influential on the resulting HV values, as it will result in a more precise dose distribution. In order to evaluate the impact of the reevaluation, the hypervolume values in the last generation before reevaluation and directly after will be compared, from which the relative reduction in HV

is deduced. Even though a reduction in HV value is expected, significantly different values after reevaluation can signal a form of overfitting on a lower number of dose calculation points. If that would be the case, more investigation would be necessary with a higher number of dose calculation points during optimization to create a fair comparison.

7.3 Empirical Results

Appendix B holds all of the figures for each of the patients for which the experiments were run. In this section, the most important results will be shown in tables, with references to important figures in the appendix.

Hypervolume Results. Table 7.3 shows the mean obtained HV results for the various algorithms on each of the 7 patients. The table shows that BRIGHT is capable of obtaining slightly better HV values, even though the results do not differ by large margins. This is to be expected as BRIGHT has more freedom in choosing what plans it adds to the approximation set, whereas Set-BezEA enforces each approximation set to be within a single niche.

When Set-BezEA is asked to produce more approximation sets, the obtained HV does reduce slightly. Depending on the patient, the reduction for Set-BezEA with 4 approximation sets as compared to 2 approximation sets differs between 3% and 6%. Set-BezEA with $b = 3$ approximation sets obtains HV values in between Set-BezEA with $b = 2$ and Set-BezEA with $b = 4$. These results are to be expected as each added approximation set increases the number of decision variables Set-BezEA has to optimize, which makes the problem more difficult.

Table 7.3: Average +- standard deviation results of approximation front HV on HDR brachytherapy after reevaluation, averaged over 10 runs. Bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction).

Patient	BRIGHT	Set-BezEA (b = 2)	Set-BezEA (b = 3)	Set-BezEA (b = 4)
1	**3.6535e-1 (± 1.2102e-2)**	3.4976e-1 (± 5.2456e-3)	3.4327e-1 (± 5.6515e-3)	3.4084e-1 (± 7.3479e-3)
2	**3.0698e-1 (± 7.7673e-3)**	2.6290e-1 (± 1.2327e-2)	2.5895e-1 (± 1.0812e-2)	2.4973e-1 (± 1.4539e-2)
3	**3.3431e-1 (± 7.0481e-3)**	3.1818e-1 (± 6.7078e-3)	3.1504e-1 (± 7.0154e-3)	3.1370e-1 (± 8.2677e-3)
4	**2.7762e-1 (± 4.2974e-2)**	**2.8622e-1 (± 8.9329e-3)**	**2.8055e-1 (± 1.4013e-2)**	2.6694e-1 (± 2.0364e-2)
5	**2.8635e-1 (± 6.3741e-3)**	2.6710e-1 (± 3.2768e-3)	2.5779e-1 (± 6.8871e-3)	2.5134e-1 (± 8.1127e-3)
6	**3.6183e-1 (± 1.5260e-2)**	3.5139e-1 (± 7.9797e-3)	3.4359e-1 (± 9.5722e-3)	3.3723e-1 (± 9.8280e-3)
7	**4.0316e-1 (± 2.3210e-3)**	3.8066e-1 (± 3.5757e-3)	3.7279e-1 (± 6.0670e-3)	3.7401e-1 (± 6.6728e-3)

Satisfying Clinical Aims. Figure B.3 shows the best and worst achieved approximation front over all runs. The approximation fronts with plans that have satisfied all clinical aims run through the first quadrant of the plots. Table 7.4 shows the average number of approximation fronts that have satisfied all clinical aims over all 10 runs. Set-BezEA ($b \in \{2, 3, 4\}$) creates b times the number of approximation sets as compared to BRIGHT. Therefore, the values in Table 7.4 seem to be rounded to different values for the different algorithms and settings.

Table 7.4 indicates that, on average, Set-BezEA is capable of obtaining treatment plans that satisfy the clinical aims. This is despite Table 7.3 indicating that Set-BezEA obtains worse hypervolume values as compared to BRIGHT. Asking Set-BezEA for more approximation sets does seem to hinder its capabilities in terms of obtaining plans that have satisfied the clinical aims, even though it still converges. This behaviour is similar to the slightly reduced hypervolume values for higher values of b as observed in Table 7.3.

It is interesting to see that the patients on which Set-BezEA ($b = 4$) (significantly) reduces its chances of satisfying the clinical aims are those for which it is harder to satisfy these aims. This claim is substantiated by Figures B.3m through B.3t for patients 4 and 5. In all cases for all algorithms, these figures show that if plans that have satisfied the clinical aims were found, these aims were only barely satisfied. In the case of patient 2, Set-BezEA ($b = 4$) is the only algorithm that does not satisfy all aims in all cases. The other algorithms are capable of doing so. However, Figure B.3e shows that the worst approximation front of BRIGHT only barely skims the golden corner. The difficulty of satisfying the aims for these patients could indicate that they are harder to optimize for due to the placement of catheters or their anatomy. Thus, since Set-BezEA ($b = 4$)

has to optimize 8 times the number of decision variables as compared to BRIGHT, the reduction in terms of capability of satisfying all aims is to be expected.

Figure B.4 shows the delta DVI plots for the runs in which the approximation sets with the best and worst values in terms of hypervolume were found. For all patients, a larger variance is observed in the $D^{Bladder}_{1.0cm^3}$ and $D^{Bladder}_{2.0cm^3}$ for each of Set-BezEA $b \in \{2, 3, 4\}$. The bladder is the largest organ of all organs considered in the evaluation, due to which it has the lowest accuracy in dose computation, since there are 4,000 dose calculation points scattered throughout bladder during optimization and the reevaluation is done with 100,000 dose calculation points. In certain cases a variance can be observed in some of the other DVIs that belong to the sparing group as described in Section 2.8.2. For the DVIs that belong to the coverage group, only the $V^{Vesicles}_{73\%}$ shows any variance. This might also be attributed to the smaller number of dwell positions in catheters that extend higher into the patients body, where the vesicles are located above the prostate. When only a few dwell positions are used, a small change in one of them can influence the outcome of a DVI faster. The other two DVIs in the coverage group, $V^{Prostate}_{100\%}$ and $D^{Prostate}_{90\%}$ hardly show any variance.

When the delta DVIs are compared to those from BRIGHT, similar results are seen for all of the patients. Furthermore, the course of the delta DVI values barely differs in most cases, with clearly visible kinks in the plots at similar locations. What does stand out is that Set-BezEA does not extend into the extremes as much. This can be attributed to the use of the reference point, which prohibits the approximation fronts from reaching below $LSI_n < -0.2$ and $LCI_n < -0.8$.

Answer of Research Question 3.1

Table 7.4 indicates that Set-BezEA is indeed capable of producing treatment plans that satisfy all clinical aims. Asking for more approximation sets does reduce the chance that all clinical aims are satisfied. However, larger population sizes at the cost of more computation time could be a solution for this problem which has currently not yet been investigated. All together, Set-BezEA has indeed been found to be capable of producing treatment plans that satisfy all clinical aims.

Table 7.4: Percentage of runs that achieved approximation fronts in which at least one plan has satisfied all clinical aims ($\text{LCI}_n > 0$ and $\text{LSI}_n > 0$)

Patient	BRIGHT	Set-BezEA ($b = 2$)	Set-BezEA ($b = 3$)	Set-BezEA ($b = 4$)
1	100%	100%	100%	100%
2	100%	100%	100%	92.5%
3	100%	100%	100%	100%
4	60%	90%	80%	50%
5	90%	85%	53.33%	17.5%
6	100%	100%	100%	100%
7	100%	100%	100%	100%

Smoothness and Dose Distribution Dissimilarity. Figures B.5a to B.5g show, for each of the 7 patients respectively, the parallel coordinates plots for the run in which the best approximation set was found in terms of obtained hypervolume in objective space. These plots can be used to denote the smoothness of the approximation sets.

Figure 7.1 shows the parallel coordinates plots for patient 1. Each of these parallel coordinates plot shows the treatment plan ID on the horizontal axis, sorted on increasing LCI values with the lowest LCI on the left and highest on the right. By doing so, plans that are adjacent on the approximation front are now adjacent in the parallel coordinates plots. The vertical axis shows the value of each of the decision variables. The lines in the plot depict the course of each of the decision variables separately. Only the 10 decision variables with the highest standard deviation are highlighted in color. The others have been given a slightly transparent grey color in order to not clutter the plots too much.

The parallel coordinate plots clearly show ragged lines for BRIGHT, indicating that it has a very low smoothness. The plans from BRIGHT originate from various modes in the fitness landscape, resulting in the large dissimilarity in the dose distribution of plans in the approximation set. When the results for BRIGHT are compared to those from Set-BezEA with $b \in \{2, 3, 4\}$, the differences are distinct. Set-BezEA produces approximation sets with smoothly varying decision variables, clearly optimized with only $q = 2$ control points as the values are interpolated in a linear fashion between the endpoints.

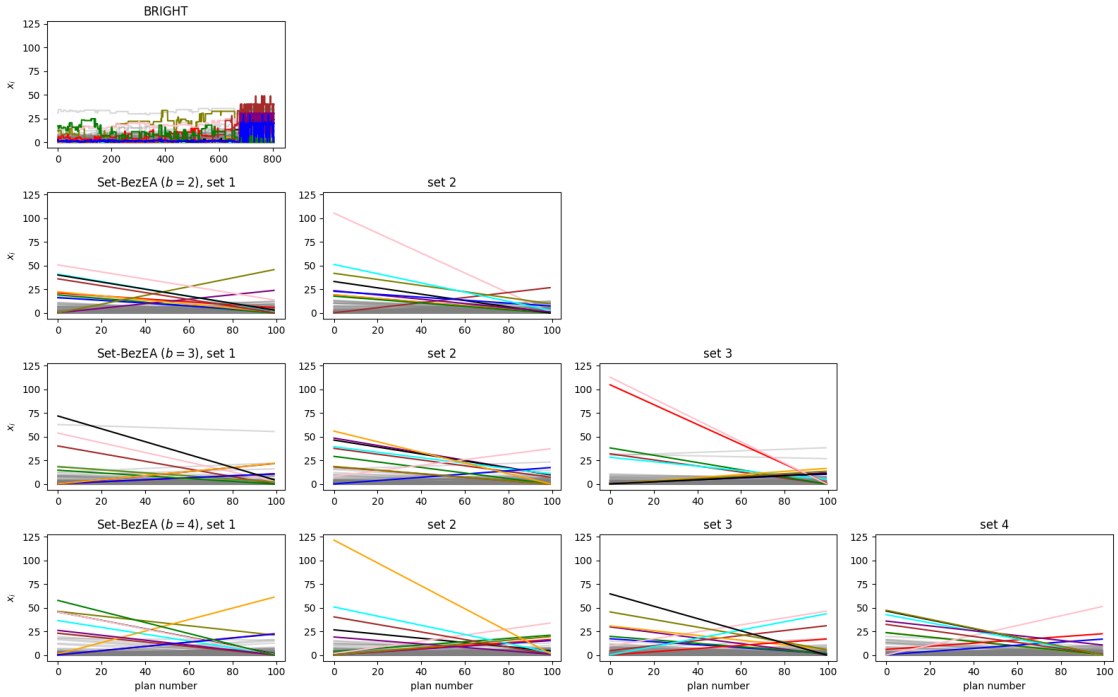


Figure 7.1: Parallel coordinates plot of best result per algorithm over 10 runs for patient 1

What is also clear from the parallel coordinates plots is the clustering step that is performed in BRIGHT. As BRIGHT has MO-RV-GOMEA as its optimization algorithm, it clusters the population into 5 groups based on their objective values. Figures B.5a, B.5b, B.5d and B.5g all show significantly higher dwell times for specific dwell positions on the right side of the parallel coordinates plots. The relative size of this particular area always seems to roughly be 20% of the treatment plans, which corresponds to 1 of the 5 clusters created in MO-RV-GOMEA in each generation. This indicates that the algorithm has successfully optimized towards higher dwell times for higher coverage values in the cluster that focuses on more coverage.

Figure B.5f that displays the parallel coordinates for Patient 6 stands out amongst the others. Many dwell positions seem to have a high variance and/or very high values. There are even some dwell positions that nearly maintain their maximal value throughout the entire approximation sets. Furthermore, the dwell times in these dwell positions has a low variance and is therefore depicted in transparent grey. This behaviour is prevalent in both the results for BRIGHT and Set-BezEA.

Table 7.5 shows the dissimilarity of all plans in the approximation sets as measured by the Euclidean distance. The table clearly indicates that when Set-BezEA is compared to BRIGHT, in all three of the cases of $b \in \{2, 3, 4\}$, the difference between plans within each of the b approximation sets is significantly larger than if all approximation sets of Set-BezEA are compared to BRIGHT. Furthermore, the last column that denotes the difference amongst plans from BRIGHT itself is significantly smaller than between approximation sets of Set-BezEA or between Set-BezEA and BRIGHT, 2 to 3 times as small to be more precise.

The results thus indicate that the second objective is working as intended by producing dissimilar plans. Even though the hypervolume values are slightly lower as seen in Table 7.3, the dissimilarity between treatment plans (solutions) is significantly larger. These two findings further emphasize that Research Question 2 has successfully been answered in the form of the Set-BezEA algorithm with the adaptations to MO-RV-GOMEA and the newly proposed fitness function in Equation 13.

Reevaluation Impact. Table 7.6 shows the mean obtained HV values in the last generation before reevaluation with 500.000 dose calculation points. Table 7.7 shows the percentual reduction between the obtained HV values of Tables 7.6 (before reevaluation) and 7.3 (after reevaluation). The average reduction seems to be similar for all three instances of Set-BezEA, denoting that using

Table 7.5: Average ± standard deviation of Euclidean distance between the b approximation sets of Set-BezEA and between Set-BezEA and BRIGHT. Bold denotes the column of which all values are significantly larger for each of Set-BezEA $b \in \{2, 3, 4\}$ separately, tested with Welch T-test ($p < 0.0001$)

	Set-BezEA ($b = 2$)		Set-BezEA ($b = 3$)		Set-BezEA ($b = 4$)		BRIGHT
Patient	**Between b sets**	To BRIGHT	**Between b sets**	To BRIGHT	**Between b sets**	To BRIGHT	To BRIGHT
1	**1.21e2(±2.13e1)**	9.28e1(±2.21e1)	**1.30e2(±2.62e1)**	1.00e2(±2.45e1)	**1.23e2(±2.70e1)**	9.76e1(±2.41e1)	4.84e1(±2.79e1)
2	**2.34e2(±3.24e1)**	1.90e2(±4.07e1)	**2.69e2(±3.59e1)**	2.16e2(±4.12e1)	**2.61e2(±4.39e1)**	2.17e2(±4.15e1)	1.00e2(±4.90e1)
3	**1.71e2(±3.36e1)**	1.26e2(±3.11e1)	**1.73e2(±3.73e1)**	1.31e2(±3.34e1)	**1.74e2(±4.03e1)**	1.33e2(±3.62e1)	6.21e1(±3.60e1)
4	**1.30e2(±3.93e1)**	1.46e2(±3.56e1)	**1.65e2(±6.18e1)**	1.62e2(±4.34e1)	**1.74e2(±5.25e1)**	1.73e2(±4.50e1)	1.12e2(±5.79e1)
5	**1.71e2(±4.36e1)**	1.30e2(±3.20e1)	**1.89e2(±4.15e1)**	1.48e2(±3.49e1)	**2.08e2(±3.87e1)**	1.62e2(±3.56e1)	6.96e1(±3.30e1)
6	**3.74e2(±4.27e1)**	2.83e2(±5.18e1)	**3.60e2(±4.56e1)**	2.92e2(±4.71e1)	**3.66e2(±4.95e1)**	3.05e2(±5.19e1)	1.55e2(±7.26e1)
7	**2.26e2(±3.40e1)**	1.70e2(±3.70e1)	**2.17e2(±3.96e1)**	1.68e2(±3.90e1)	**2.18e2(±4.16e1)**	1.71e2(±4.01e1)	8.72e1(±4.81e1)

more or less approximation sets do not significantly impact the reduction of HV after reevaluation.

For patients 1, 3 and 7, the reduction seems similar between BRIGHT and Set-BezEA. For patient 2, Set-BezEA seems to have a larger reduction as compared to BRIGHT. On the other hand, for patient 4, 5, and 6, BRIGHT seems to obtain a larger average reduction in obtained HV. In these cases, larger reduction can signal possible overfitting on the lower number of dose calculation points during optimization (20.000).

The reduction is largest on patient 4, where BRIGHT on average loses just over 15% in terms of HV. In Table 7.3, BRIGHT seems to obtain slightly lower HV values as compared to Set-BezEA with $b \in \{2, 3\}$. In this specific instance this could actually be an instance of overfitting which could be indicated by the large reduction in obtained HV. However, the boxplots of Figures B.1g and B.1h show that there are a few outliers in terms of obtained HV for an approximation set for BRIGHT. These fall significantly below the whiskers of the boxplot, due to which underlying causes for the large differences cannot be ruled out, such as the fact that the anatomy of the patient could be difficult to optimize for.

Overall it seems that there is no sign of overfitting as for most patients the percentual reduction seems nearly equal over all patients for both algorithms. Furthermore, BRIGHT still obtains slightly higher hypervolume values after reevaluation, signalling that it has optimized well. If it would have overfitted, BRIGHT should have gotten significantly lower HV values than Set-BezEA after reevaluation.

Table 7.6: Average +- standard deviation results of approximation front HV on HDR brachytherapy before reevaluation, averaged over 10 runs. Bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction)

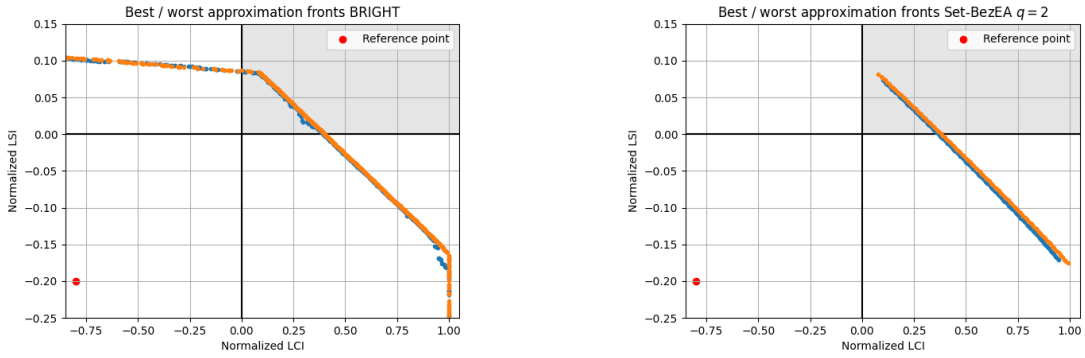
Patient	BRIGHT	Set-BezEA ($b = 2$)	Set-BezEA ($b = 3$)	Set-BezEA ($b = 4$)
1	**3.8698e-1 (± 4.2591e-3)**	3.6495e-1 (± 2.1878e-3)	3.5966e-1 (± 3.2784e-3)	3.5654e-1 (± 3.5257e-3)
2	**3.3526e-1 (± 9.2887e-3)**	2.9997e-1 (± 4.5612e-3)	2.9407e-1 (± 5.0329e-3)	2.8930e-1 (± 6.6853e-3)
3	**3.6113e-1 (± 8.3861e-3)**	3.5216e-1 (± 2.4213e-3)	3.4741e-1 (± 2.8266e-3)	3.4511e-1 (± 3.3478e-3)
4	**3.2740e-1 (± 3.6546e-2)**	3.1124e-1 (± 7.7427e-3)	3.0674e-1 (± 1.3397e-2)	2.9293e-1 (± 1.8680e-2)
5	**3.1916e-1 (± 8.6753e-3)**	2.7869e-1 (± 2.4864e-3)	2.7051e-1 (± 6.7123e-3)	2.6306e-1 (± 9.3195e-3)
6	**3.9956e-1 (± 1.8225e-2)**	3.6345e-1 (± 3.7015e-3)	3.5160e-1 (± 6.6463e-3)	3.4668e-1 (± 9.9455e-3)
7	**4.1428e-1 (± 3.6507e-3)**	3.9017e-1 (± 1.3737e-3)	3.8622e-1 (± 2.2453e-3)	3.8525e-1 (± 2.6251e-3)

Table 7.7: Percentual reduction of approximation front HV on HDR brachytherapy after reevaluation, averaged over 10 runs

Patient	BRIGHT	Set-BezEA ($b = 2$)	Set-BezEA ($b = 3$)	Set-BezEA ($b = 4$)
1	5.59%	4.16%	4.56%	4.41%
2	8.43%	12.36%	11.94%	13.68%
3	7.43%	9.65%	9.32%	9.10%
4	15.21%	8.04%	8.54%	8.87%
5	10.28%	4.16%	4.70%	4.46%
6	9.44%	3.32%	2.28%	2.73%
7	2.68%	2.44%	3.48%	2.92%

7.4 Higher Order Bézier Curve Parameterizations

The results for patient 7 of the previous section indicated that, in all three settings of Set-BezEA $b \in \{2, 3, 4\}$, the approximation fronts abruptly stopped in the middle of the first quadrant, i.e., golden corner. This can be seen in Figures B.3z to B.3ab that display the approximation fronts. Table 7.3 also indicates that for this specific patient, the highest hypervolume values out of all patients were found. This possibly indicates that this patient is comparatively easy to optimize for as compared to the other patients.



(a) BRIGHT

(b) Set-BezEA ($b = 2, q = 2$)

Figure 7.2: Patient 7 approximation front

Set-BezEA tries to optimize the hypervolume of the area below the approximation front bound by the reference point. Figure 7.2a indicates that the point at which the approximation fronts of Set-BezEA abruptly stop, the point in Figure 7.2b with the highest LSI_n, is at one of the kinks in the approximation front of BRIGHT. This could possibly be caused by the fact that Set-BezEA found that extending the approximation fronts to lower LCI_n and higher LSI_n values did not impact the overall hypervolume value enough due to the marginal increase in terms of LSI_n versus a sharp decrease in terms of LCI_n. However, it could very well be the case that the plans in the golden corner that are currently not found, as the approximation front abruptly stops, can be of clinical interest since they would satisfy all clinical aims. Therefore, further experiments are required to see if the approximation front can be made to fully cover the golden corner.

Setup. In order to evaluate if Set-BezEA is capable of covering the golden corner completely with an approximation front, more experiments are run. The setup is similar to the one described in Table 7.3 of Section 7.1, but with slightly different settings for Set-BezEA. Now, Set-BezEA is always asked to produce $b = 2$ approximation sets, whereas the number of control points is varied with $q \in \{2, 3, 4\}$. The experiments are only conducted for patient 7, where the undesired behaviour of the approximation front stopping within the golden corner was noticed.

Results. Figure 7.3a shows the behaviour in terms of hypervolume performance. Set-BezEA with $q = 4$ control points and $b = 2$ approximation sets has the same number of decision variables as Set-BezEA with $b = 4$ approximation sets and $q = 2$ control points. The figure shows that when increasing the number of control points it is slightly harder to obtain high-quality approximation fronts as compared to increasing the number of approximation sets. This could be caused by the fact that it is harder to correctly unfold a higher order Bézier curve in order to have all test solutions be undominated. For example, in the case of $q = 3$ control points, in objective space one can visualize the third control point to be in between the objective values of the two control points that parameterize the endpoints of the approximation set. As it is harder to optimize the decision variables in such a way that this third control point is in the correct position, it takes more evaluations before Set-BezEA obtains any hypervolume value.

Figures 7.3b and 7.3c, respectively showing the hypervolume boxplots before and after reevaluation, combined with Table 7.8 show the behaviour in terms of reduced hypervolume after reevaluation for Set-BezEA. This behaviour is similar to what is observed in the original experimental

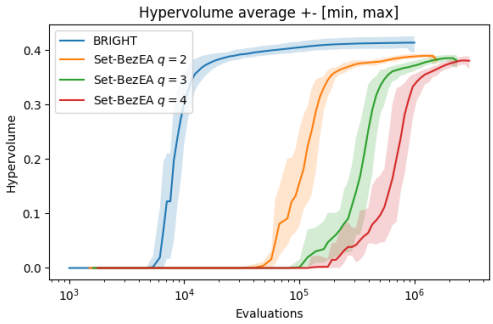
setup and its results in Tables 7.3, 7.6 and 7.7, that display the hypervolume before, after and percentual reduction respectively. All of these results combined indicate that it is probable that the lower hypervolume results are more likely caused by an increase in decision variables that are to be optimized, which adds more complexity to the problem. Since, either asking Set-BezEA to produce more approximation sets, i.e., setting b to higher values, or asking Set-BezEA to optimize approximation sets through higher order Bézier curves, i.e., setting q to higher values, will result in lower hypervolume values for the approximation fronts. However, the results also denote the fact there are hardly any performance differences between asking approximation sets with more control points or asking more approximation sets in total, as the obtained approximation fronts are of similar quality on average when b or q are set to equal values.

Most importantly, Figures 7.3d and 7.3e, for $q = 3$ and $q = 4$ control points respectively, shows that the best and worst approximation fronts now all fully cover the entire golden corner. They no longer abruptly stop within the golden corner, but cut through it and end outside of the region where all clinical aims are satisfied. This is exactly the desired behaviour as plans that fall slightly outside of the golden corner could potentially be of clinical interest if these plans better suit the preference of the clinicians based on patient-specific characteristics. However, it is interesting to see that the approximation fronts are spread out further with larger steps in LCI_n between adjacent plans along their nearly horizontal part, i.e., the part that was previously missed. It seems that the plans that are in this particular area quickly increase the normalized LCI for small reductions in terms of normalized LSI.

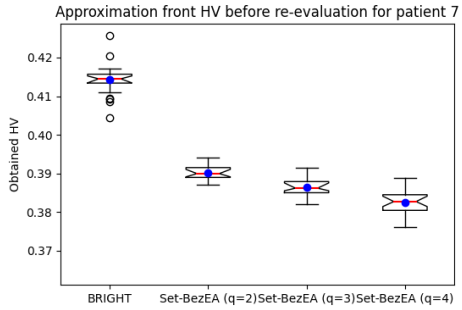
Figure 7.3f displays the parallel coordinates plots for BRIGHT and Set-BezEA with higher order Bézier curves. The plots indicate the large differences in dwell times in the plans originating from BRIGHT. They also show the differences between the various order of Bézier curves. With $q = 2$, linear interpolation of the dwell times is observed with either monotonically increasing or decreasing values. For $q = 3$ parabolic shapes appear in some lines as some decision variables have a minimum or maximum in the middle of the approximation set, denoting the non-linearity. The highest order Bézier curves with $q = 4$ show their added freedom of parameterizing approximation sets as some decision variables now having saddle points in the course of their values.

Table 7.8: Hypervolume results for patient 7

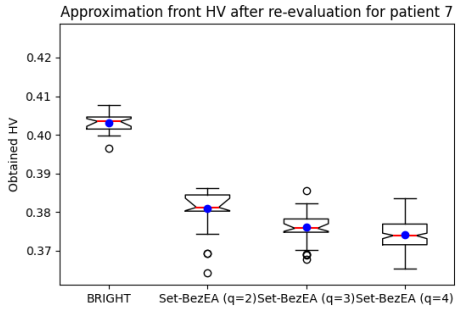
Value	BRIGHT	Set-BezEA ($q = 2$)	Set-BezEA ($q = 3$)	Set-BezEA ($q = 4$)
Before reevaluation	4.1428e-1 (± 3.6507e-3)	3.9017e-1 (± 1.3737e-3)	3.8642e-1 (± 2.0920e-3)	3.8244e-1 (± 2.8443e-3)
After reevaluation	4.0316e-1 (± 2.3210e-3)	3.8066e-1 (± 3.5757e-3)	3.7606e-1 (± 4.0116e-3)	3.7423e-1 (± 4.3300e-3)
Reduction	2.68%	2.44%	2.68%	2.15%



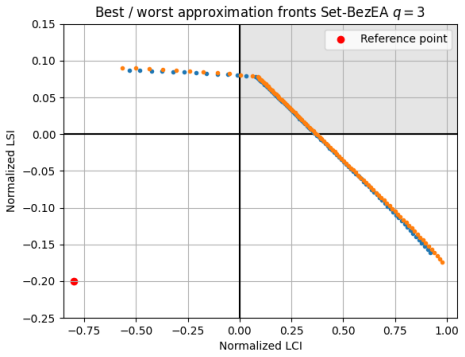
(a) Hypervolume performance



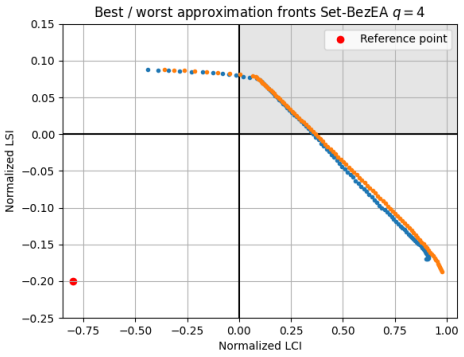
(b) Hypervolume boxplot before reevaluation. Red is median, blue is mean



(c) Hypervolume boxplot after reevaluation. Red is median, blue is mean

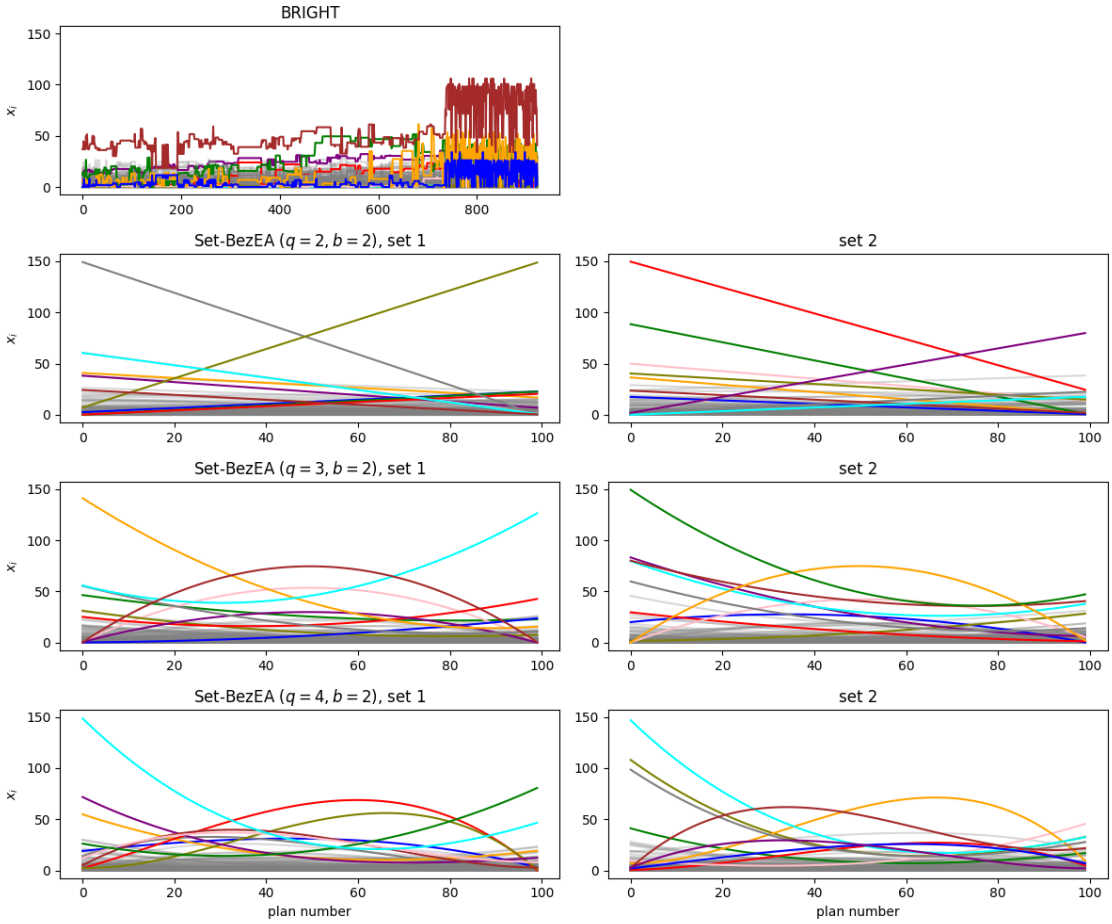


(d) Patient 7 approximation front Set-BezEA ($b = 2, q = 3$)



(e) Patient 7 approximation front Set-BezEA ($b = 2, q = 4$)

Figure 7.3: Results for higher order Bézier curves for patient 7



(f) Patient 7 parallel coordinates plot Set-BezEA

Figure 7.3: Results for higher order Bézier curves for patient 7

7.5 Clinical Evaluation of Plans

A radiation oncologist and medical physicist were asked to evaluate the plans of Set-BezEA in order to get a clinical evaluation of the treatment plans the algorithm produced. Even though performance indicators can concretize if an algorithm can outperform another, they cannot tell how good a treatment plan is for actual use in the clinic. Furthermore, treatment plans that were created through a Bézier parameterization of approximation sets were not yet evaluated by clinicians in earlier work [72]. Therefore, a clinical evaluation by experts is required in order to see if the treatment plans are truly acceptable in the clinic or not. The two clinicians asked to evaluate plans from BRIGHT and Set-BezEA for a comparison are:

- MD PhD Bradley Pieters, a medical specialist in radiation oncology at the Amsterdam UMC.
- MSc Danique Barten, a medical radiotherapy physicist at the Amsterdam UMC.

Setup For each of the 7 patients, one plan from Set-BezEA and aone plan from BRIGHT were selected to be evaluated. Although more treatment plans could, of course, be evaluated, in the interest of limited availability of the clinicians only one comparison is made per patient. The plans from Set-BezEA and BRIGHT were filtered based on them having satisfied all clinical aims. Here, the results of all three Set-BezEA settings were aggregated into a single set of plans. Of the resulting two sets of plans, i.e., all from Set-BezEA and all from BRIGHT, one plan was chosen from each. All of the Euclidean distances in normalized LCI and LSI ($\mathbb{R}^2$) are calculated for all pairs of plans between the two sets. The pair of plans, one from Set-BezEA and one from BRIGHT, with the smallest distance were selected for the evaluation.

The setup of the comparison is as follows. For each patient a visualization of the treatment plans was made using the commercial Oncentra Brachy software [73] (Elekta AB, Stockholm, Sweden) that is used in many clinics. As no side-by-side viewer is available in Oncentra Brachy, there is no option of creating a direct comparison between both treatment plans for a patient on the screen. Therefore, a plan from either BRIGHT or Set-BezEA was shown on screen to both clinicians. They could then go through the various slices of an MRI scan of the patient with the dose distribution information overlayed on top. All of their remarks regarding the clinical acceptability of the plans are transcribed during this process. Afterwards, the treatment plan from the other algorithm for the same patient is shown and their remarks again transcribed. Lastly, their preference between the two treatment plans is asked. The clinicians did not know which plan originated from which algorithm to provide a fair comparison based on treatment plan characteristics.

Results Table 7.9 shows all of the algorithm specific and general remarks made. The general remarks are remarks that apply to treatment plans originating from both algorithms. Table 7.10 shows the preferred plan per patient.

The remarks in Table 7.9 specific to the algorithms are predominantly about the dose distribution, which can be found to be slightly better in a plan from one of the two algorithms. However, the table does show some interesting general remarks, which can be summarized into two major points of interest:

- Dwell times of dwell positions that are adjacent to each other in a catheter vary too much
- Radiation outside of target volumes is usually too high

Figure 7.4 shows two screenshots taken from the Oncentra Brachy visualization software. The isodose lines give an indication of the dwell times of the activated dwell positions that are denoted by the red dots. The actual dwell times are displayed in Figure 7.4b, where the dwell times for the dwell positions in a specific catheter are shown.

Figure 7.4a shows that there are quite a few dwell positions outside of the prostate, with high amounts of radiation administered there. Especially on the left and right sides of the prostate, the normal tissue is irradiated too much in those locations. This should be prevented as much as possible according to the clinicians remarks since the healthy tissue should not incur these amounts of radiation. The cause can be found in the current clinical protocol, on which the bi-objective problem formulation is based. This protocol currently does not specify planning aims for the healthy tissue, due to which the algorithms do not optimize on the radiation in those

locations. Another of the clinicians remarks was targeted towards the activated dwell positions in the catheters, which are activated by clinicians before optimization. The algorithms simply take the activated dwell positions as input and optimize based on that. Therefore, this remark was targeted towards the clinicians themselves as they should not have enabled some of the dwell positions so far outside of the target volumes.

Figure 7.4b shows the dwell times for each of the dwell positions in a catheter. The figure shows a table from Oncentra Brachy that displayes the information about the dwell positions in one catheter. As can clearly be seen in the plot, position 7 has a dwell time of zero seconds, whereas position 8 has nearly 40 seconds. These sharp differences also occur around position 13. The clinicians indicated that the large dissimilarity between two adjacent dwell positions is not preferential.

The cause of both general remarks can be found in the current bi-objective problem formulation based on the protocol employed in the clinic. The protocol currently does not constrain radiation outside of the organs. A similar explanation can be found for the dwell times, as no constraint is currently put in place in the protocol to enforce a smooth distribution of dwell times within a single catheter. However, these issues have been researched and solutions have been proposed. Dwell Time Modulation Restriction can be used to smooth the dwell time distribution within catheters [14, 21] and normal tissue constraints can be added to the DV indices in order to prevent the irradiation of healthy tissue [21]. These solutions will be included in the future work section in order to further improve the characteristics of the treatment plans.

Table 7.10, that shows the plan preferences, indicates that there is hardly a difference between plans originating from either algorithm. In 2 out of 7 cases BRIGHT was found to be slightly better and in 1 case Set-BezEA. In the case of patient 3, a slight preference for BRIGHT came solely from the remark made for Set-BezEA (see Table 7.9) that one particular dwell position inside the bladder had a high dwell time, a so-called local hotspot. Even though the current protocol allows this, since the plan met all clinical aims of Table 2.2, it is of course more beneficial to not irradiate the bladder any more than necessary.

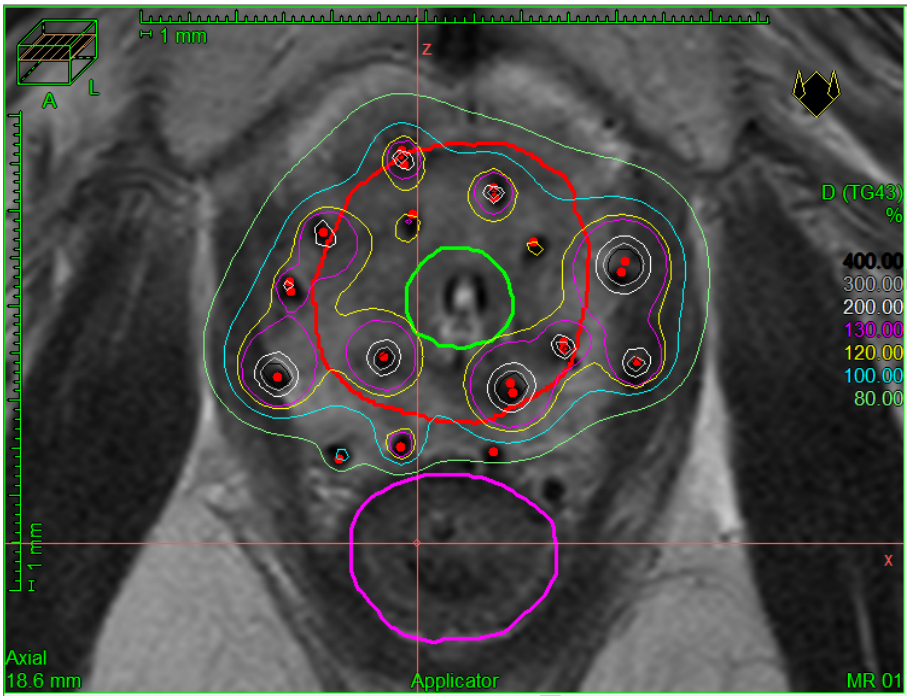
In the case of patient 5, BRIGHT was found to produce a slightly more preferential coverage of the entire prostate volume as compared to Set-BezEA. This was the sole reason for slightly preferring BRIGHT over Set-BezEA, since both treatment plans met the aspiration values of the DV indices. A similar situation occurred in the case of patient 7. However, here the roles are reversed, with Set-BezEA producing a slightly more preferential coverage of the prostate volume.

Answer of Research Question 3.2

Tables 7.9 and 7.10 indicate that plans from BRIGHT are not preferable in the clinical setting when compared to Set-BezEA. Based on the findings reported in the tables, Set-BezEA can be found to be comparable to BRIGHT, since there is no clear winner and the treatment plans from both algorithms received the same general remarks.

Answer of Research Question 3

It is the clinical protocol at the basis of the bi-objective problem formulation that allows room for the shortcomings of the optimized treatment plans as mentioned by clinicians. The clinical evaluation showed that the protocol has to be adjusted to include the additional desirable characteristics of treatment plans. Even though Set-BezEA can be further improved to overcome some of its current limitations, it produces treatment plans which were judged just as good as BRIGHT in the clinical setting. BRIGHT does produce slightly better hypervolume results, but in a comparison by clinicians no clear winner could be pointed out. Therefore, Set-BezEA is found to perform comparable to BRIGHT on the current bi-objective problem formulation.



(a) Screenshot from Oncentra Brachy showing a lot of radiation outside of prostate target volume. The thick red contour line denotes the delineation of the prostate, with the urethra in green and the rectum in purple. Thinner lines are isodose lines that denote areas that receive x% of the aimed planning dose. Isodose line legend is shown on the right. Red dots are dwell positions.

Catheter	Dwell pos.	Time [s]	Dwell weight
16	5	4.70	
16	6	0.29	
16	7	0.00	
16	8	38.38	4.84
16	9	16.83	2.12
16	10	2.16	
16	11	0.00	
16	12	1.12	
16	13	9.47	1.19
16	14	1.60	
16	15	6.94	
16	16	0.00	
16	17	0.00	
16	18	0.43	

(b) Screenshot from Oncentra Brachy showing large differences in dwell times within a catheter. Table displays all dwell positions within a catheter and the corresponding dwell time and weight, which is proportional to the dwell time.

Figure 7.4: Visualizations from Oncentra Brachy that show the cause of the general remarks by clinicians.

Table 7.9: Remarks from clinicians for investigated treatment plans per patient

Patient	**Set-BezEA specific**	**BRIGHT specific**	**General**
1	Slightly better distribution of dwell times.	Slightly better sparing of rectum.	Many dwell positions outside of prostate, making optimization harder. Differences in dwell times of adjacent positions in catheters are too large.
2	A lot of radiation within prostate.	Slightly more radiation outside of prostate.	Very high dwell times overall. Too many activated dwell positions outside of prostate, where it could have been beneficial to enable some currently inactive dwell positions.
3	One dwell position with high dwell time inside bladder.	Slightly better coverage inside prostate.	Lower coverage of upper part of prostate in order to spare urethra. Large dissimilarity of dwell times for adjacent dwell positions in catheters.
4	None.	One dwell position in prostate with huge dwell time.	Large dissimilarity of dwell times for adjacent dwell positions in catheters. Too much radiation outside of prostate.
5	Many dwell positions with impractically small dwell times.	Slightly higher dose outside of prostate.	Dwell positions far away from target volume are active. Large dissimilarity of dwell times for adjacent dwell positions in catheters.
6	Slightly better coverage in centre of prostate.	Some dwell positions with high dwell times.	Difficult patient with very few dwell positions inside the prostate, thus required to use many dwell positions at the edge or outside of the prostate.
7	Vesicles not completely covered.	Hotspots slightly larger.	Large dissimilarity of dwell times for adjacent dwell positions in catheters. Enabling dwell positions two slices below target volume is not acceptable.

Table 7.10: Preferences from clinicians for investigated treatment plans per patient

Patient	**Preference**
1	None
2	None
3	Slightly for BRIGHT
4	None
5	Slightly for BRIGHT
6	Slightly for Set-BezEA
7	None

Chapter 8

Discussion

The empirical benchmarks of Chapter 6 of this thesis show promising results for both MM-BezEA and Set-BezEA. This was further underlined for Set-BezEA by the experiments in Chapter 7. However, the used metrics and techniques behind Bézier curve parameterizations and employed performance indicators have some limitations. These will be discussed in the paragraphs below. The future works section touches on some further research topics that could potentially improve the performance of EAs that build on Bézier curve parameterizations.

8.1 Bézier Curve and Algorithm Limitations

Hyperparameter Settings MM-BezEA and Set-BezEA have been used to optimize various problems, ranging from simpler ones to very complex problem instances like HDR brachytherapy. The same set of hyper parameters has been used in all problem instances, e.g., equal population size. However, especially in higher dimensional problems like HDR brachytherapy, it could very well be the case that the optimal performance of Set-BezEA or MM-BezEA has not yet been reached and could be improved with better problem specific parameter settings. However, the computational cost of optimizing the hyper parameters in computationally expensive optimization made it intractable of doing so during this thesis.

Discontinuous Objective Functions MM-BezEA and Set-BezEA did not cover the endpoints of the PSs in the case of MMF12, which resulted in lower HV values. This can be caused by the fact that the Bézier fitness function will constrain a solution when one of its control points is dominated in objective space by one of the test points. As the endpoints of each part of the discontinuous PF are close to being dominated, i.e., close to the constraint space, it can lead to not entirely capturing the discontinuous pieces of the PF and thus resulting in a lower HV.

8.2 Indicator Limitations

Optimal μ-distribution The use of the HV indicator to score individuals in MM-BezEA and Set-BezEA is useful in the sense that it is a Pareto compliant indicator [25, 79], but it does suffer from a downside. In some situations the endpoints of the approximation sets cannot reach the endpoints of the Pareto set [4]. Even in cases where the number of test points will be set to infinity, the reference point can never be set so that the extremes are captured. This issue comes from the distributions of the test solutions along an approximation front, in literature also called μ-distributions [4], in certain cases where objective functions have asymptotic limits. Then, the optimal distribution of test solutions that maximizes the hypervolume will never include the endpoints of a Pareto set as the hypervolume contribution will become zero at the extremes.

Smoothness Indicator Even though the smoothness indicator tries to determine whether an approximation set is smooth by measuring the detour length, it comes down to determining the angle between solutions. This implies that it only considers linear curves to be perfectly smooth, i.e., when there is a straight angle between a solution and its adjacent solutions in decision space. If the approximation sets increase in size, more solutions will be spread over an equal range in

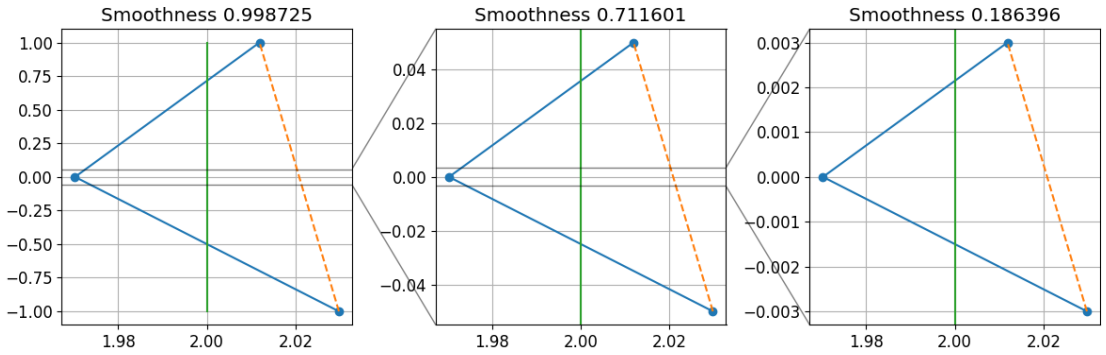


Figure 8.1: Plot showing the issue with the used Smoothness indicator. From left to right, plots have smaller y-axis range. Approximation set in blue, with the Pareto set in green.

decision space with smaller distances in between. Then, the angles between the solutions will become smaller if an approximation set oscillates slightly around a PS, which results in a lower smoothness. Figure 8.1 illustrates this issue in three plots with decreasing y axis ranges from left to right, similar to having more solutions in the approximation set as a result of which the distance between the solutions reduces. In cases where the niches can be separated in a good manner, another definition of smoothness that considers the oscillation around the PS might be more useful.

8.3 Future Work

Even though the three research questions that were posed in this thesis have successfully been answered, many more questions and ideas arose for future work. In this section these questions and ideas are shown in a few subsections that subdivides them into several key areas.

8.3.1 Algorithm Hyperparameters

Set-BezEA and MM-BezEA both have been run with basic parameter settings, which could be further tuned to achieve better results.

Population Size. The influence of the population size on MM-BezEA and Set-BezEA has not been thoroughly tested. As the initial population size determines the number of modes MM-BezEA can search through, it could influence the number of restarts required for MM-BezEA to find all approximation sets. In the case of Set-BezEA, especially when tasked with finding more approximation sets, the number of parameters the algorithm has to optimize grows quickly. With more parameters, a higher population size is sometimes required in order to be able to obtain high quality approximation sets. Thus, future work should look into the optimal population size instead of the currently used default of 76, equal to the original population size in BezEA [72].

Niching Radius. The population size, in the case of Set-BezEA, also influences the expected edge length that is used to determine the minimal separation criterion for the fitness sharing. For this reason, future work that investigates further upon the population size could conclude that another solution might be required for the niche radius. Another solution to this could for example be the Bézier Hill-Valley test, as niche radii are difficult to define properly [15]. This test could support partial variations in order to minimize its impact on the number of evaluations as the pairwise combinations between each of the b Bézier curves are known. Each of the pairwise tests can be stored and only some of them are partially updated when a partial evaluation on one of the curves is performed.

8.3.2 HDR Brachytherapy

The current bi-objective problem formulation is capable of producing treatment plans that address all of the clinical aims as originally set in the Amsterdam UMC. However, these aims do not capture all of the preferred characteristics of a treatment plan and room was left for undesirable behaviour like too much radiation outside of the target volumes and high dissimilarity of dwell times of adjacent dwell positions within a catheter. Parallel to the research performed in this thesis, several works have been published that try to address these issues [14, 21].

Normal Tissue Constraints Normal tissue constraints [21] have been implemented in a bi-objective formulation with an adaptive objective configuration for HDR brachytherapy for cervix cancer. These normal tissue constraints take any healthy tissue outside of the target volumes and OARs and measure the radiation in those locations. New aims were added to the sparing group of the existing clinical protocol for HDR brachytherapy for cervix cancer in order to keep the radiation in the normal tissue below acceptable levels.

Dwell Times Dissimilarity The large dissimilarity in dwell times for adjacent dwell positions has also been researched parallel to this thesis [14]. Several techniques like Dwell Time Modulation Restriction, Dwell Length Duration Modulation and Hotspot Size Index (HSI) were researched. All of these measures are found to improve the homogeneity of treatment plans, with the HSI showing the most promising results. These solutions all add a third objective, thus making it even more relevant to work towards expanding the capabilities of Bézier parameterizations to three objectives.

GPU Parallelization BRIGHT, which is currently used in the clinical practice, makes use of GPU parallelization. Set-BezEA and MM-BezEA currently do not make use of parallelization, but they do have the capacity for it. For example, parallelization techniques could be used to more efficiently calculate the dose distribution for the q control points. This would be similar to the parallelization technique employed for BRIGHT [10, 11]. More specific for Bézier curve parameterizations in the case of HDR brachytherapy, equation 2.10 could be parallelized using the GPU in order to gain more performance in the evaluation of each of the p test solutions. Combining both of these parallelization techniques could significantly speed up the evaluation of a Bézier parameterized solution set, reducing the currently required computation time of up to 24 hours on a single CPU to possibly a computation time to under a quarter of an hour, whereas BRIGHT currently uses up to 5 minutes.

8.3.3 Bézier Curve Parameterizations

The performance and behaviour of MM-BezEA and Set-BezEA has only been investigated on bi-objective problems, whereas enabling them to optimize for more objectives would be preferred. Furthermore, higher order Bézier curves still have some limitations that can be alleviated to enhance the performance when optimizing for non-linear Pareto sets. After these changes, further experiments can be conducted to more thoroughly investigate the performance on optimization problems with more objectives and non-linear Pareto sets.

Higher Order Bézier Curves. Future work could investigate the use of the Bézier parameterizations with more control points to approximate non-linear Pareto sets. Only the test points of the Bézier solution set $S_{p,q}(C_q)$ have to be within the bounds of the search space. Therefore, future work could research the option of having intermediary control points outside of the feasible region, or possibly even outside of the search space, as long as the test points remain within the feasible region. This could allow higher order Bézier curves to be used in order to approximate non-linear Pareto sets which skim the edges of the feasible region.

Figure 8.2 illustrates the issue for $q \in \{3, 4\}$ control points. Both sub figures have control points that are at the limits of the search space with domain $[0, 1]^2$. However, in the example for 3 control points, the curve starts at the lower bound of the y-axis in the search space and then only goes up to 0.5 before going down to the lower bound at the other extreme of the x-axis. In the case of 4 control points, the curve does go up further by reaching 0.75, but still falls short of the edges of the search space. If control points could be drawn outside of the search space, as

long as the test points remain within it, this could help better parameterize approximations for non-linear Pareto sets.

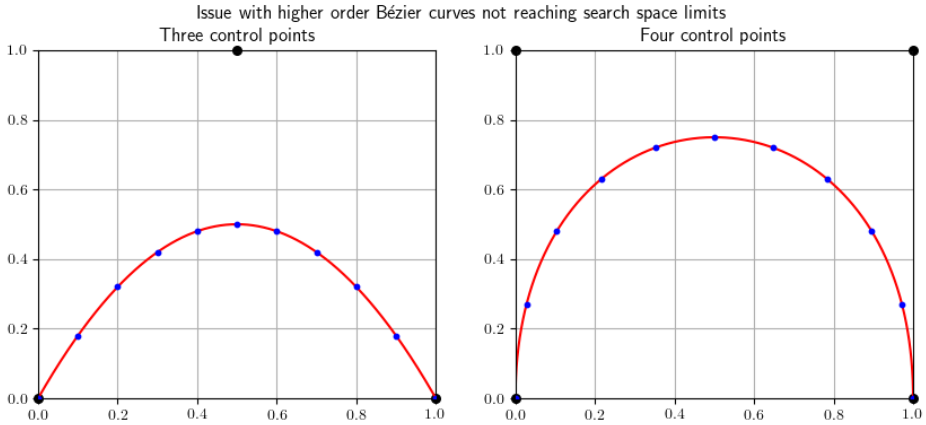


Figure 8.2: Plot showing the issue with higher order Bézier curves not reaching search space limits. Black is control points, red the curve and blue the test points.

Three Objectives. Bézier curves can be expanded upon in the form of Bézier surfaces [54] or simplexes [34]. These might be usable for problems with more than two objectives. Figure 8.3 shows an example of a Bézier surface, parameterized by a 4x4 grid of control points. Using these Bézier parameterizations in 3 dimensions, tri-objective optimization problems can be tackled by encoding Bézier solution set $S_{p,q}(C_q)$ as planes in three dimensional space instead of lines in two dimensional space. This could open the way for applying Bézier parameterizations to more elaborate optimization problems, e.g. more recent versions of the HDR brachytherapy problem definitions that use a third objective for various reasons. GPU parallelization for the calculation of the points on a Bézier surface has also been researched to allow for fast interpolation and evaluation of Bézier surfaces [54].

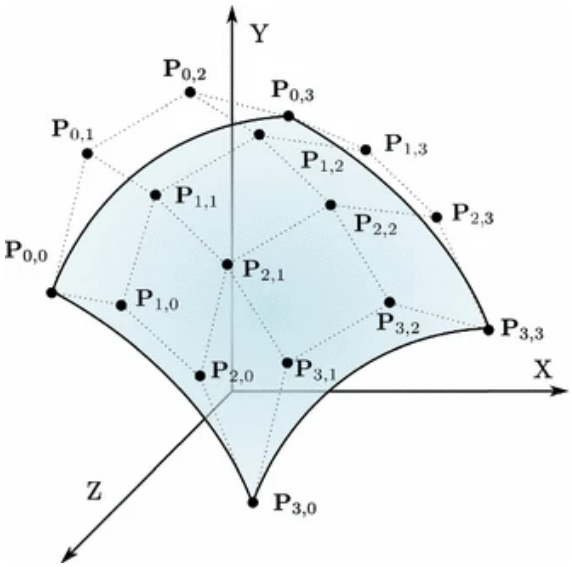


Figure 8.3: Visualization of Bézier surface that can be employed in three objective problems [54].

Chapter 9

Conclusion

In this thesis two algorithms have been proposed, MM-BezEA and Set-BezEA. Both of these algorithms search for Bézier curve parameterizations of approximation sets that define smooth curves in the decision space for bi-objective multi-modal optimization problems. In (real-world) multi-modal problems, good navigability of approximation sets is desired by decision makers. In highly multi-modal problem, the navigability can be further improved if approximation sets for a specific number of modes can be requested as inspecting all approximation sets then becomes intractable.

The empirical results show that both algorithms were outperformed by MO-HillVallEA and MO_Ring_PSO_SCD in problems with non-linear Pareto sets, since if the non-linearity of a mode surpasses the Bézier order, it can no longer be captured by the Bézier curve. However, only low-order Bézier curves were researched. The proposed future research for higher order Bézier curves could potentially improve the performance on non-linear Pareto sets, which has been partially demonstrated in the experiment for the HDR brachytherapy problem.

MM-BezEA has shown good performance in bi-objective problems of moderate multi-modality through its linear search of all available modes. However, on highly multi-modal problems, MM-BezEA deteriorates in its performance with lower obtained hypervolume values as the population has to be divided over each found mode. On the other hand, the encoding of b curves in Set-BezEA produces an exact number of approximation sets that are of high quality for highly multi-modal problems. Both algorithms produce smoothly navigable approximation sets, as the results demonstrated that the smoothness is indeed enforced by the Bézier parameterizations.

On the topic of HDR brachytherapy, MM-BezEA was found to be incapable of optimizing multiple approximation sets for highly multi-modal problems. Set-BezEA on the other hand has shown very promising results in its capabilities to find high-quality approximation sets. Chapter 7 shows that even though the BRIGHT algorithm has produced slightly better results in terms of hypervolume, the clinicians did not see a clear difference between the treatment plans originating from either BRIGHT or Set-BezEA.

All together, it can be concluded that the original goal of this thesis has successfully been achieved in the form of the Set-BezEA algorithm. The clinicians required a way to smoothly navigate approximation sets in order to be able to choose the best plan for a patient. The dissimilarity of approximation sets also had to be enforced in order to allow clinicians to choose from the diversity of treatment plans from the original BRIGHT algorithm. Set-BezEA is found to be capable of obtaining a predefined number of approximation sets in highly multi-modal bi-objective optimization problems and compete with state-of-the-art results from the BRIGHT algorithm. The smooth navigability that is desired by decision makers is inherently caused by the Bézier parameterization, whereas the second objective in the new fitness function formulation ensures distinct approximation sets in order to preserve the diversity.

After successfully answering the research questions posed in this thesis, more research questions have arisen. Future work could: improve Bézier parameterizations by enabling the optimization of a third objective; improve performance through the use of GPU parallelization; and allow more freedom in choosing control points. It would be exciting to see what the capabilities of Bézier inspired optimization algorithms will be in the future, as both MM-BezEA and Set-BezEA not only have been shown to produce very good results, but also have great potential for further improvements.

Bibliography

[1] A. Ahrari, K. Deb, and M. Preuss, "Multimodal optimization by covariance matrix self-adaptation evolution strategy with repelling subpopulations," *Evol. Comput.*, vol. 25, no. 3, p. 439–471, sep 2017. [Online]. Available: https://doi.org/10.1162/evco_a_00182

[2] T. Alderliesten, P. A. N. Bosman, and A. Bel, "Getting the most out of additional guidance information in deformable image registration by leveraging multi-objective optimization," in *Medical Imaging 2015: Image Processing*, S. Ourselin and M. A. Styner, Eds., vol. 9413, International Society for Optics and Photonics. SPIE, 2015, pp. 469 – 475. [Online]. Available: https://doi.org/10.1117/12.2081438

[3] A. Auger and N. Hansen, "A restart CMA evolution strategy with increasing population size," in *2005 IEEE Congress on Evolutionary Computation*, vol. 2, 2005, pp. 1769–1776 Vol. 2.

[4] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler, "Theory of the hypervolume indicator: Optimal μ-distributions and the choice of the reference point," in *Proceedings of the Tenth ACM SIGEVO Workshop on Foundations of Genetic Algorithms*, ser. FOGA '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 87–102. [Online]. Available: https://doi.org/10.1145/1527125.1527138

[5] D. Barten, A. Bouter, N. van Wieringen, B. Pieters, K. Hinnen, G. Westerveld, S. Maree, M. van der Meer, T. Alderliesten, P. Bosman, Y. Niatsetski, and A. Bel, "Artificial intelligence based planning of hdr prostate brachytherapy: first clinical experience." *Radiotherapy and Oncology*, vol. 161, pp. S653–S655, 8 2021.

[6] R. Berghammer, T. Friedrich, and F. Neumann, "Convergence of set-based multi-objective optimization, indicators and deteriorative cycles," *Theoretical Computer Science*, vol. 456, pp. 2–17, Oct 2012.

[7] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221706005443

[8] P. A. N. Bosman, "The anticipated mean shift and cluster registration in mixture-based EDAs for multi-objective optimization," in *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 351–358. [Online]. Available: https://doi.org/10.1145/1830483.1830549

[9] P. A. N. Bosman, J. Grahl, and D. Thierens, "Benchmarking parameter-free AMaLGaM on functions with and without noise," *Evolutionary Computation*, vol. 21, pp. 445–469, 9 2013.

[10] A. Bouter, T. Alderliesten, A. Bel, C. Witteveen, and P. A. N. Bosman, "Large-scale parallelization of partial evaluations in evolutionary algorithms for real-world problems," in *GECCO 2018 - Proceedings of the 2018 Genetic and Evolutionary Computation Conference*, Jul. 2018, pp. 1199–1206.

[11] A. Bouter, T. Alderliesten, B. R. Pieters, A. Bel, Y. Niatsetski, and P. A. N. Bosman, "GPU-accelerated bi-objective treatment planning for prostate high-dose-rate brachytherapy," *Medical Physics*, vol. 46, no. 9, pp. 3776–3787, 2019. [Online]. Available: https://doi.org/10.1145/3205455.3205610

[12] A. Bouter, T. Alderliesten, C. Witteveen, and P. A. N. Bosman, "Exploiting linkage information in real-valued optimization with the real-valued gene-pool optimal mixing evolutionary algorithm," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 705–712. [Online]. Available: https://doi.org/10.1145/3071178.3071272

[13] A. Bouter, N. H. Luong, C. Witteveen, T. Alderliesten, and P. A. N. Bosman, "The multi-objective real-valued gene-pool optimal mixing evolutionary algorithm," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 537–544. [Online]. Available: https://doi.org/10.1145/3071178.3071274

[14] J. Commandeur, "Improving the homogeneity of brachytherapy treatment plans generated by bright using a hotspot registration method based on connected component analysis," Master Thesis, Delft University of Technology, 11 2021, mentor P. A. N. Bosman.

[15] P. Darwen and X. Yao, "A dilemma for fitness sharing with a scaling function," in *Proceedings of 1995 IEEE International Conference on Evolutionary Computation*, vol. 1, 1995, pp. 166–.

[16] L. De Boeck, J. Beliën, and W. Egyed, "Dose optimization in high-dose-rate brachytherapy: A literature review of quantitative models from 1990 to 2010," *Operations Research for Health Care*, vol. 3, no. 2, pp. 80–90, 2014, special Issue of the 2012 conference of the EURO working group Operational Research Applied To Health Services (ORAHS). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211692313000428

[17] L. de Castro, *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*, ser. Chapman & Hall/CRC Computer and Information Science Series. Taylor & Francis, 2006. [Online]. Available: https://books.google.nl/books?id=0gT1zMLEDloC

[18] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, 4 2002.

[19] K. Deb, "An efficient constraint handling method for genetic algorithms," in *Computer Methods in Applied Mechanics and Engineering*, 1998, pp. 311–338.

[20] K. Deb and S. Tiwari, "Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1062–1087, Mar 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221706006291

[21] L. R. M. Dickhoff, E. M. Kerkhof, H. H. Deuzeman, C. L. Creutzberg, T. Alderliesten, and P. A. N. Bosman, "Adaptive objective configuration in bi-objective evolutionary optimization for cervical cancer brachytherapy treatment planning," 2022.

[22] A. Dinkla, R. Laarse, E. Kaljouw, B. R. Pieters, K. Koedooder, N. Wieringen, and A. Bel, "A comparison of inverse optimization algorithms for HDR/PDR prostate brachytherapy treatment planning," *Brachytherapy*, vol. 14, 10 2014.

[23] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion," in *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 62–76.

[24] M. Emmerich, A. Deutz, and N. Beume, "Gradient-based/evolutionary relay hybrid for computing Pareto front approximations maximizing the S-metric," in *Hybrid Metaheuristics*, T. Bartz-Beielstein, M. J. Blesa Aguilera, C. Blum, B. Naujoks, A. Roli, G. Rudolph, and M. Sampels, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 140–156.

[25] M. Fleischer, "The measure of Pareto optima applications to multi-objective metaheuristics," in *Evolutionary Multi-Criterion Optimization*, C. M. Fonseca, P. J. Fleming, E. Zitzler, L. Thiele, and K. Deb, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 519–533.

[26] J. Gallier, *Curves and Surfaces in Geometric Modeling: Theory and Algorithms*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999.

[27] D. E. Goldberg, K. Deb, and J. Horn, "Massive multimodality, deception, and genetic algorithms," in *Parallel Problem Solving from Nature 2, PPSN-II, Brussels, Belgium, September 28-30, 1992*, R. Männer and B. Manderick, Eds. Elsevier, 1992, pp. 37–48.

[28] A. P. Guerreiro, C. M. Fonseca, and L. Paquete, "Greedy hypervolume subset selection in low dimensions," *Evolutionary Computation*, vol. 24, pp. 521–544, Sep 2016.

[29] K. Harada, J. Sakuma, and S. Kobayashi, "Local search for multiobjective function optimization: Pareto descent method," in *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 659–666. [Online]. Available: https://doi.org/10.1145/1143997.1144115

[30] G. R. Harik and F. G. Lobo, "A parameter-less genetic algorithm," in *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation - Volume 1*, ser. GECCO'99. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999, p. 258–265.

[31] J. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. University of Michigan Press, 1975.

[32] P. J. Hoskin, A. Colombo, A. Henry, P. Niehoff, T. Paulsen Hellebust, F.-A. Siebert, and G. Kovacs, "GEC/ESTRO recommendations on high dose rate afterloading brachytherapy for localised prostate cancer: An update," *Radiotherapy and Oncology*, vol. 107, no. 3, pp. 325–332, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167814013002004

[33] C. Igel, N. Hansen, and S. Roth, "Covariance matrix adaptation for multi-objective optimization," *Evolutionary Computation*, vol. 15, no. 1, pp. 1–28, 2007.

[34] K. Kobayashi, N. Hamada, A. Sannai, A. Tanaka, K. Bannai, and M. Sugiyama, "Bézier simplex fitting: Describing Pareto fronts of simplicial problems with small samples in multi-objective optimization," in *Proceedings of the 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, the 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*. Palo Alto, CA, USA: AAAI press, Jan 2019, pp. 2304–2313.

[35] I.-K. K. Kolkman-Deurloo, X. G. Deleye, P. P. Jansen, and P. C. Koper, "Anatomy based inverse planning in HDR prostate brachytherapy," *Radiotherapy and Oncology*, vol. 73, no. 1, pp. 73–77, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016781400400369X

[36] T. Kondo and T. Tatsukawa, "Efficient search techniques using adaptive discretization of design variables on real-coded evolutionary computations," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 697–704. [Online]. Available: https://doi.org/10.1145/3205455.3205636

[37] X. Li, M. G. Epitropakis, K. Deb, and A. Engelbrecht, "Seeking multiple solutions: An updated survey on niching methods and their applications," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 4, pp. 518–538, Aug 2017.

[38] J. J. Liang, C. T. Yue, and B. Y. Qu, "Multimodal multi-objective optimization: A preliminary study," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, 2016, pp. 2454–2461.

[39] J. J. Liang, B. Y. Qu, D. W. Gong, and C. T. Yue, "Problem definitions and evaluation criteria for the CEC 2019 special session on multimodal multiobjective optimization," *Computational Intelligence Laboratory, Zhengzhou University*, Nov 2019.

[40] M. S. Litwin and H.-J. Tan, "The Diagnosis and Treatment of Prostate Cancer: A Review," *JAMA*, vol. 317, no. 24, pp. 2532–2542, 06 2017. [Online]. Available: https://doi.org/10.1001/jama.2017.7248

[41] H. N. Luong and P. A. N. Bosman, "Elitist archiving for multi-objective evolutionary algorithms: To adapt or not to adapt," in *Parallel Problem Solving from Nature - PPSN XII*, C. A. C. Coello, V. Cutello, K. Deb, S. Forrest, G. Nicosia, and M. Pavone, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 72–81.

[42] N. H. Luong, T. Alderliesten, B. R. Pieters, A. Bel, Y. Niatsetski, and P. A. N. Bosman, "Fast and insightful bi-objective optimization for prostate cancer treatment planning with high-dose-rate brachytherapy," *Applied Soft Computing*, vol. 84, p. 105681, 11 2019.

[43] N. H. Luong, T. Alderliesten, A. Bel, Y. Niatsetski, and P. A. N. Bosman, "Application and benchmarking of multi-objective evolutionary algorithms on high-dose-rate brachytherapy planning for prostate cancer treatment," *Swarm and Evolutionary Computation*, vol. 40, pp. 37–52, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210650217306685

[44] N. H. Luong, A. Bouter, M. C. van der Meer, Y. Niatsetski, C. Witteveen, A. Bel, T. Alderliesten, and P. A. N. Bosman, "Efficient, effective, and insightful tackling of the high-dose-rate brachytherapy treatment planning problem for prostate cancer using evolutionary multi-objective optimization algorithms," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO '17, ACM. New York, NY, USA: Association for Computing Machinery, 2017, pp. 1372–1379. [Online]. Available: https://doi.org/10.1145/3067695.3082491

[45] N. H. Luong, H. La Poutré, and P. A. N. Bosman, "Multi-objective gene-pool optimal mixing evolutionary algorithms," in *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 357–364. [Online]. Available: https://doi.org/10.1145/2576768.2598261

[46] S. W. Mahfoud, "Niching methods for genetic algorithms," Ph.D. dissertation, University of Illinois at Urbana-Champaign, USA, 1996, uMI Order No. GAX95-43663.

[47] S. C. Maree, T. Alderliesten, and P. A. N. Bosman, "Real-valued evolutionary multi-modal multi-objective optimization by hill-valley clustering," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 568–576. [Online]. Available: https://doi.org/10.1145/3321707.3321759

[48] S. C. Maree, T. Alderliesten, D. Thierens, and P. A. N. Bosman, "Real-valued evolutionary multi-modal optimization driven by hill-valley clustering," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 857–864. [Online]. Available: https://doi.org/10.1145/3205455.3205477

[49] S. C. Maree, T. Alderliesten, and P. A. N. Bosman, "Benchmarking HillVallEA for the GECCO 2019 competition on multimodal optimization," *Computing Research Repository*, vol. abs/1907.10988, 2019. [Online]. Available: http://arxiv.org/abs/1907.10988

[50] S. C. Maree, N. H. Luong, E. S. Kooreman, N. van Wieringen, A. Bel, K. A. Hinnen, H. Westerveld, B. R. Pieters, P. A. N. Bosman, and T. Alderliesten, "Evaluation of bi-objective treatment planning for high-dose-rate prostate brachytherapy—A retrospective observer study," *Brachytherapy*, vol. 18, pp. 396–403, 5 2019.

[51] G. C. Morton, R. Sankreacha, P. Halina, and A. Loblaw, "A comparison of anatomy-based inverse planning with simulated annealing and graphical optimization for high-dose-rate prostate brachytherapy," *Brachytherapy*, vol. 7, no. 1, pp. 12–16, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1538472107002528

[52] R. Nath, L. L. Anderson, G. Luxton, K. A. Weaver, J. F. Williamson, and A. S. Meigooni, "Dosimetry of interstitial brachytherapy sources: Recommendations of the AAPM radiation therapy committee task group no. 43," *Medical Physics*, vol. 22, no. 2, pp. 209–234, 1995. [Online]. Available: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.597458

[53] M. Nicolini, "A two-level evolutionary approach to multi-criterion optimization of water supply systems," in *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 736–751.

[54] R. Palomar, J. Gómez-Luna, F. A. Cheikh, J. Olivares-Bueno, and O. J. Elle, "High-performance computation of bézier surfaces on parallel and heterogeneous platforms," *International Journal of Parallel Programming*, vol. 46, pp. 1035–1062, 12 2018.

[55] M. Pelikan, M. W. Hauschild, and D. Thierens, "Pairwise and problem-specific distance metrics in the linkage tree genetic algorithm," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 1005–1012. [Online]. Available: https://doi.org/10.1145/2001576.2001713

[56] M. Pelikan, K. Sastry, and D. E. Goldberg, "Multiobjective HBOA, clustering, and scalability," in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 663–670. [Online]. Available: https://doi.org/10.1145/1068009.1068122

[57] M. Preuss, "Niching the CMA-ES via nearest-better clustering," in *Proceedings of the 12th Annual Conference Companion on Genetic and Evolutionary Computation*, ser. GECCO '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 1711–1718. [Online]. Available: https://doi.org/10.1145/1830761.1830793

[58] M. Preuss, B. Naujoks, and G. Rudolph, "Pareto set and EMOA behavior for simple multimodal multiobjective functions," in *Parallel Problem Solving from Nature – PPSN IX*, ser. PPSN'06. Berlin, Heidelberg: Springer-Verlag, 2006, p. 513–522. [Online]. Available: https://doi.org/10.1007/11844297_52

[59] M. W. Przewozniczek and M. M. Komarnicki, "Empirical linkage learning," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, ser. GECCO '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 23–24. [Online]. Available: https://doi.org/10.1145/3377929.3397771

[60] M. J. Rivard, B. M. Coursey, L. A. DeWerd, W. F. Hanson, M. Saiful Huq, G. S. Ibbott, M. G. Mitch, R. Nath, and J. F. Williamson, "Update of AAPM task group no. 43 report: A revised AAPM protocol for brachytherapy dose calculations," *Medical Physics*, vol. 31, no. 3, pp. 633–674, 2004. [Online]. Available: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.1646040

[61] G. Rudolph, B. Naujoks, and M. Preuss, "Capabilities of EMOA to detect and preserve equivalent pareto subsets," in *Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007, Matsushima, Japan, March 5-8, 2007, Proceedings*, ser. Lecture Notes in Computer Science, vol. 4403. Berlin, Heidelberg: Springer, 2006, pp. 36–50. [Online]. Available: https://doi.org/10.1007/978-3-540-70928-2_7

[62] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209–249, Feb. 2021. [Online]. Available: https://doi.org/10.3322/caac.21660

[63] D. Thierens, "The linkage tree genetic algorithm," in *Proceedings of Parallel Problem Solving from Nature, PPSN XI*, ser. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, Sep 2010, pp. 264–273.

[64] D. Thierens and P. A. N. Bosman, "Optimal mixing evolutionary algorithms," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 617–624. [Online]. Available: https://doi.org/10.1145/2001576.2001661

[65] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization," *IEEE Computational Intelligence Magazine*, vol. 12, no. 4, pp. 73–87, Nov 2017.

[66] C. Touré, N. Hansen, A. Auger, and D. Brockhoff, "Uncrowded hypervolume improvement: COMO-CMA-ES and the sofomore framework," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 638–646. [Online]. Available: https://doi.org/10.1145/3321707.3321852

[67] R. Ursem, "Multinational evolutionary algorithms," in *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, vol. 3, 1999, pp. 1633–1640 Vol. 3.

[68] J. Venselaar, D. Baltas, A. Meigooni, P. Hoskin, and B. Thomadsen, "Comprehensive brachytherapy: Physical and clinical aspects." *Medical physics*, vol. 40, p. 117302, 11 2013.

[69] S. C. Maree, T. Alderliesten, and P. A. N. Bosman, "Uncrowded Hypervolume-based Multi-objective Optimization with Gene-pool Optimal Mixing," *Evolutionary Computation*, pp. 1–24, Dec 2021. [Online]. Available: https://doi.org/10.1162/evco_a_00303

[70] P. A. N. Bosman, "Lecture notes," in *CS4205 - Evolutionary Algorithms*. Delft, The Netherlands: Technical University of Delft, Apr 2019.

[71] Kalyanmoy Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. USA: John Wiley & Sons, Inc., 2001.

[72] S. C. Maree, T. Alderliesten, and P. A. N. Bosman, "Ensuring smoothly navigable approximation sets by Bézier curve parameterizations in evolutionary bi-objective optimization," in *Parallel Problem Solving from Nature – PPSN XVI*. Cham: Springer International Publishing, 2020, pp. 215–228.

[73] J. Yang, "Oncentra brachytherapy planning system," *Medical Dosimetry*, vol. 43, no. 2, pp. 141–149, 2018. [Online]. Available: https://doi.org/10.1016/j.meddos.2018.02.011

[74] C. Yue, B. Qu, and J. Liang, "A multi-objective particle swarm optimizer using ring topology for solving multimodal multi-objective problems," *IEEE Transactions on Evolutionary Computation*, vol. 22, pp. 805–817, Sep 2017.

[75] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, pp. 712–731, 12 2007.

[76] E. Zitzler, D. Brockhoff, and L. Thiele, "The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration," in *Evolutionary Multi-Criterion Optimization*, S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, and T. Murata, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 862–876.

[77] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evol. Comput.*, vol. 8, no. 2, p. 173–195, jun 2000. [Online]. Available: https://doi.org/10.1162/106365600568202

[78] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, Nov 1999.

[79] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, May 2003.

Appendix A

Empirical Test Problems

This appendix denotes the various objective functions that are used in the empirical benchmarks of Chapter 6. It also defines the analytical formulas from which the Pareto sets were drawn in order to determine the various indicators. Everything is displayed in the same order of appearance as in Table 6.1.

A.1 Objective Functions

MinDist [47]. For $x_i \in [-4, 4]$:

$$f_1 = \min\{||\boldsymbol{x} - [1.0, 2.0]||, ||\boldsymbol{x} - [-1.0, -2.0]||\}$$
$$f_2 = \min\{||\boldsymbol{x} - [-1.0, 2.0]||, ||\boldsymbol{x} - [1.0, -2.0]||\}$$

OmniTest [20]. For $x_i \in [0, 6]$:

$$f_1 = \sum_{i=1}^{n} \sin \pi x_i$$
$$f_2 = \sum_{i=1}^{n} \cos \pi x_i$$

TwoOnOne [58]. For $x_i \in [-3, 3]$, with $c = 10, d = 0, k = 0, l = 0$:

$$f_1 = x_1^4 + x_2^4 - x_1^2 + x_2^2 - cx_1x_2 + dx_1 + 20$$
$$f_2 = (x_1 - k)^2 + (x_2 - l)^2$$

SymPart 1,2 and 3 [61]. For these problems, for $x_i \in [-20, 20]$, with $a = 1, b = 10, c = 8, \omega = 45°$ and $\epsilon = 0.001$, there are some supporting functions for the objectives:

$$\hat{t}_1(x_1) = sign(x_1) * \left\lfloor \frac{|x_1| - (a + \frac{c}{2})}{2a + c} \right\rfloor$$

$$\hat{t}_2(x_2) = sign(x_2) * \left\lfloor \frac{|x_2| - \frac{b}{2}}{b} \right\rfloor$$

$$\text{Tile identifier } t_i(x_i) = sign(\hat{t}_i(x_i)) * \min\left(|\hat{t}_i(x_i)|, 1\right)$$

$$\text{Rotation } r(\boldsymbol{x}) = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \boldsymbol{x}$$

$$\text{Distortion } d(x_1, x_2) = x_1 * \left(\frac{x_2 + 20 + \epsilon}{40}\right)^{-1}$$

$$\text{Default objective function } g(x_1, x_2) = (x_1 + a)^2 + x_2^2$$

$$\text{Default objective function } h(x_1, x_2) = (x_1 - a)^2 + x_2^2$$

SymPart 1 is then:

$$f_1 = g(x_1 - t_1 * (c + 2a), x_2 - t_2 * b)$$
$$f_2 = h(x_1 - t_1 * (c + 2a), x_2 - t_2 * b)$$

SymPart 2 is then:

$$f_1 = g(r_1(\boldsymbol{x}) - t_1(r_1(\boldsymbol{x})) * (c + 2a), r_2(\boldsymbol{x}) - t_2(r_2(\boldsymbol{x})) * b)$$
$$f_2 = h(r_1(\boldsymbol{x}) - t_1(r_1(\boldsymbol{x})) * (c + 2a), r_2(\boldsymbol{x}) - t_2(r_2(\boldsymbol{x})) * b)$$

SymPart 3 is then:

$$f_1 = g(r_1(d(\boldsymbol{x}), x_2) - t_1(r_1(d(\boldsymbol{x}), x_2)) * (c + 2a), r_2(\boldsymbol{x}) - t_2(r_2(\boldsymbol{x})) * b)$$
$$f_2 = h(r_1(d(\boldsymbol{x}), x_2) - t_1(r_1(d(\boldsymbol{x}), x_2)) * (c + 2a), r_2(\boldsymbol{x}) - t_2(r_2(\boldsymbol{x})) * b)$$

MMF1 [39]. For $x_1 \in [1, 3]$, $x_2 \in [-1, 1]$:

$$f_1 = |x_1 - 2|$$
$$f_2 = 1 - \sqrt{|x_1 - 2|} + 2\left(6\pi|x_1 - 2| + \pi\right)^2$$

MMF2 [39]. For $x_1 \in [0, 1]$, $x_2 \in [0, 2]$:

$$f_1 = x_1$$

$$f_2 = \begin{cases} 1 - \sqrt{x_1} + 2\left(4(x_2 - \sqrt{x_1})^2 - 2\cos\left(\frac{20(x_2 - \sqrt{x_1})\pi}{\sqrt{2}}\right)\right) & 0 \le x_2 \le 1 \\ 1 - \sqrt{x_1} + 2\left(4(x_2 - 1 - \sqrt{x_1})^2 - 2\cos\left(\frac{20(x_2 - 1 - \sqrt{x_1})\pi}{\sqrt{2}}\right)\right) & 0 < x_2 \le 2 \end{cases}$$

MMF12 [39]. For $x_i \in [0,1]$:

$$f_1 = x_1$$
$$f_2 = g(x_2) * h(f_1(x_1), g(x_2))$$

Where:

$$g(x) = 2 - \exp\left[-2\log(2) * \left(\frac{x-0.1}{0.8}\right)^2\right] * \sin(n_p \pi x)$$
$$h(f_1, g) = 1 - \left(\frac{f_1}{g}\right)^2 - \frac{f_1}{g}\sin(2\pi 4 f_1)$$

MMF14 [39]. For $x_i \in [0,1]$:

$$f_1 = \cos(\frac{\pi}{2}x_1) * (1 + g(x_2))$$
$$f_2 = \sin(\frac{\pi}{2}x_1) * (1 + g(x_2))$$

Where:

$$g(x) = 2 - \sin^2(n_p \pi(x))$$

MMF15 [39]. For $x_i \in [0,1]$:

$$f_1 = \cos(\frac{\pi}{2}x_1) * (1 + g(x_2))$$
$$f_2 = \sin(\frac{\pi}{2}x_1) * (1 + g(x_2))$$

Where:

$$g(x) = 2 - \exp\left[-2\log(2)\left(\frac{x-0.1}{0.8}\right)^2\right]\sin^2(n_p \pi x)$$

A.2 Pareto Set Formulas

MinDist [47]:

$$PS_1 : x_1 = -2.0 \;\; x_2 \in [-1, 1]$$
$$PS_2 : x_1 = 2.0 \;\; x_2 \in [-1, 1]$$

OmniTest [20]:

$$PS : x_i \in [2m + 1, 2m + \frac{3}{2}] \text{ for } m \in \{0, 1, 2\}$$

TwoOnOne [58]. Here the linear approximation is given as the analytical formula is currently unavailable. This approximation has a deviation less than 0.045 from the true convex-concave curve.

$$PS : x_1 \in \left[-\frac{1}{2}\sqrt{\sqrt{101}+1}, \frac{1}{2}\sqrt{\sqrt{101}+1}\right]$$
$$x_2 = \frac{\sqrt{101}-1}{10}x_1$$

Sympart 1, 2 and 3 [61]. Starting with $y_1 = \mu - c * \lambda$ with $\mu \in [-1, 1]$ and $\lambda \in \{-1, 0, 1\}$, and $y_2 \in \{-b, 0, b\}$. Then, using the function definitions from the objective formulas, the PSs are as follows.
Sympart 1:

$$PSs : x_1 = y_1, x_2 = y_2$$

Sympart 2:

$$PSs : x_1 = r_1(y_1, y_2), x_2 = r_2(y_1, y_2)$$

Sympart 2:

$$PSs : x_1 = d(r_1(y_1, y_2)), x_2 = r_2(y_1, y_2)$$

MMF1 [39].

$$\begin{aligned} PSs : x_1 \in [1, 2], x_2 &= \sin(6\pi|x_1 - 2| + \pi) \\ x_1 \in [2, 3], x_2 &= \sin(6\pi|x_1 - 2| + \pi) \end{aligned}$$

MMF2 [39].

$$PSs : x_2 \in [0, 1]$$
$$x_1 = \begin{cases} x_2^2 & 0 \leq x_2 \leq 1 \\ (x_2 - 1)^2 & 0 < x_2 \leq 2 \end{cases}$$

MMF12 [39].

$$PSs : x_1 \in [0, 1], x_2 = \frac{1}{2n_p} + \frac{1}{n_p} * (i - 1) \text{ for } i \in \{1, 2\}$$

MMF14 [39].

$$PSs : x_1 \in [0, 1], x_2 = \frac{1}{2n_p} + \frac{1}{n_p} * (i - 1) \text{ for } i \in \{1, 2\}$$

MMF15 [39].

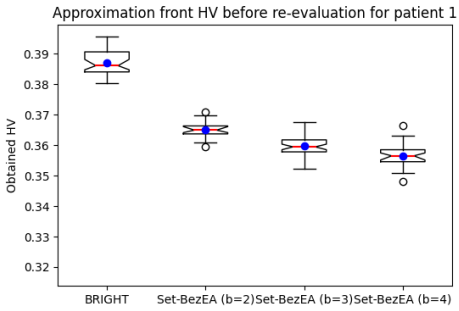
$$\begin{aligned} \text{Global } PS : x_1 \in [0, 1], x_2 &= \frac{1}{2n_p} + \frac{1}{n_p} \\ \text{Local } PS : x_1 \in [0, 1], x_2 &= \frac{1}{2n_p} + \frac{1}{n_p} \end{aligned}$$

Appendix B

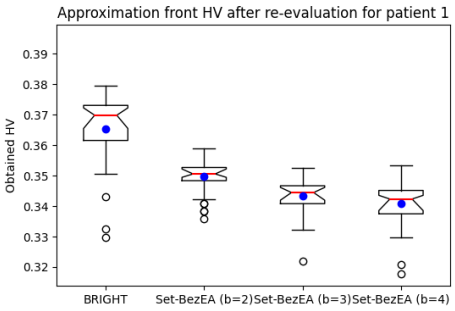
High-Dose-Rate Brachytherapy Results

B.1 Hypervolume Box and Whisker Plots with Outliers

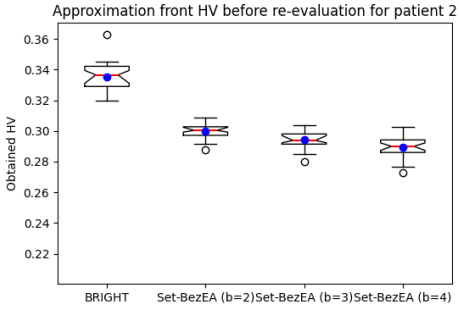
Figures B.1 shows the box and whisker plots for patients 1 to 7 before and after re evaluation in Figures B.1a to B.1n respectively. The plots show the mean as the red dotted line, with the blue point indicating the mean. The boxplots show notches for the 95% confidance interval of the median after bootstrapping the results 10.000 times. Furthermore, these boxes show the Inter Quartile Range (IQR) in which 50% of the results fall, i.e. the 25'th percentile to the 75'th percentile. The whiskers show the minimum and maximum obtained HV, unless a result is more than 1.5 IQR away from the 25'th or 75'th percentile. In that case, all results that fall outside of the whiskers are shown as singular outliers in the form of circles.



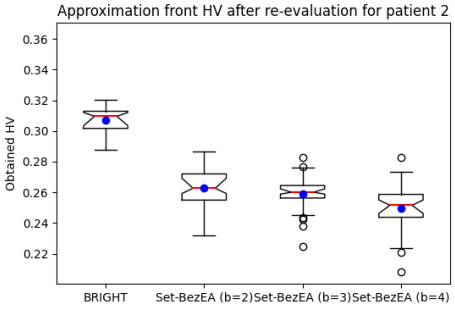
(a) Patient 1 before re-evaluation



(b) Patient 1 after re-evaluation

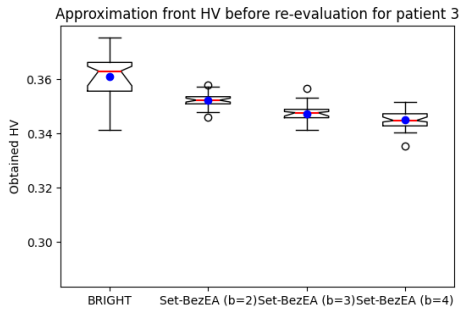


(c) Patient 2 before re-evaluation

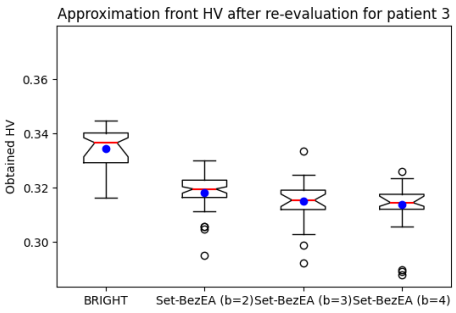


(d) Patient 2 after re-evaluation

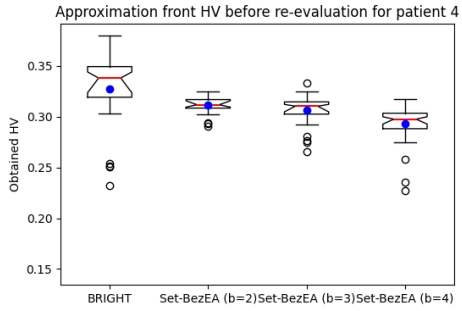
Figure B.1: Obtained hypervolume box and whisker plots over all 10 runs, red is median and blue mean. Notches indicate 95% confidence interval of median after bootstrapping 10,000 times.



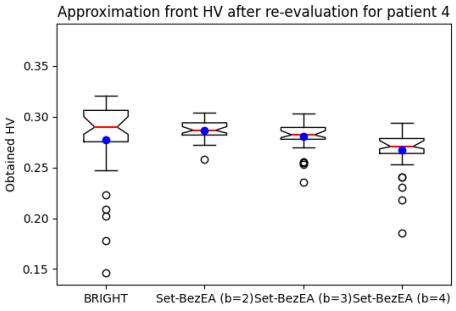
(e) Patient 3 before re-evaluation



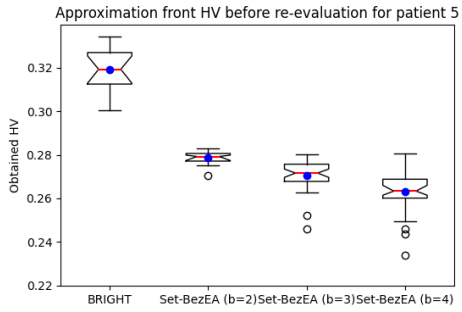
(f) Patient 3 after re-evaluation



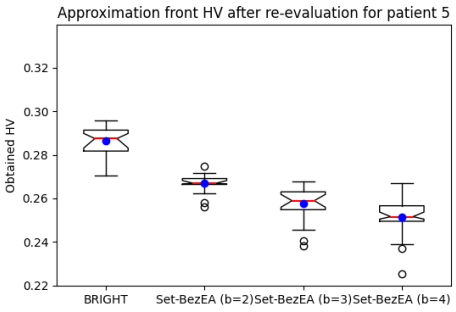
(g) Patient 4 before re-evaluation



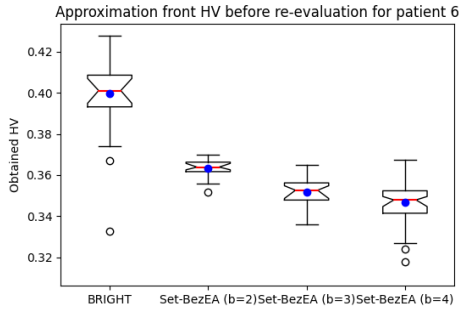
(h) Patient 4 after re-evaluation



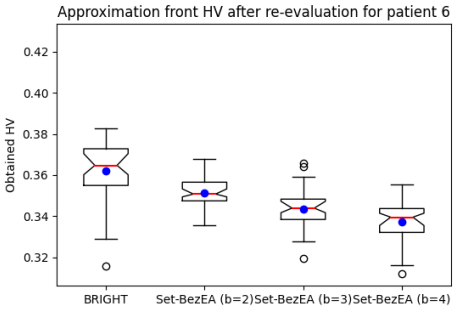
(i) Patient 5 before re-evaluation



(j) Patient 5 after re-evaluation

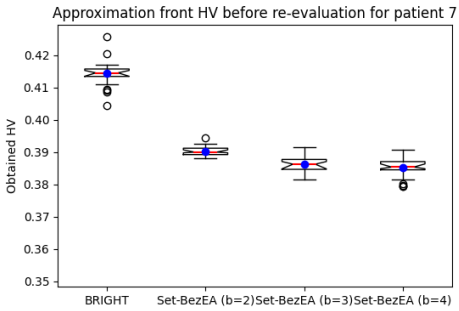


(k) Patient 6 before re-evaluation

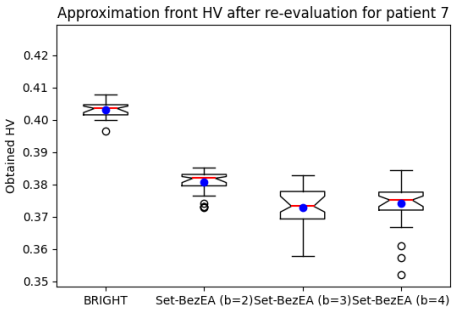


(l) Patient 6 after re-evaluation

Figure B.1: Obtained hypervolume box and whisker plots over all 10 runs, red is median and blue mean. Notches indicate 95% confidence interval of median after bootstrapping 10,000 times.



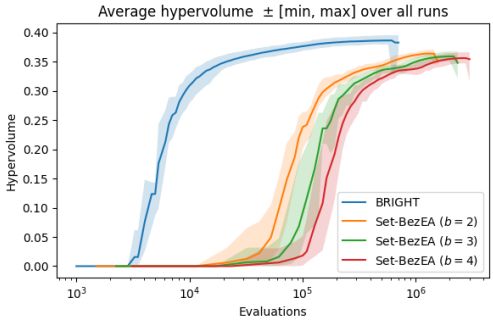
(m) Patient 7 before re-evaluation



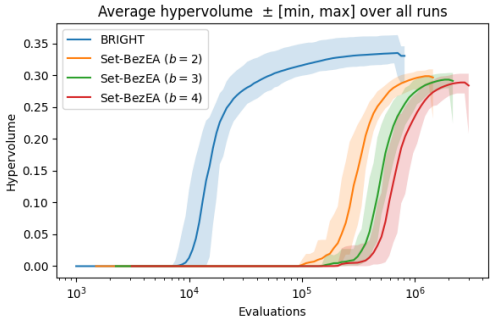
(n) Patient 7 after re-evaluation

Figure B.1: Obtained hypervolume box and whisker plots over all 10 runs, red is median and blue mean. Notches indicate 95% confidence interval of median after bootstrapping 10,000 times.

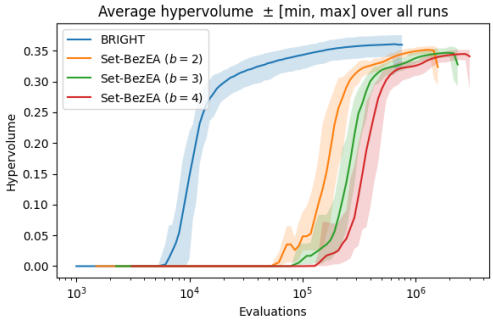
B.2 Hypervolume Performance Comparison



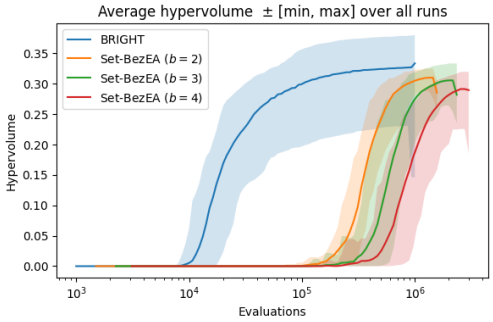
(a) Patient 1



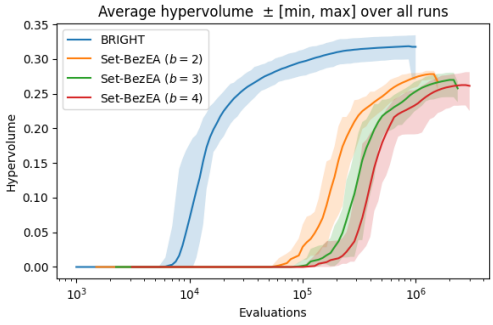
(b) Patient 2



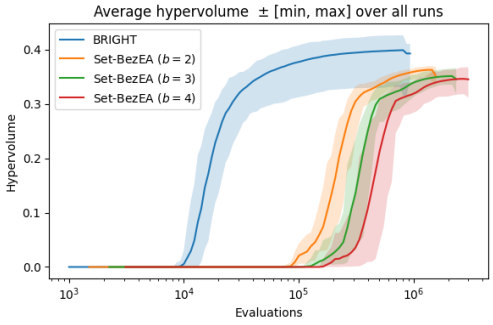
(c) Patient 3



(d) Patient 4

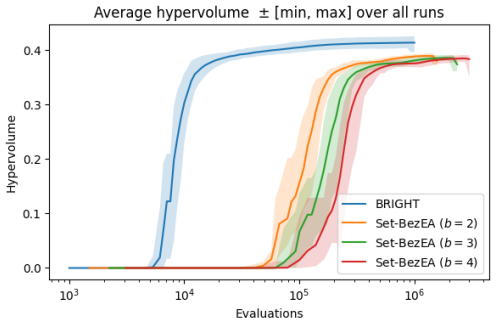


(e) Patient 5



(f) Patient 6

Figure B.2: Hypervolume vs evaluations (performance plots) for all patients.

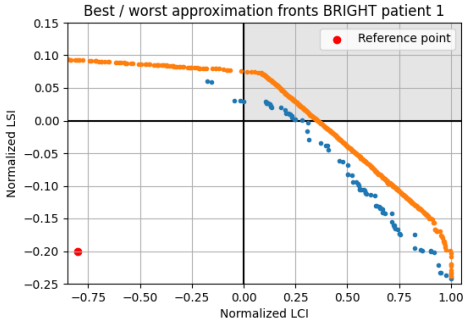


(g) Patient 7

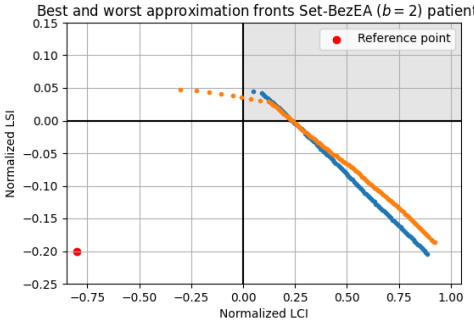
Figure B.2: Hypervolume vs evaluations (performance plots) for all patients.

B.3 Best and Worst Approximation Fronts

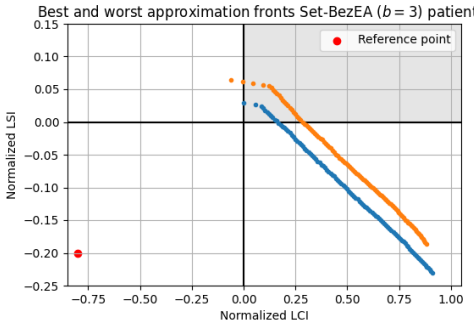
This section shows the worst and best approximation fronts after re-evaluation in terms of normalized LCI vs normalized LSI. Figure B.3 shows the best and worst approximation fronts in terms of obtained hypervolume for patients 1 to 7 respectively. All plots have a similar range in axis to compare the quality of the approximation fronts between patients. Each plot shows the reference point in red, which is used to bound the area below the approximation front for the hypervolume calculation. In grey is the quadrant that has positive values for both the normalized LSI and normalized LCI, indicating that the plans that lie within these ranges have met all clinical aims. For BRIGHT, the fronts extend further outward outside of the shown plots, with plans having normalized LCI values down to -4 and normalized LSI down to -0.6. However, most of these plans that lie outside of the range of the shown plots are usually not of clinical interest since they fall to far outside of the region in which the aims are satisfied.



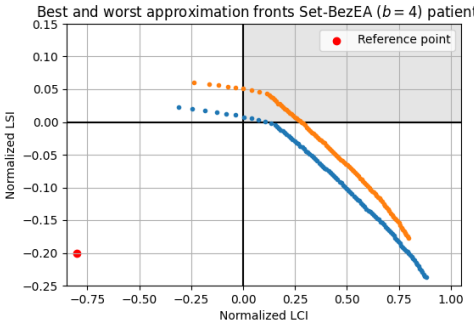
(a) Patient 1 BRIGHT



(b) Patient 1 Set-BezEA ($b = 2$)

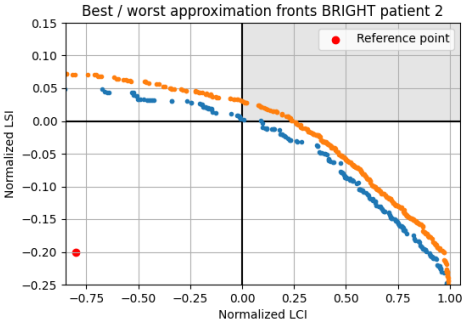


(c) Patient 1 Set-BezEA ($b = 3$)

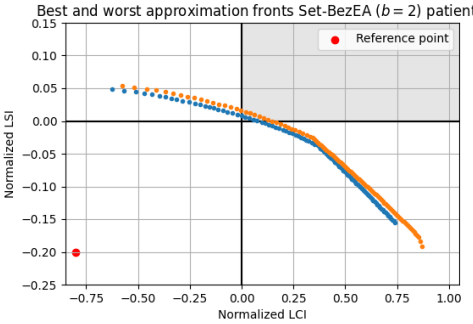


(d) Patient 1 Set-BezEA ($b = 4$)

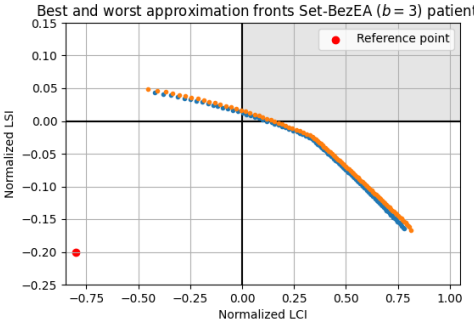
Figure B.3: Best and worst approximation fronts over all 10 runs.



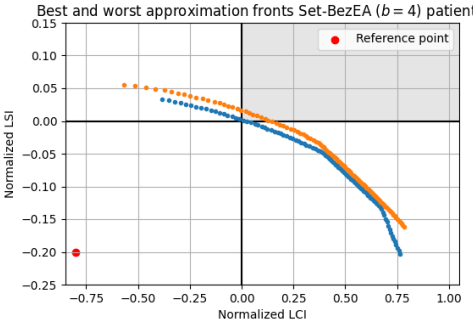
(e) Patient 2 BRIGHT



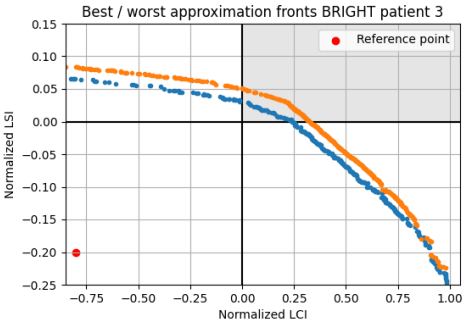
(f) Patient 2 Set-BezEA ($b = 2$)



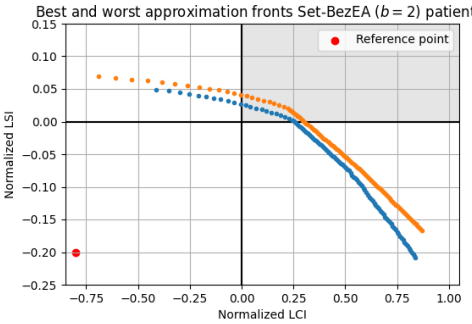
(g) Patient 2 Set-BezEA ($b = 3$)



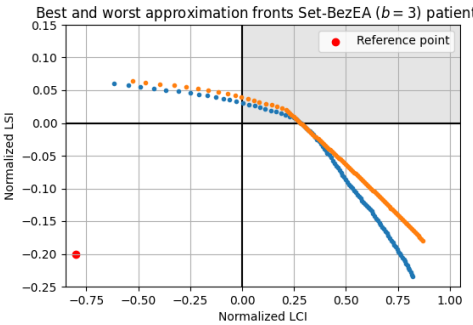
(h) Patient 2 Set-BezEA ($b = 4$)



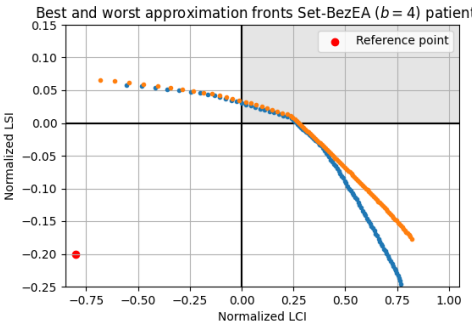
(i) Patient 3 BRIGHT



(j) Patient 3 Set-BezEA ($b = 2$)

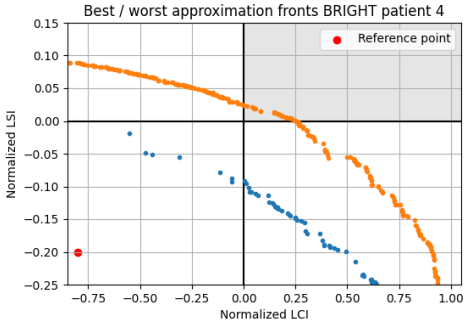


(k) Patient 3 Set-BezEA ($b = 3$)

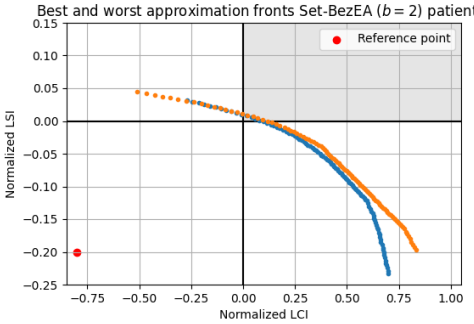


(l) Patient 3 Set-BezEA ($b = 4$)

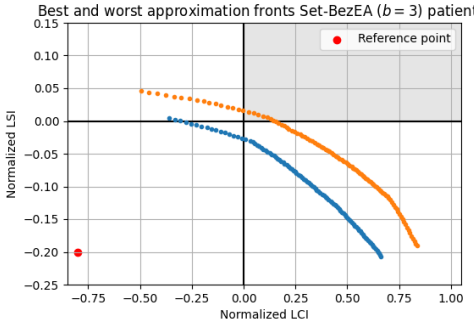
Figure B.3: Best and worst approximation fronts over all 10 runs.



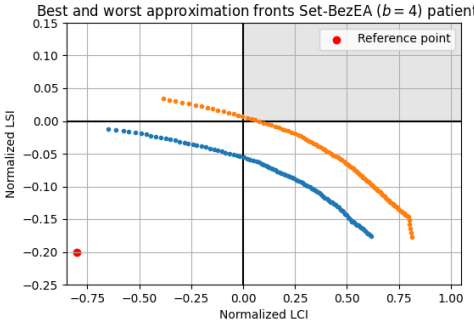
(m) Patient 4 BRIGHT



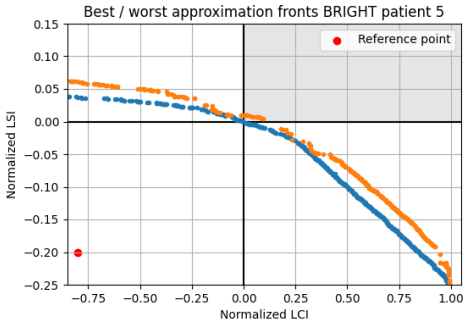
(n) Patient 4 Set-BezEA ($b = 2$)



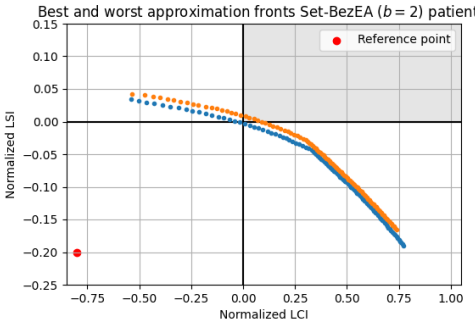
(o) Patient 4 Set-BezEA ($b = 3$)



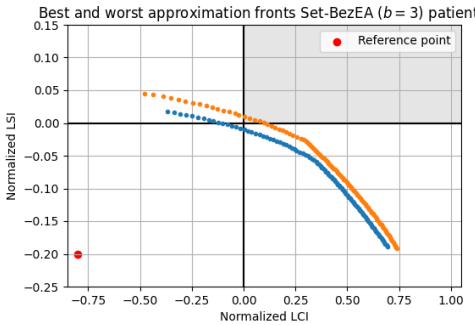
(p) Patient 4 Set-BezEA ($b = 4$)



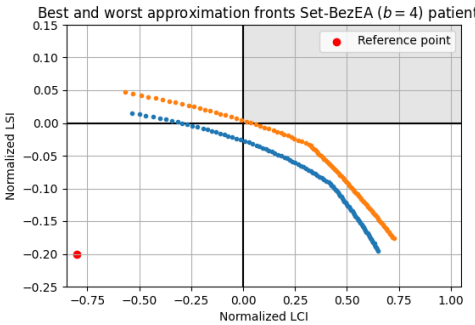
(q) Patient 5 BRIGHT



(r) Patient 5 Set-BezEA ($b = 2$)

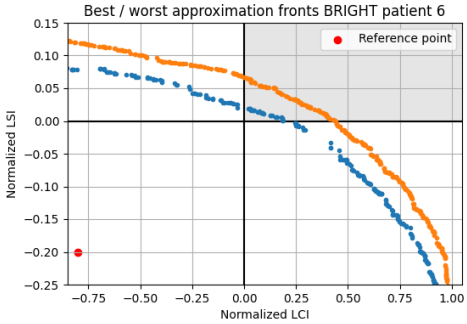


(s) Patient 5 Set-BezEA ($b = 3$)

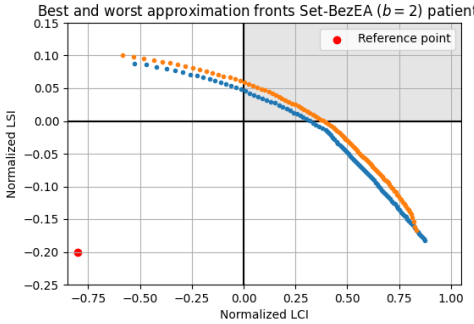


(t) Patient 5 Set-BezEA ($b = 4$)

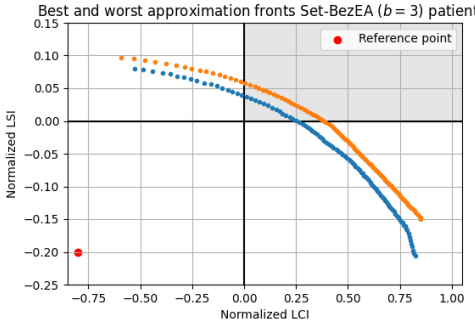
Figure B.3: Best and worst approximation fronts over all 10 runs.



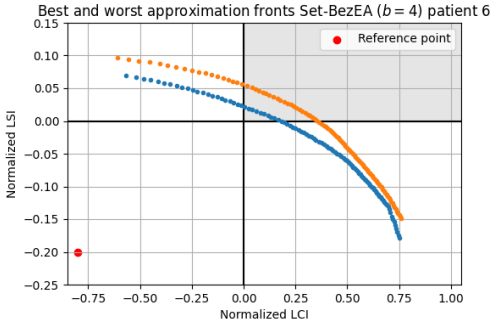
(u) Patient 6 BRIGHT



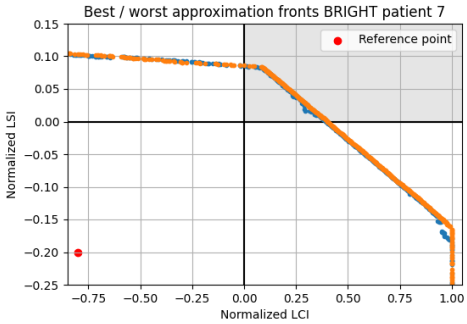
(v) Patient 6 Set-BezEA ($b = 2$)



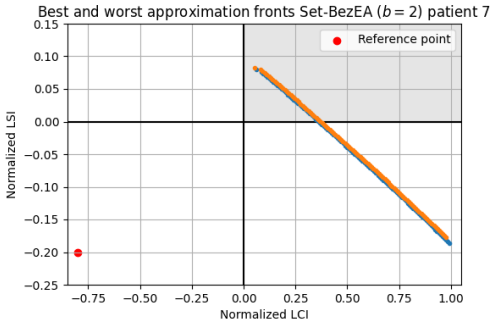
(w) Patient 6 Set-BezEA ($b = 3$)



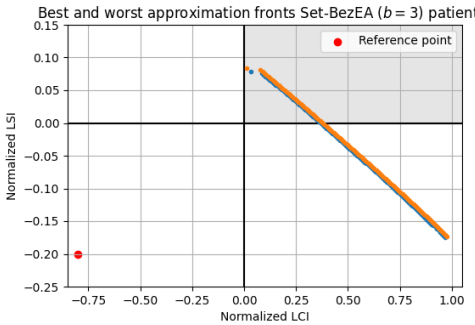
(x) Patient 6 Set-BezEA ($b = 4$)



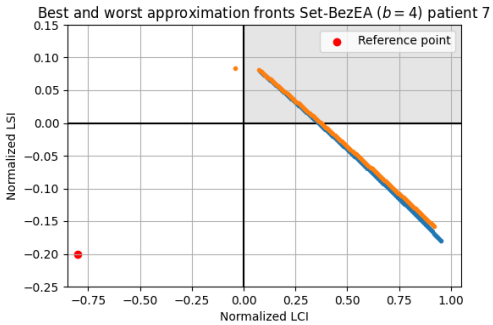
(y) Patient 7 BRIGHT



(z) Patient 7 Set-BezEA ($b = 2$)



(aa) Patient 7 Set-BezEA ($b = 3$)

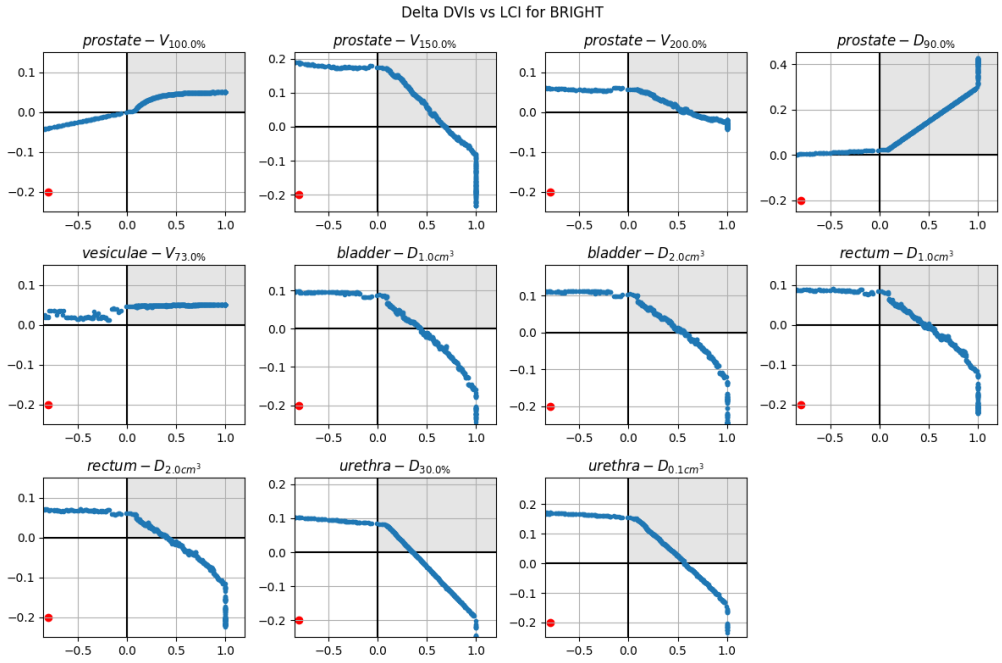


(ab) Patient 7 Set-BezEA ($b = 4$)

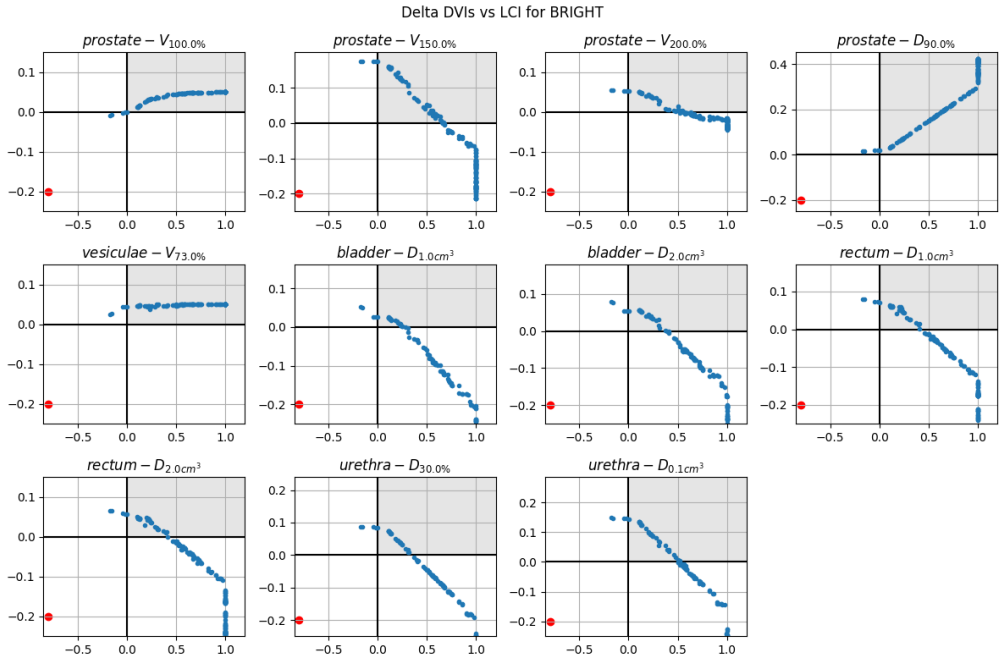
Figure B.3: Best and worst approximation fronts over all 10 runs.

B.4 Delta Dose-Volume Indices

The plots in this section show the Delta DVI values obtained for the best and worst approximation sets for each of the 7 patients. These plots show the difference between the clinical aim and what has been achieved for each DVI on the y axis and the normalized LCI value on the x axis All 11 DVIs are shown, where the observed behaviour for coverage aims is that they are increasing, whereas the sparing aims are decreasing with higher normalized LCI values.

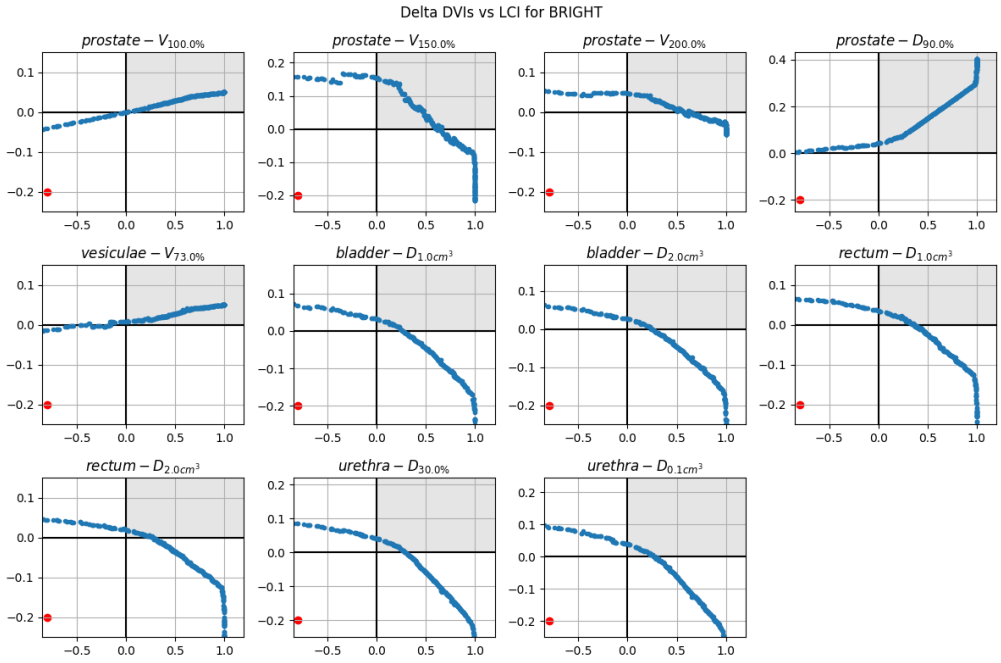


(a) Patient 1 BRIGHT best approximation front delta DVIs

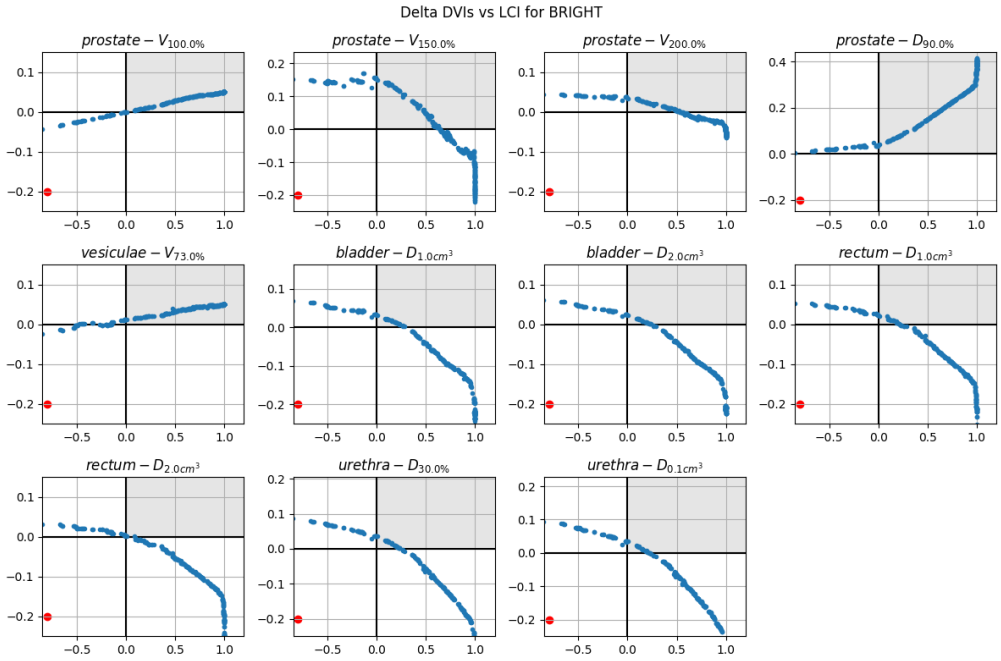


(b) Patient 1 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

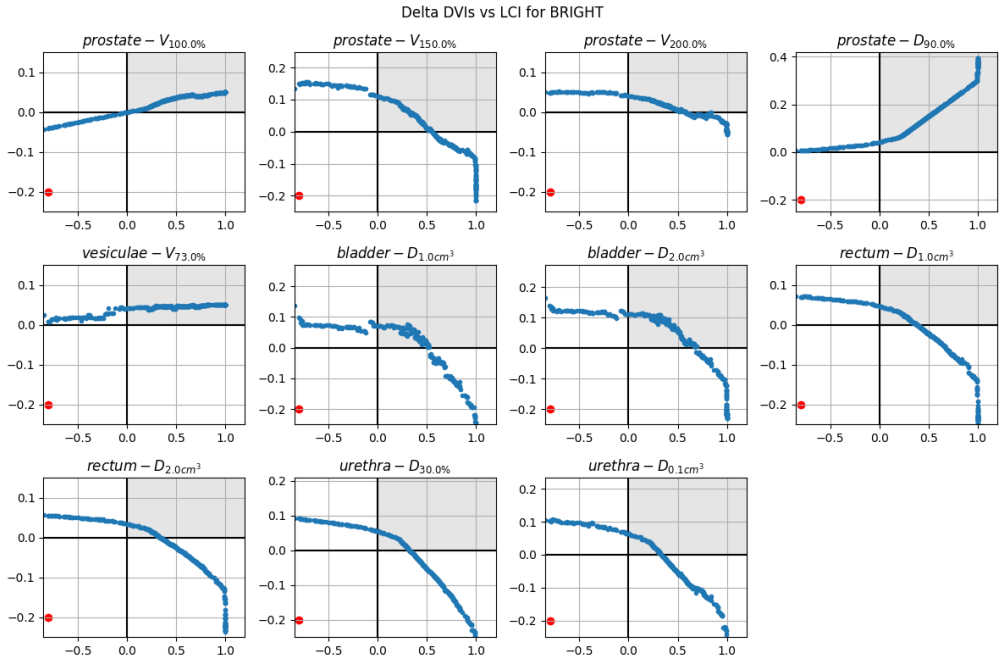


(c) Patient 2 BRIGHT best approximation front delta DVIs

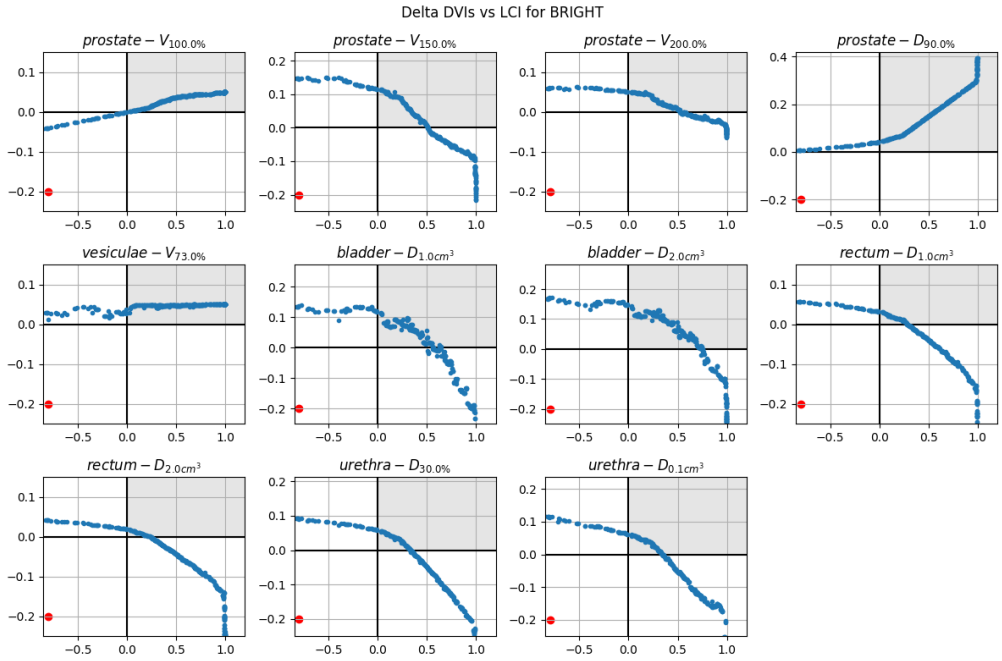


(d) Patient 2 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

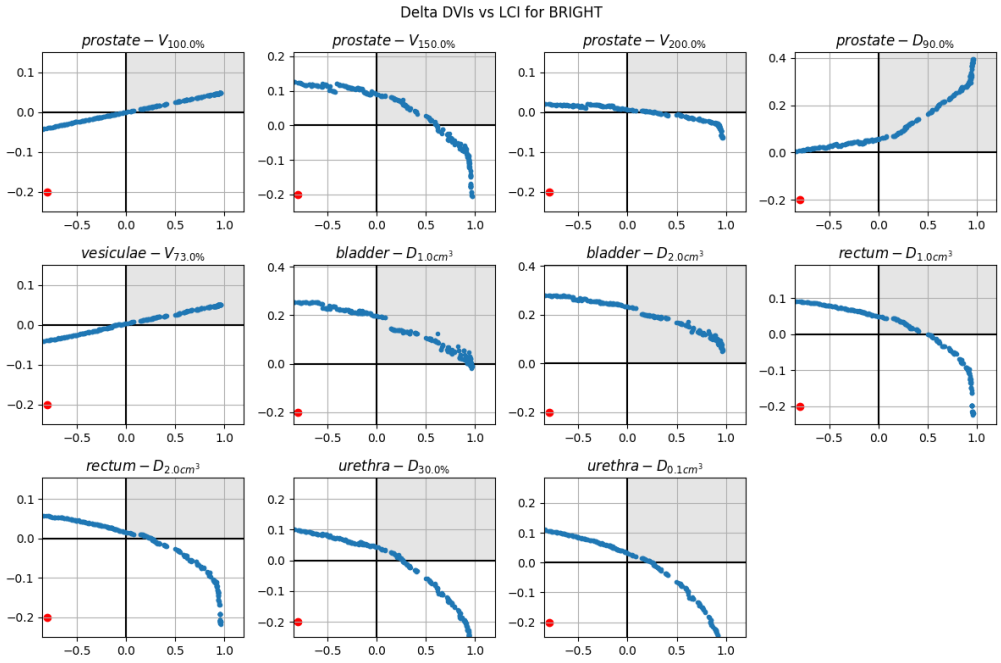


(e) Patient 3 BRIGHT best approximation front delta DVIs

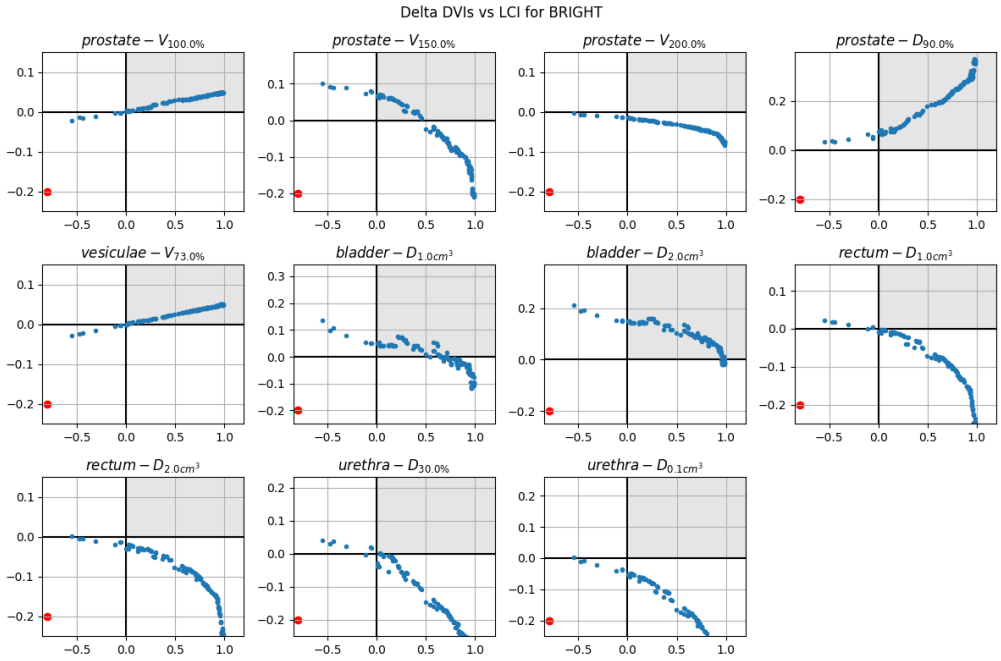


(f) Patient 3 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

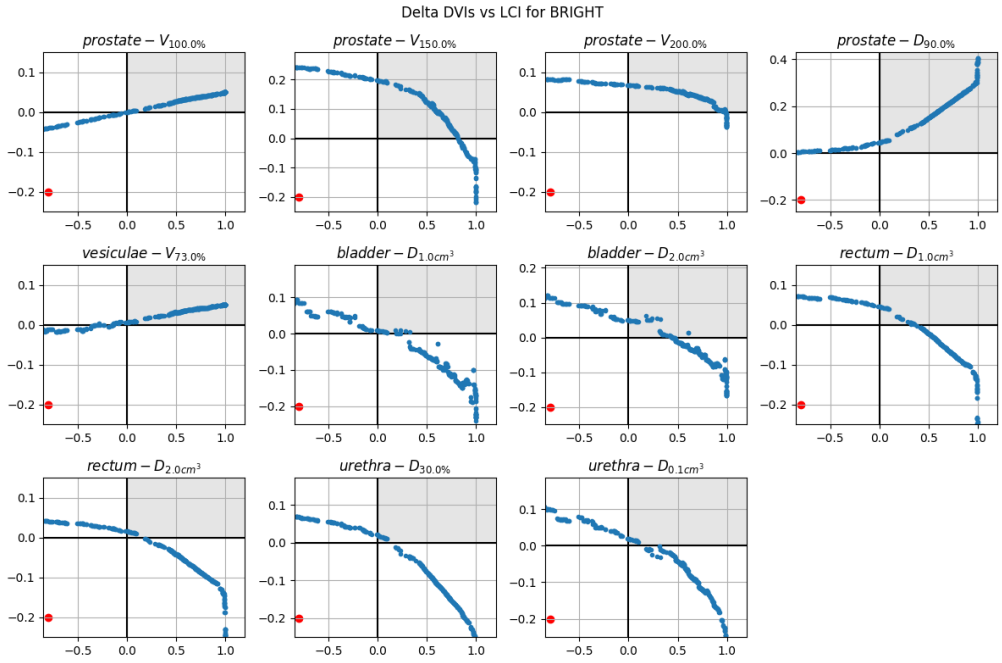


(g) Patient 4 BRIGHT best approximation front delta DVIs

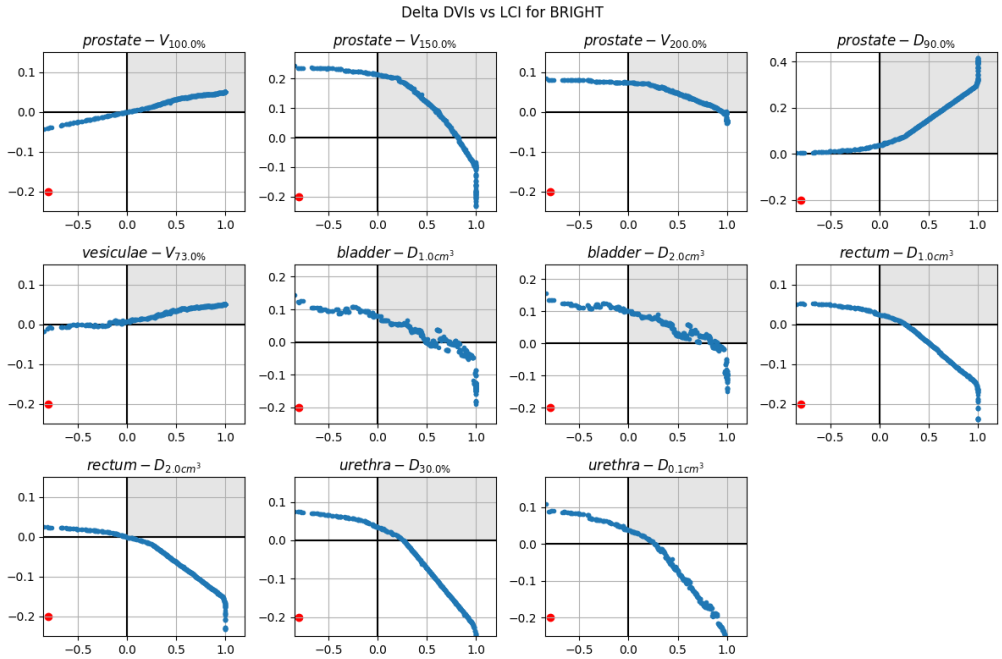


(h) Patient 4 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

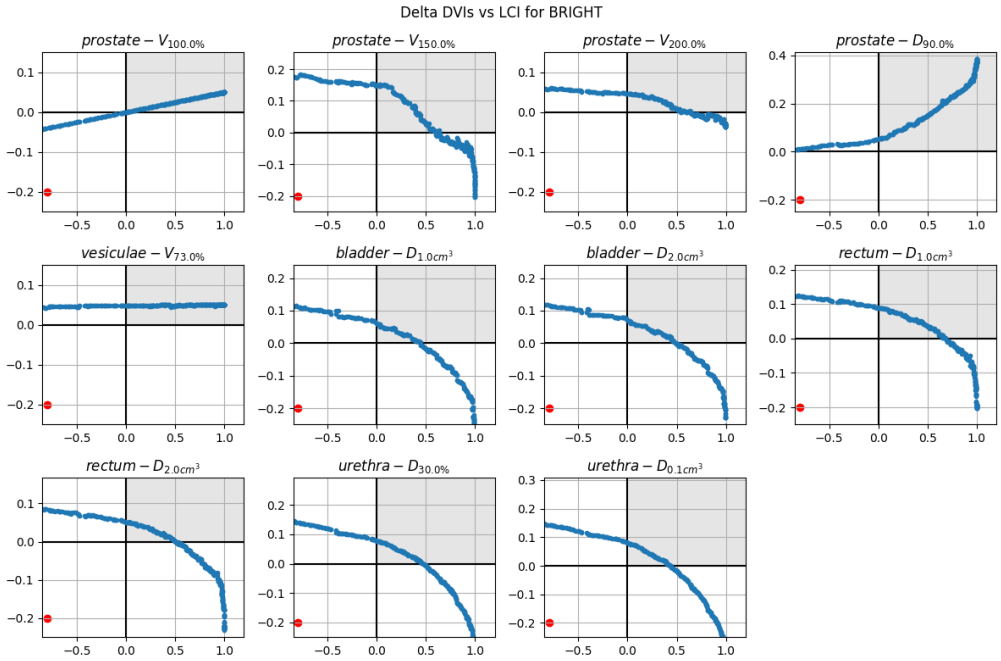


(i) Patient 5 BRIGHT best approximation front delta DVIs

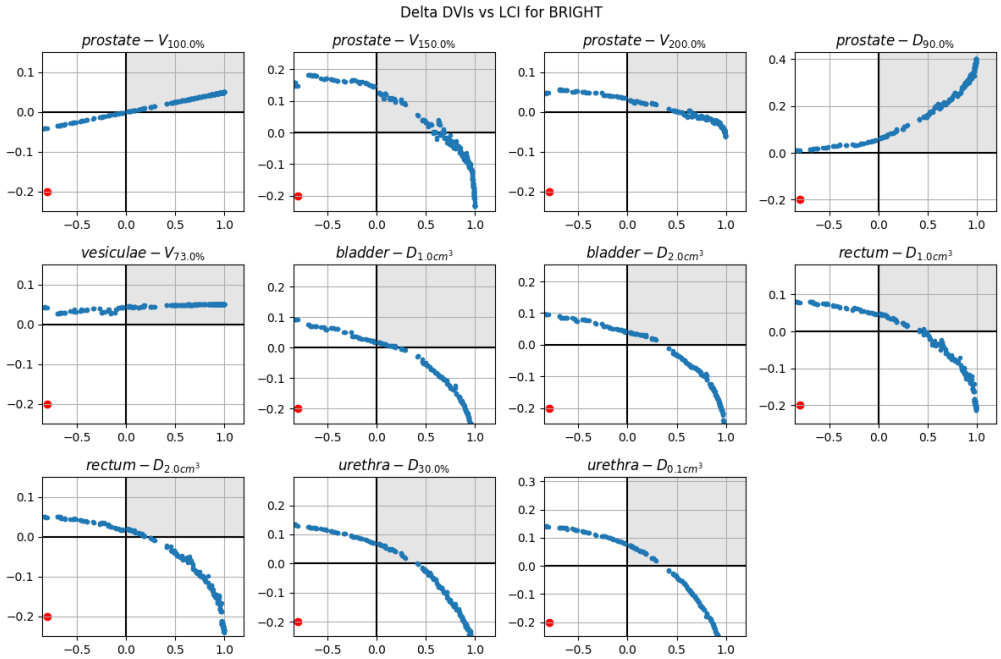


(j) Patient 5 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

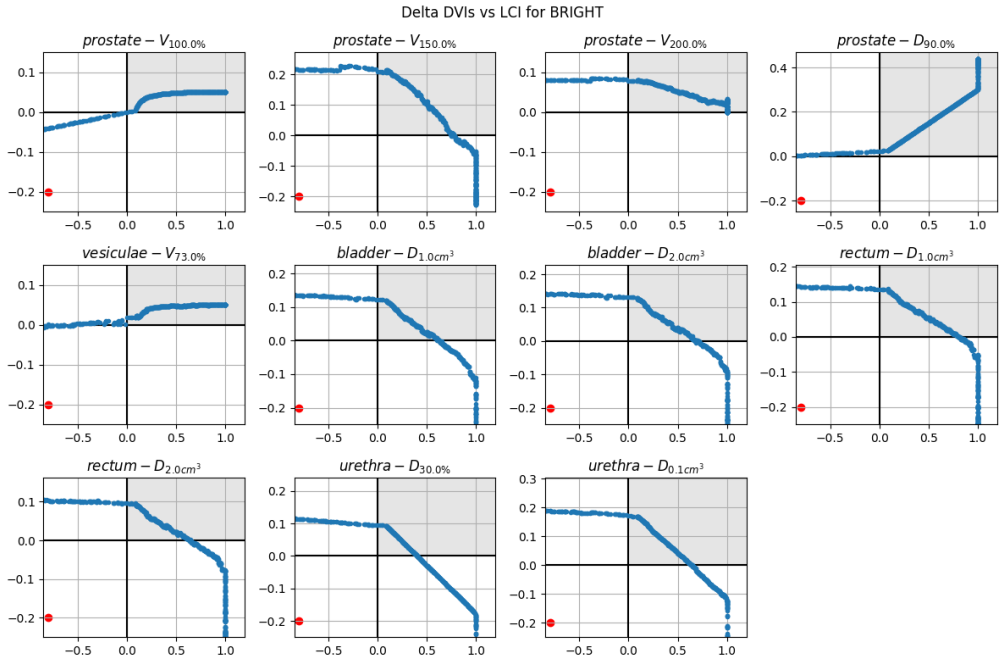


(k) Patient 6 BRIGHT best approximation front delta DVIs

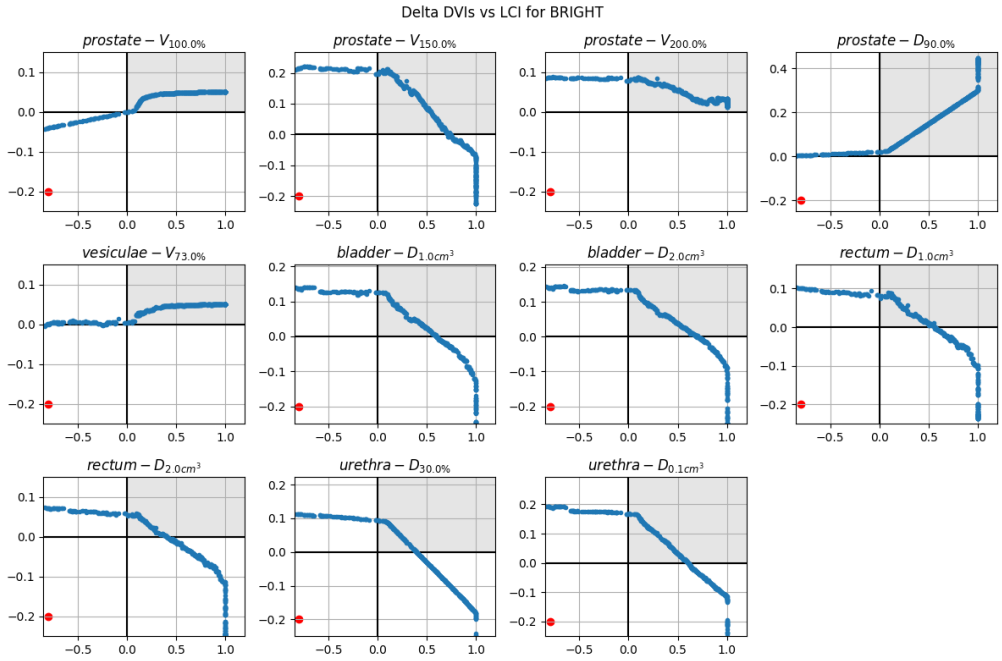


(l) Patient 6 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

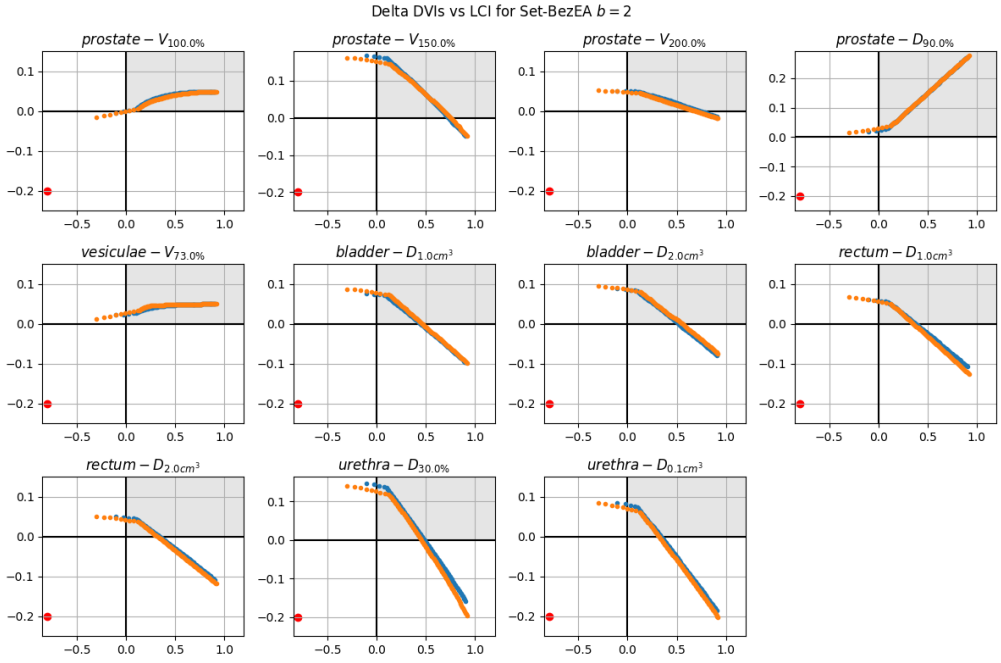


(m) Patient 7 BRIGHT best approximation front delta DVIs

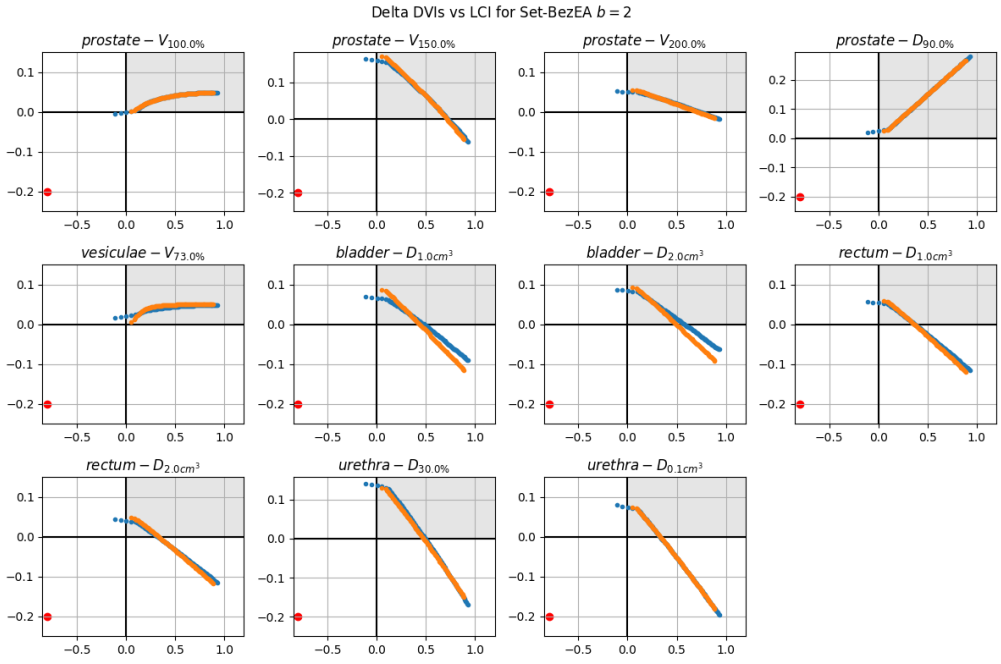


(n) Patient 7 BRIGHT worst approximation front delta DVIs

Figure B.4: Best and worst delta DVIs over all 10 runs.

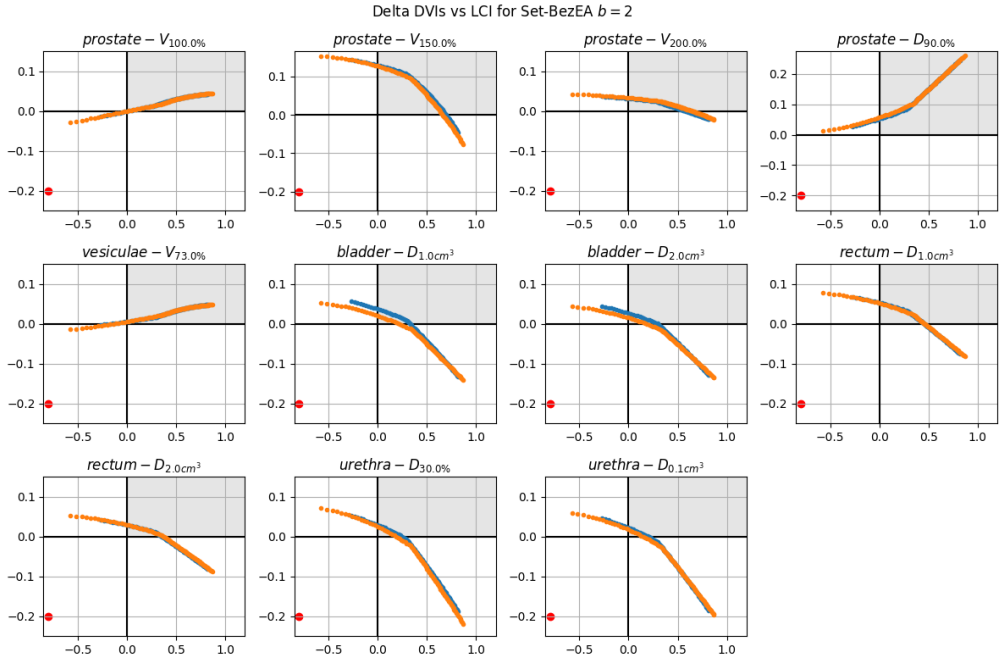


(o) Patient 1 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

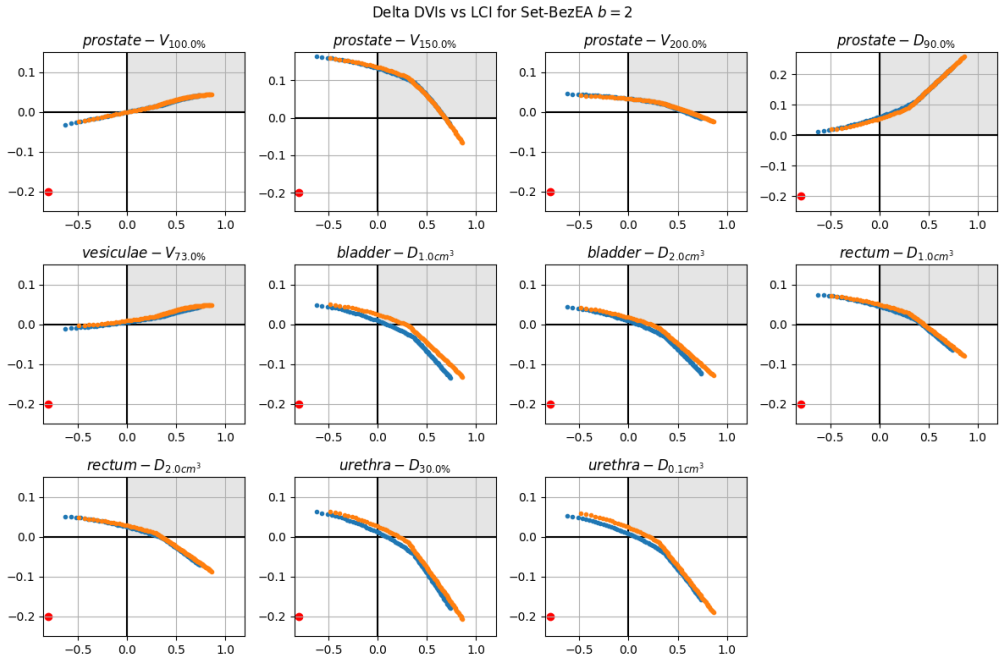


(p) Patient 1 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

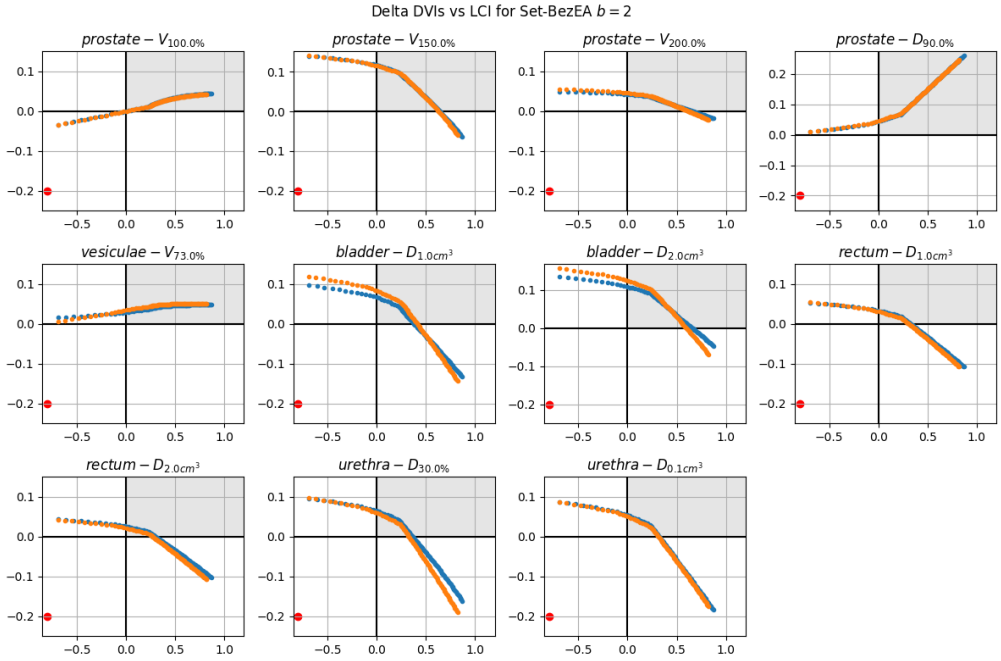


(q) Patient 2 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

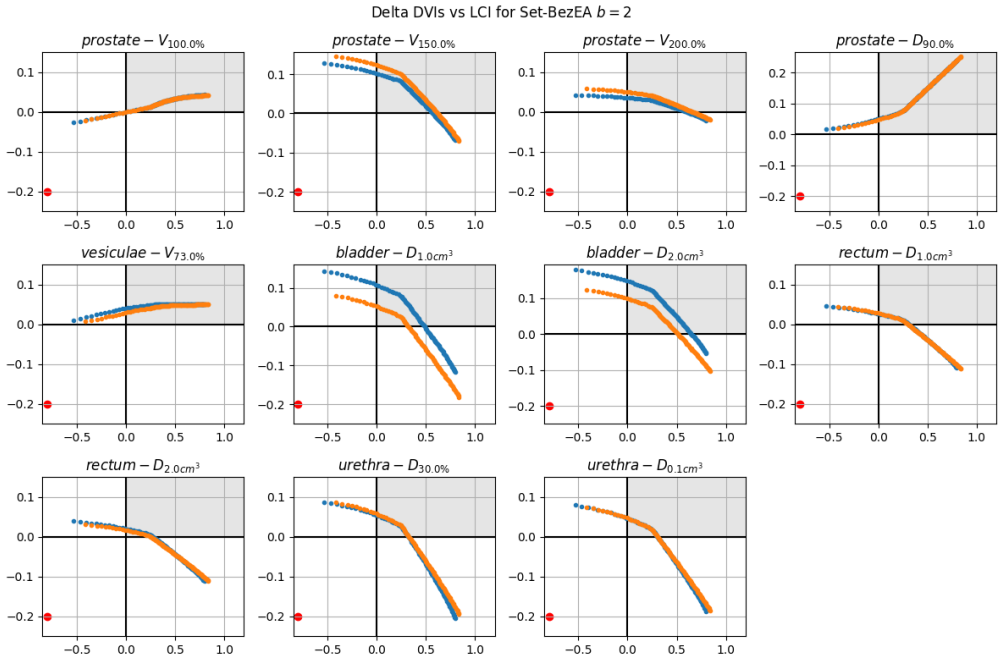


(r) Patient 2 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

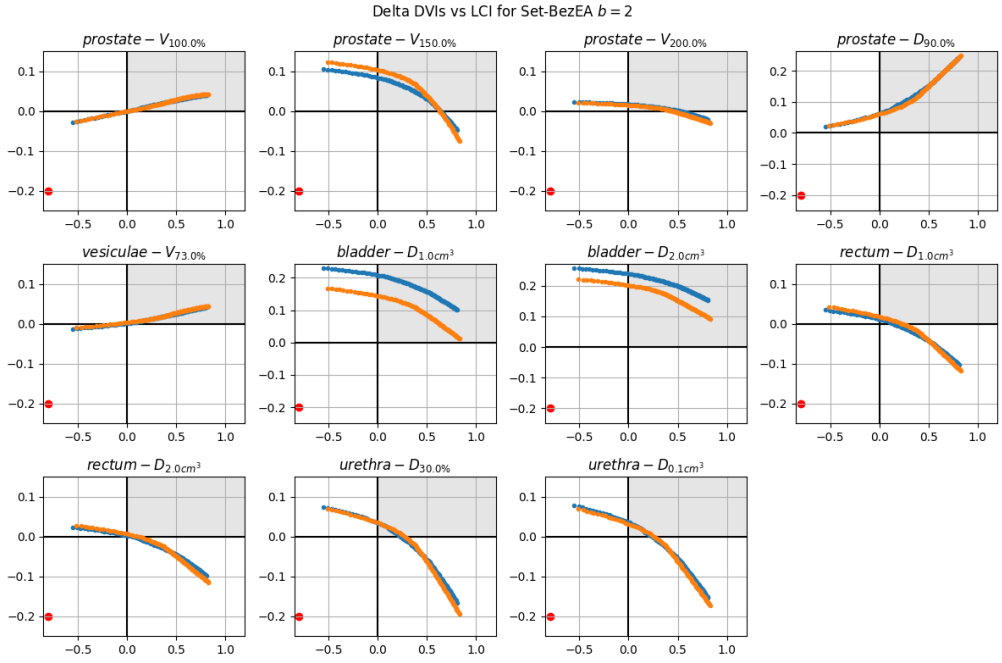


(s) Patient 3 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

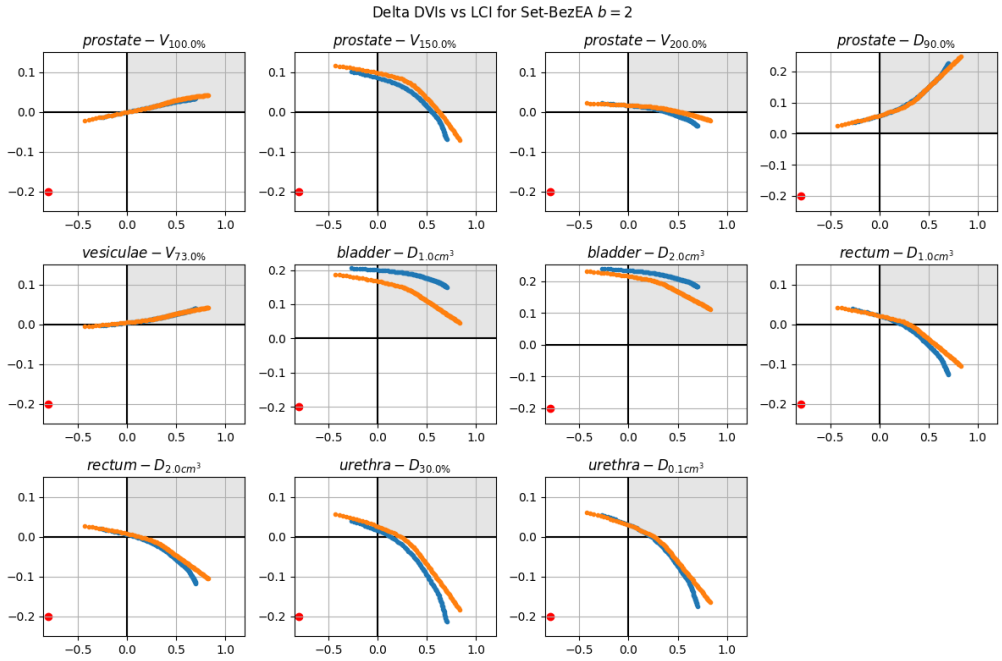


(t) Patient 3 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

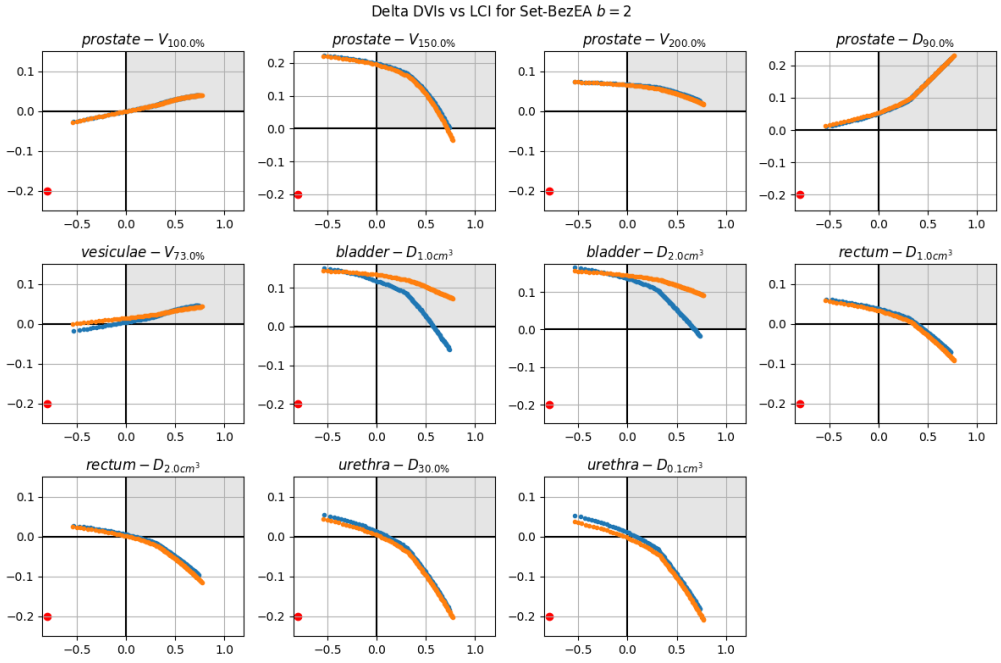


(u) Patient 4 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

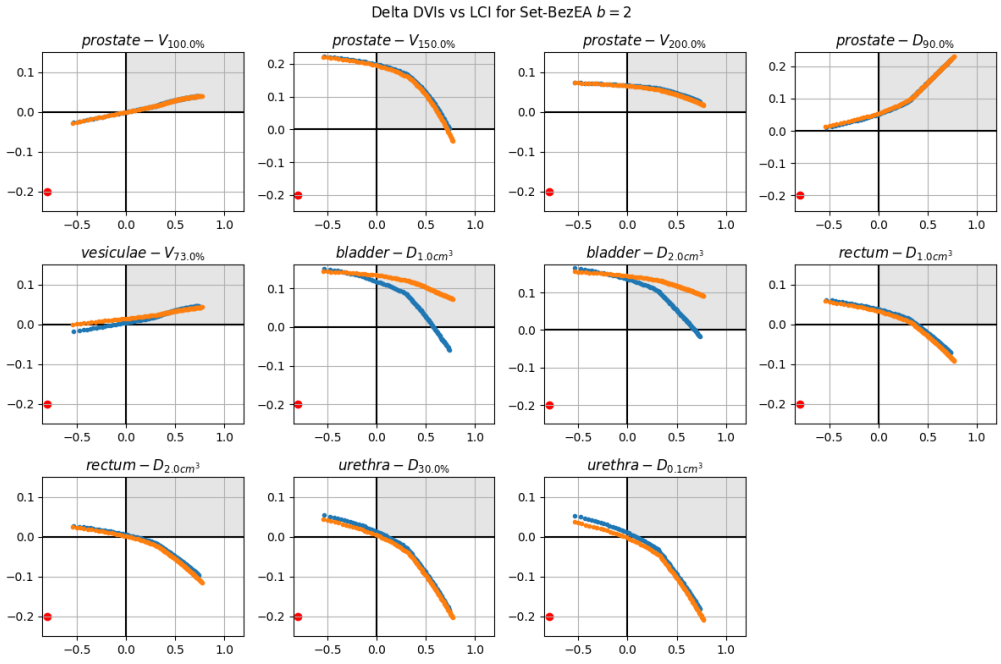


(v) Patient 4 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

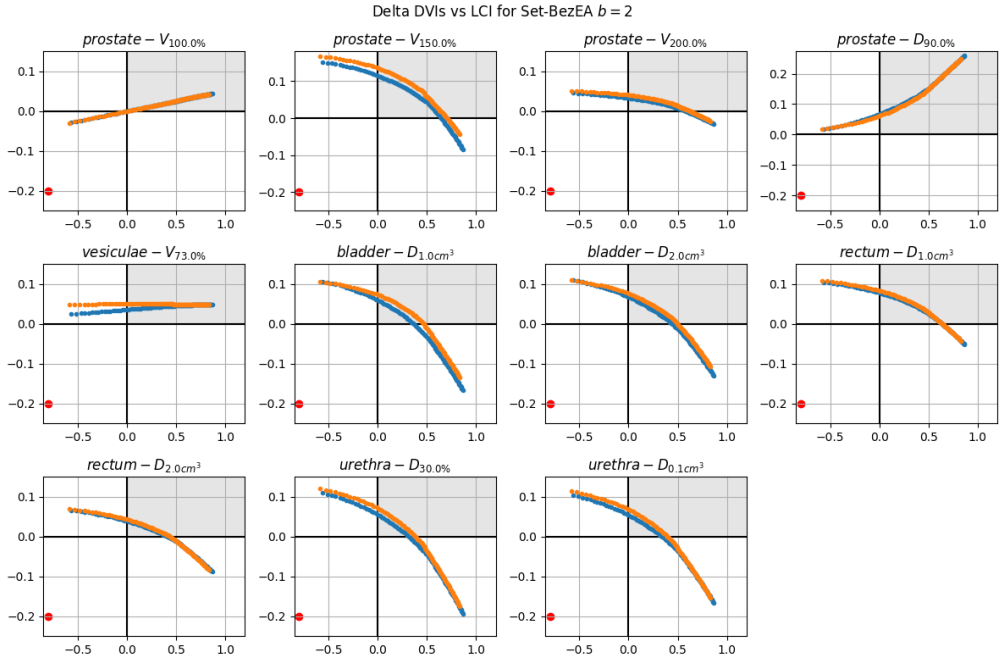


(w) Patient 5 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

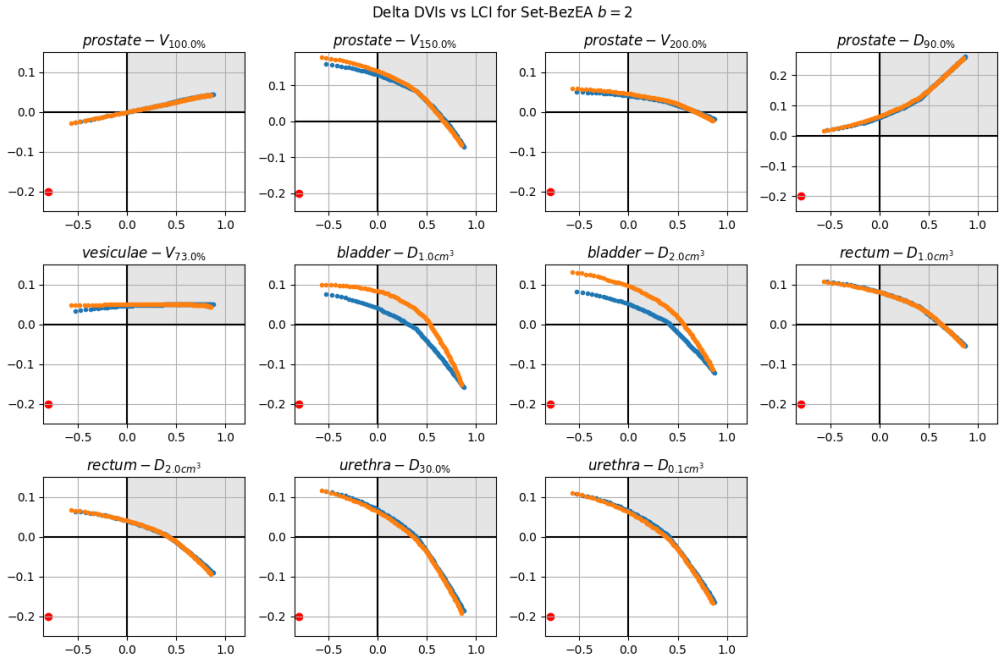


(x) Patient 5 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

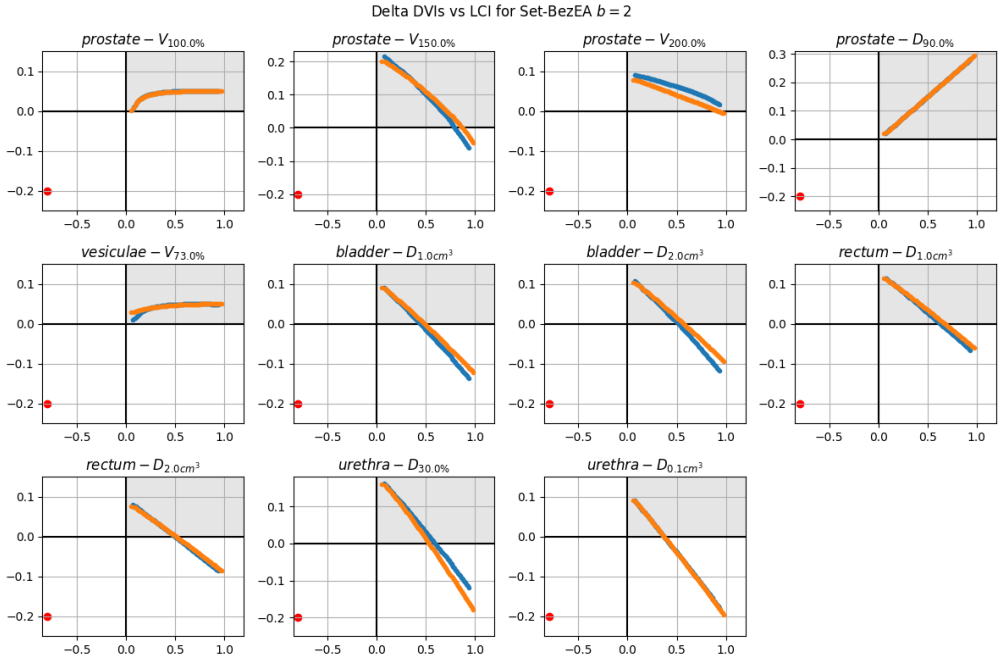


(y) Patient 6 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

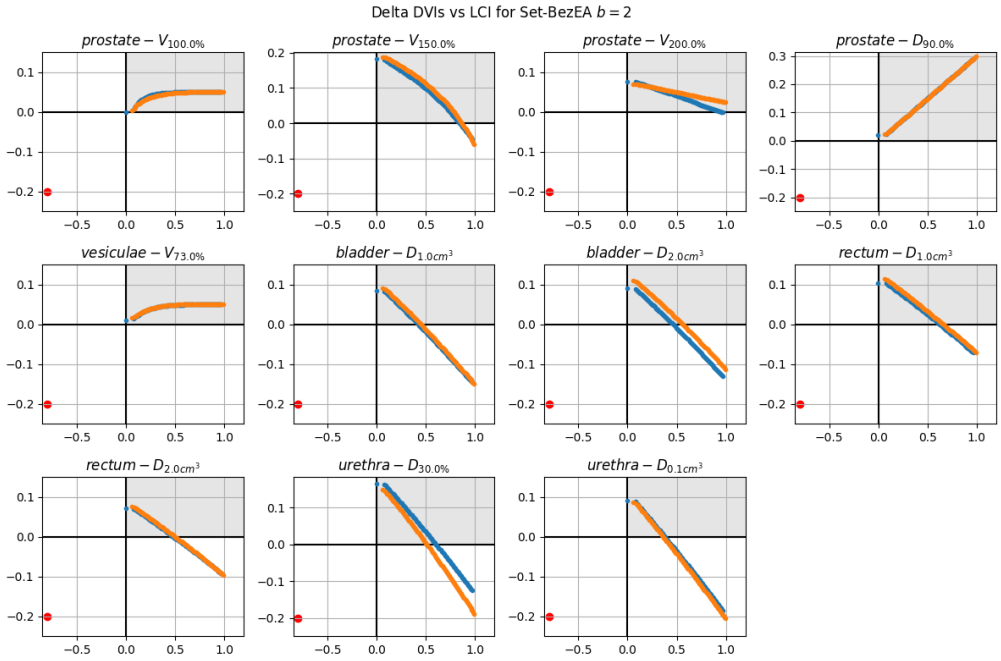


(z) Patient 6 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

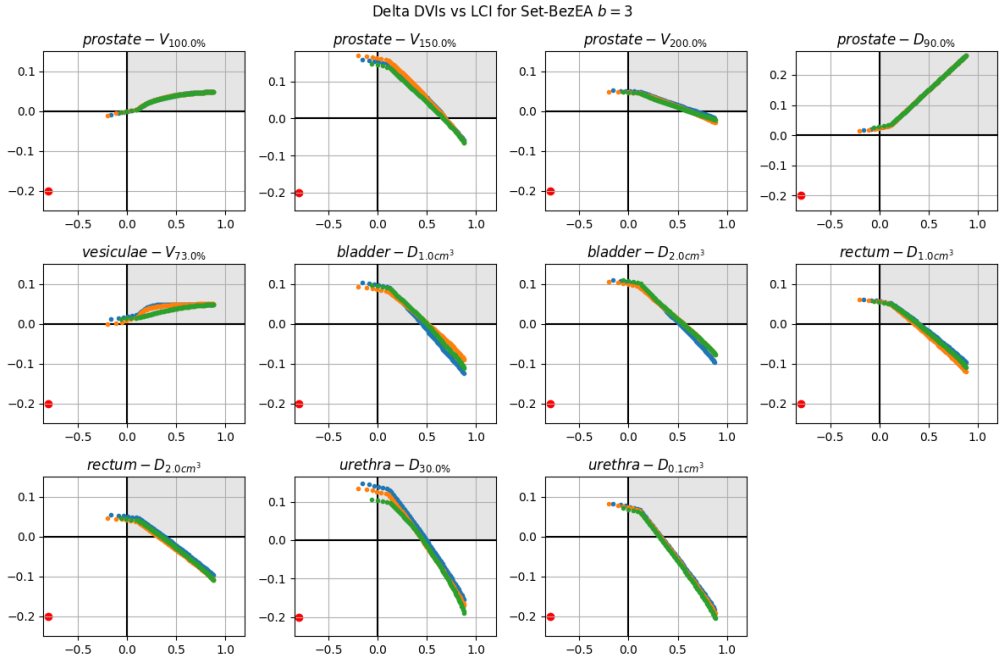


(aa) Patient 7 Set-BezEA ($b = 2$) approximation front delta DVIs of best run. One color for each of the b fronts.

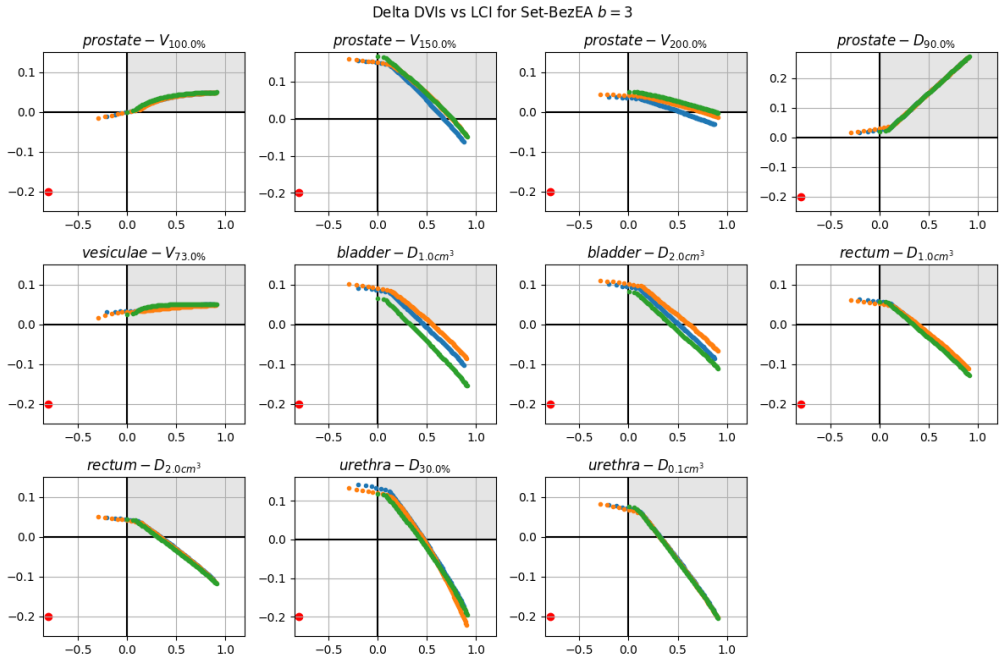


(ab) Patient 7 Set-BezEA ($b = 2$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

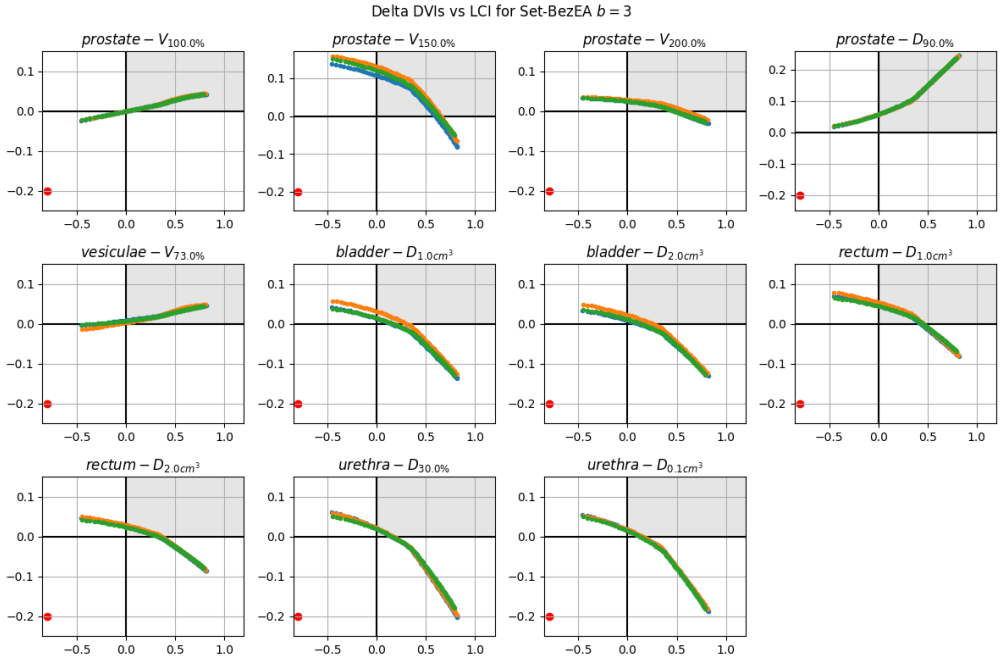


(ac) Patient 1 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

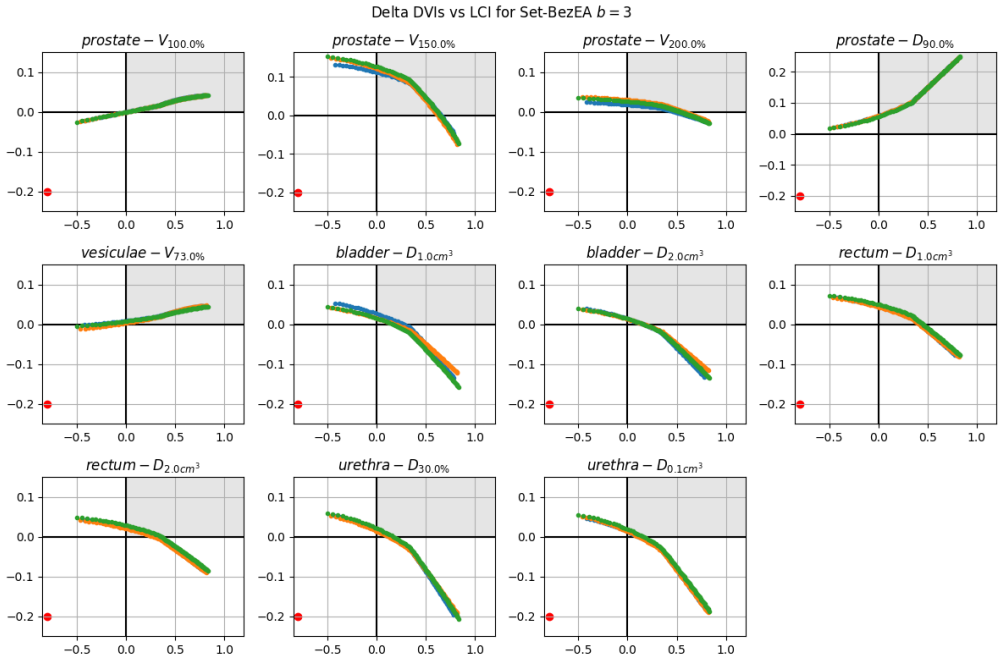


(ad) Patient 1 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

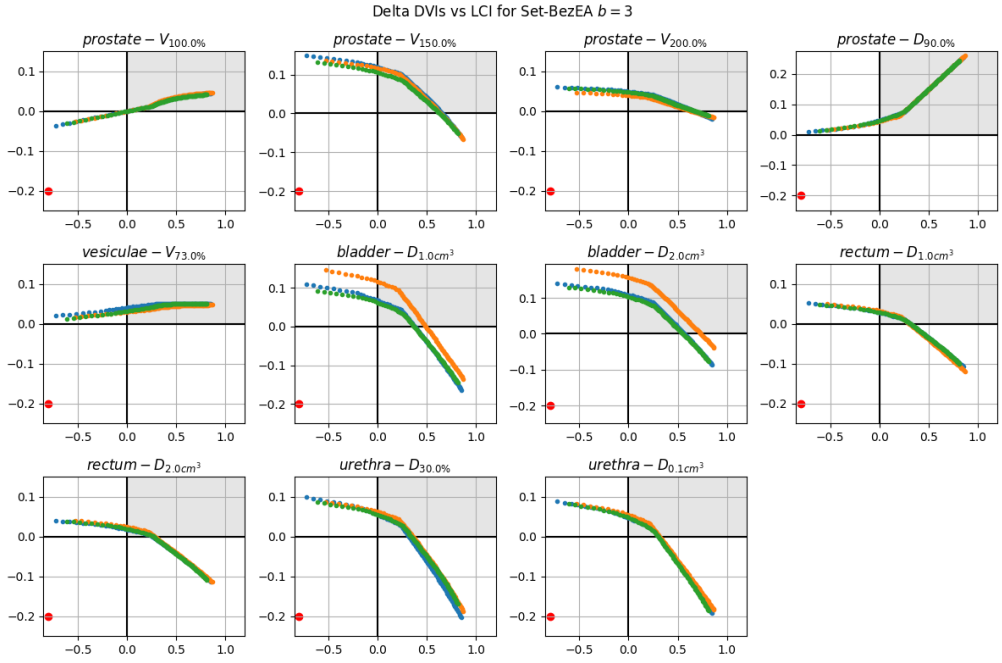


(ae) Patient 2 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

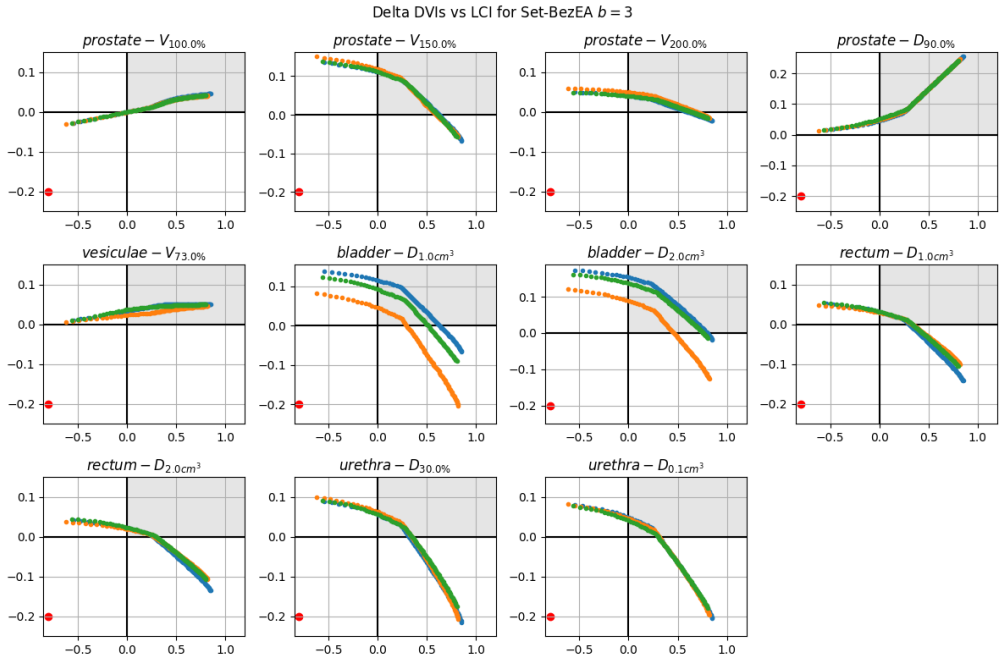


(af) Patient 2 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

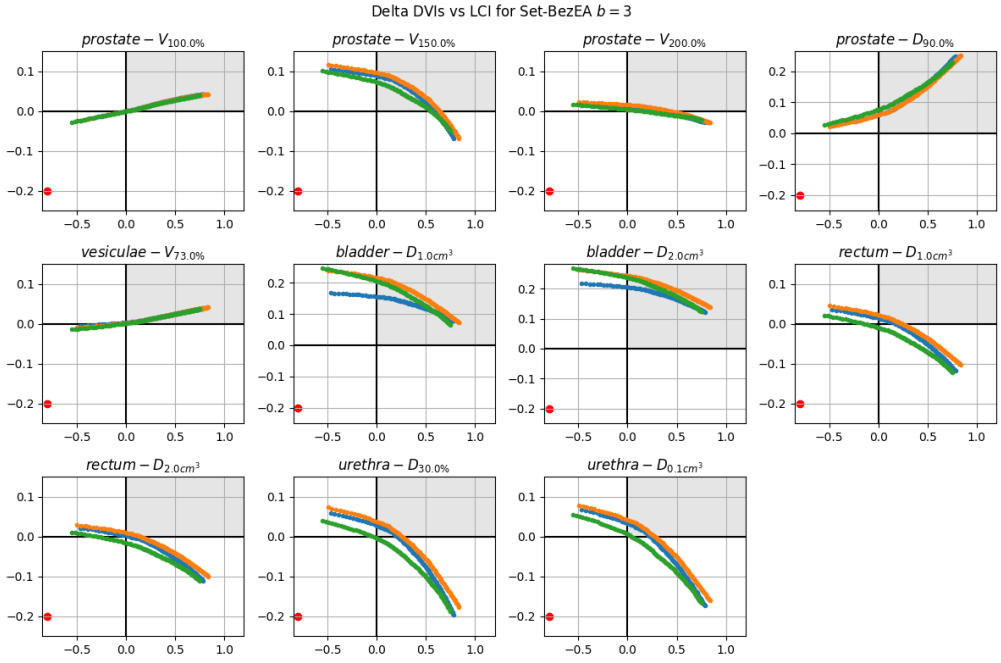


(ag) Patient 3 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

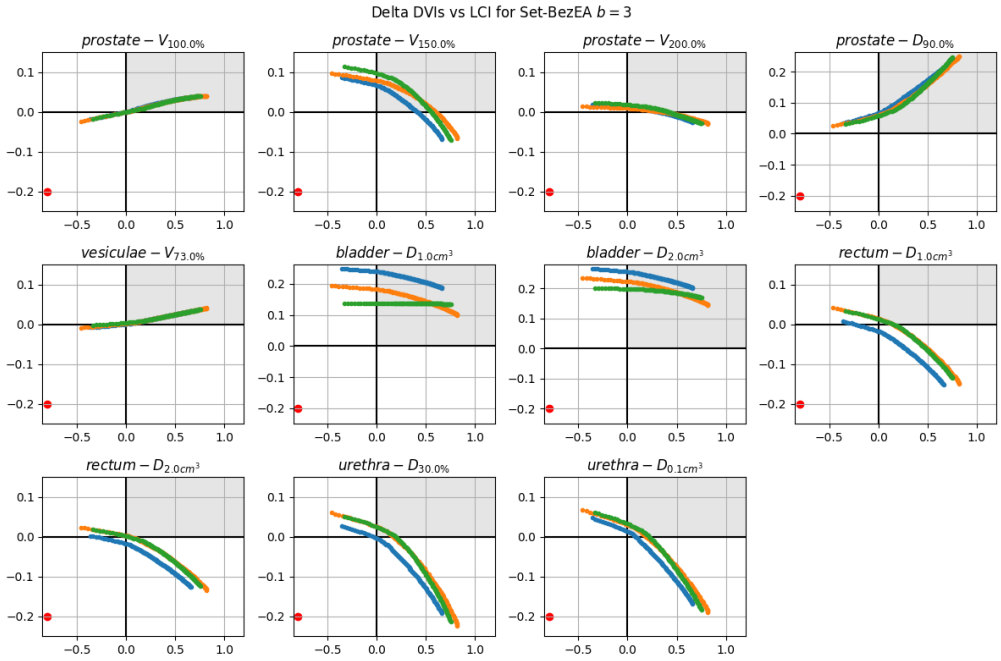


(ah) Patient 3 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

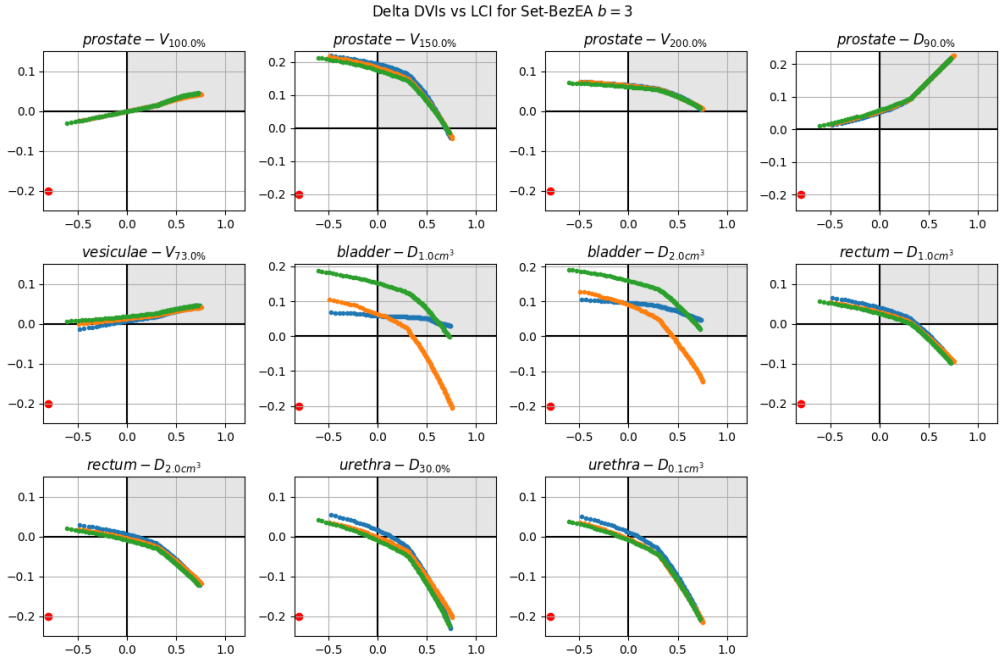


(ai) Patient 4 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

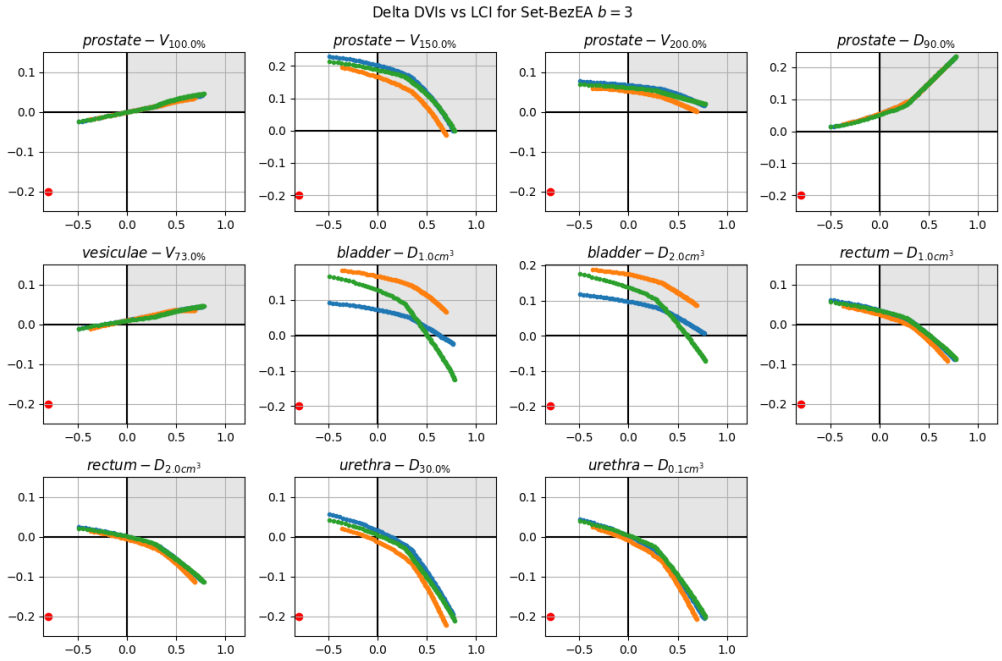


(aj) Patient 4 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

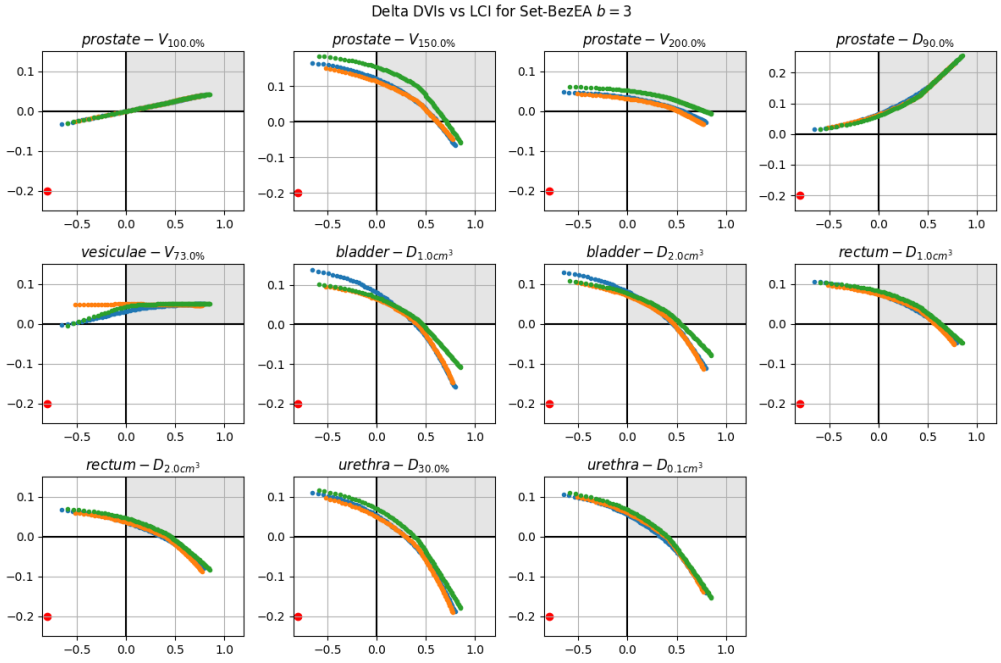


(ak) Patient 5 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

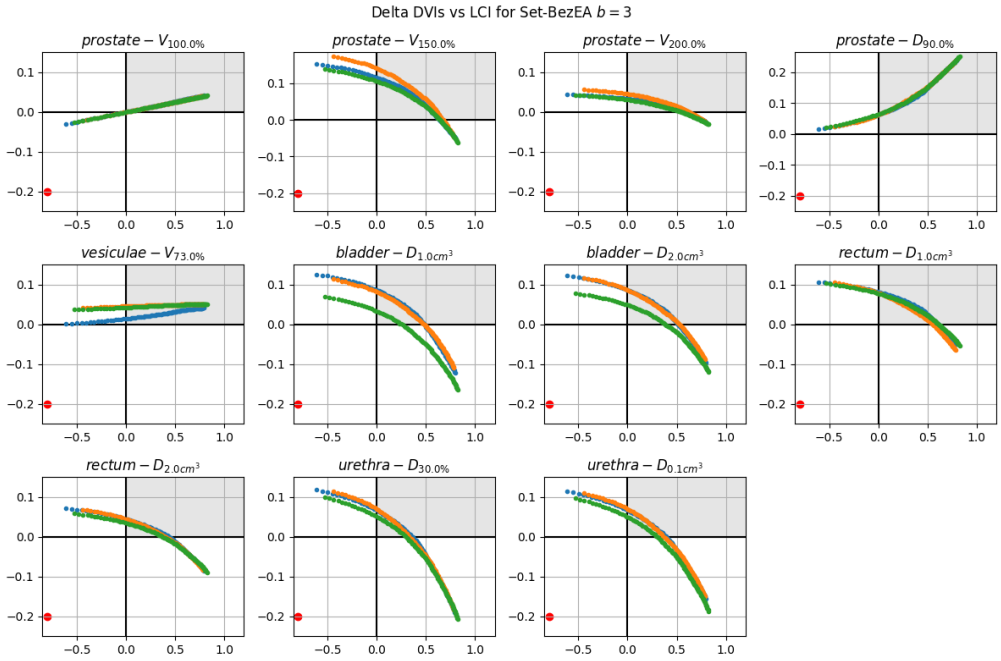


(al) Patient 5 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

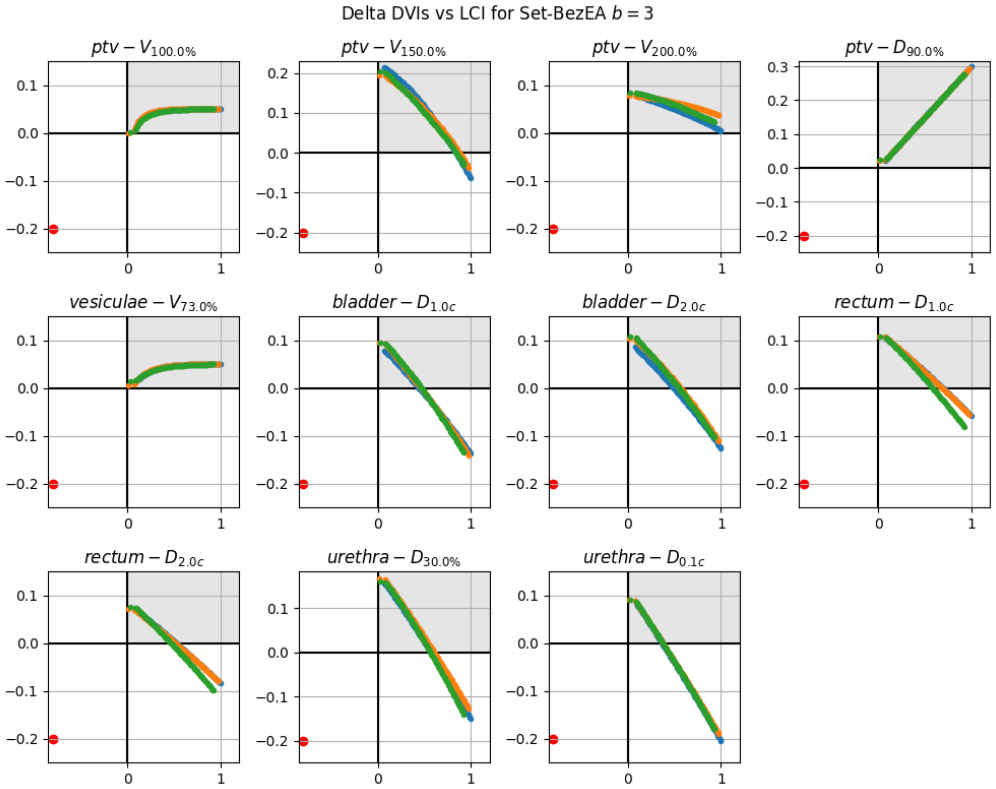


(am) Patient 6 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

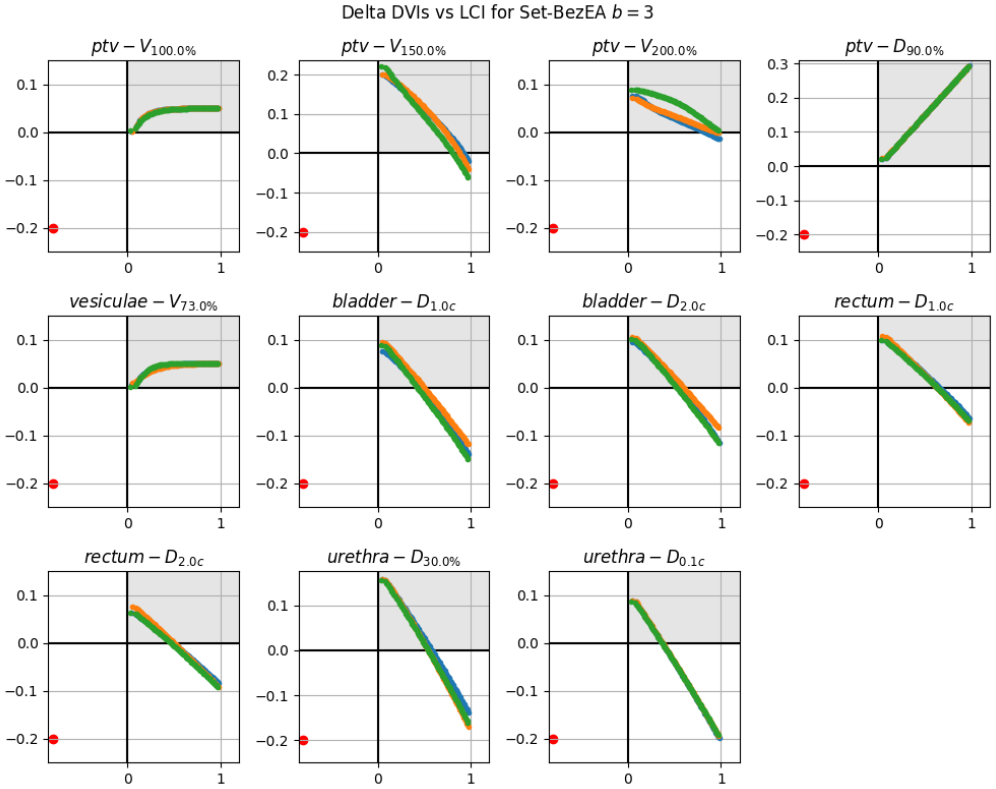


(an) Patient 6 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

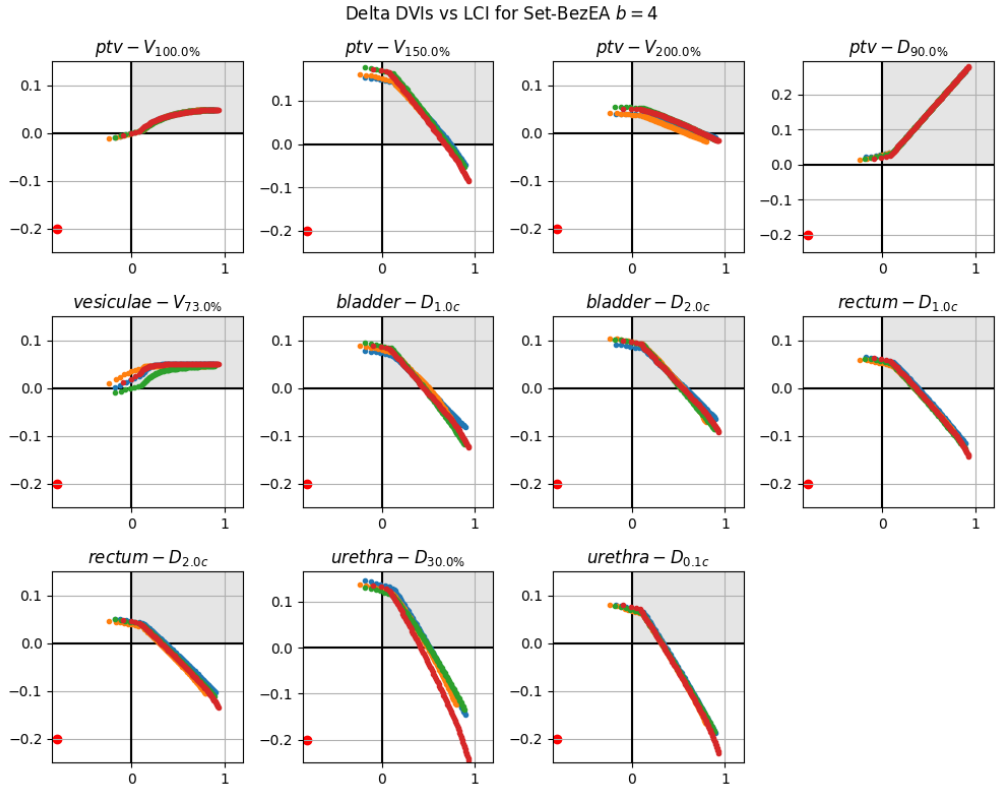


(ao) Patient 7 Set-BezEA ($b = 3$) approximation front delta DVIs of best run. One color for each of the b fronts.

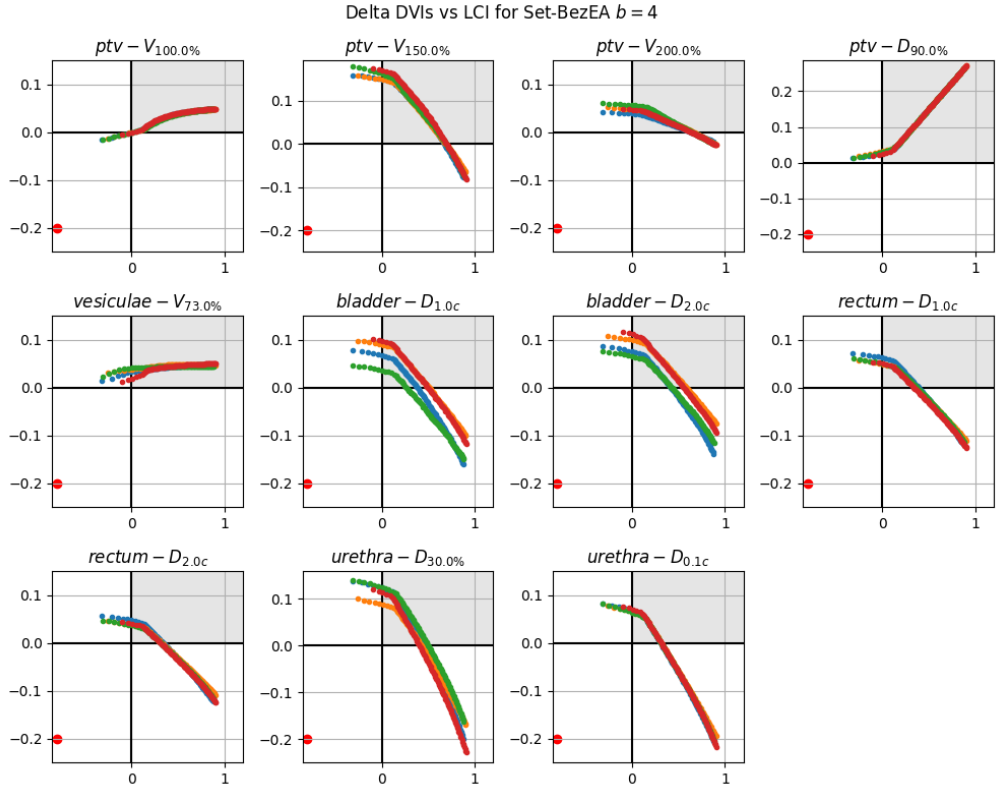


(ap) Patient 7 Set-BezEA ($b = 3$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

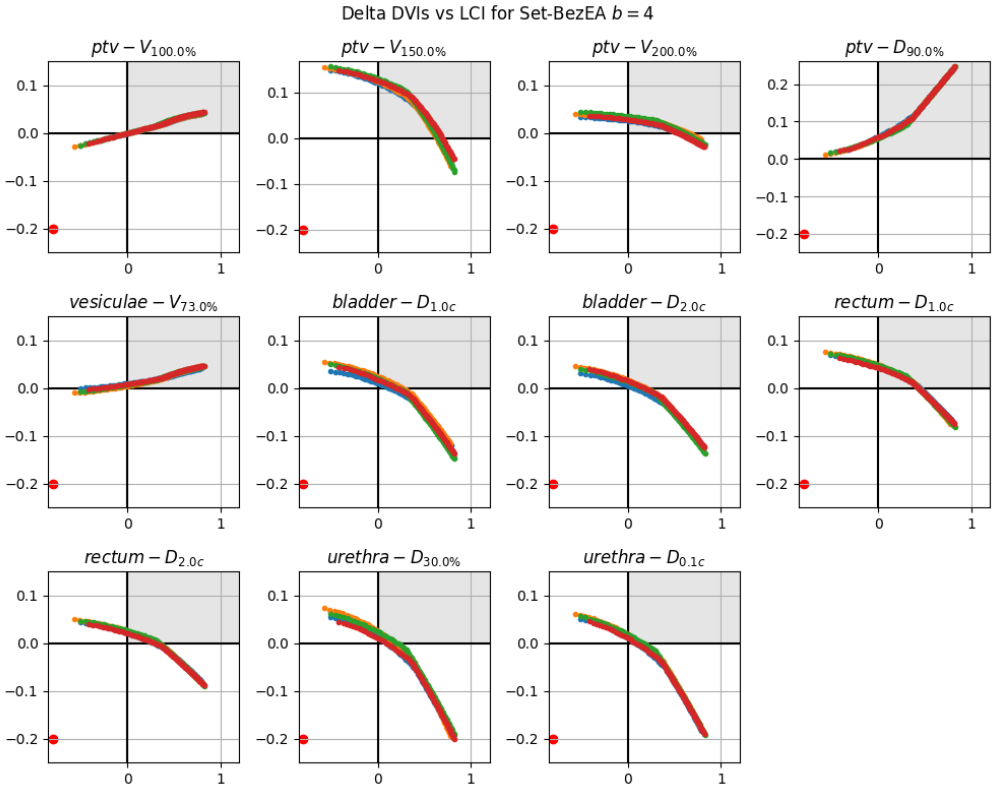


(aq) Patient 1 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

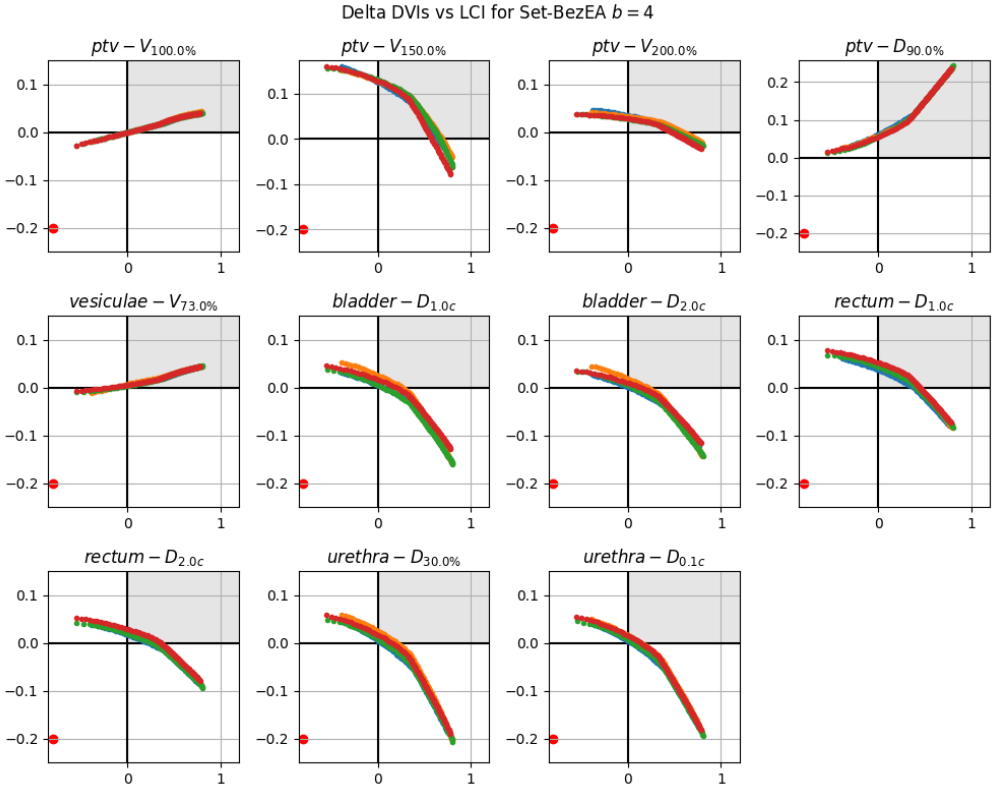


(ar) Patient 1 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

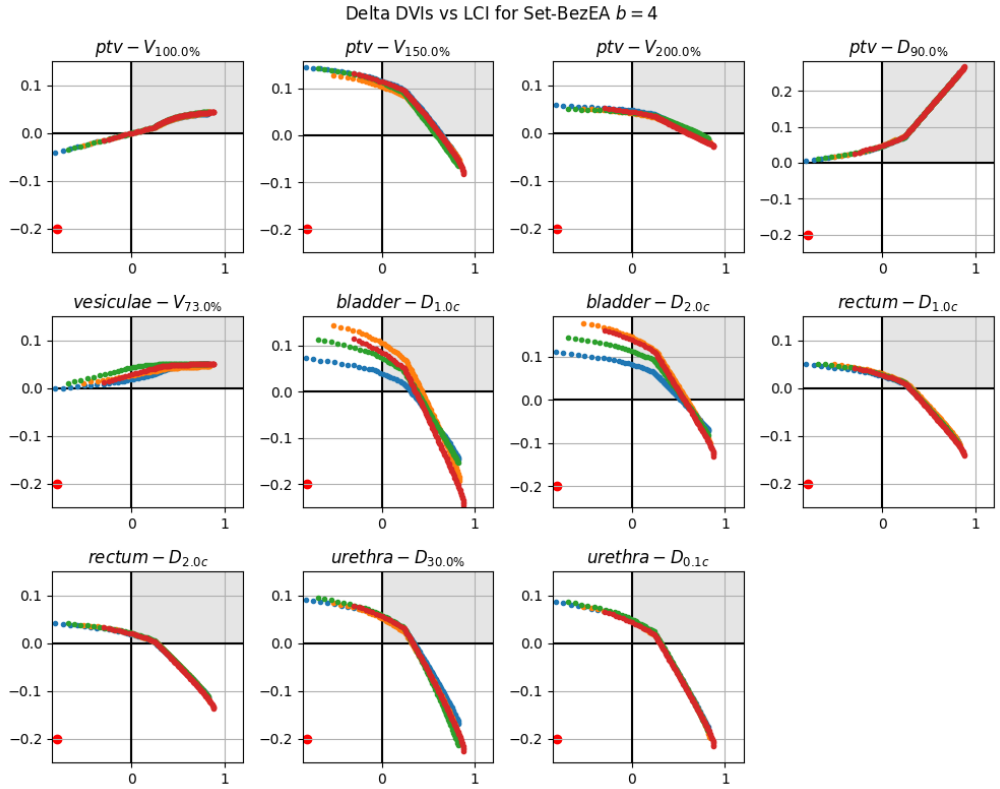


(as) Patient 2 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

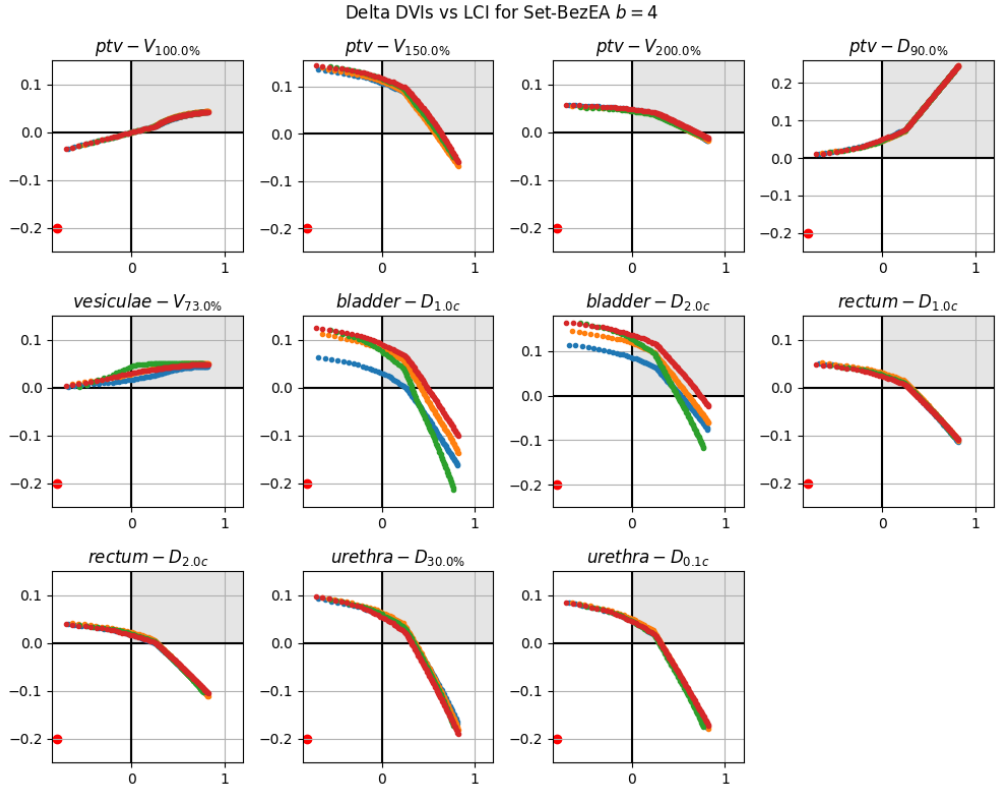


(at) Patient 2 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

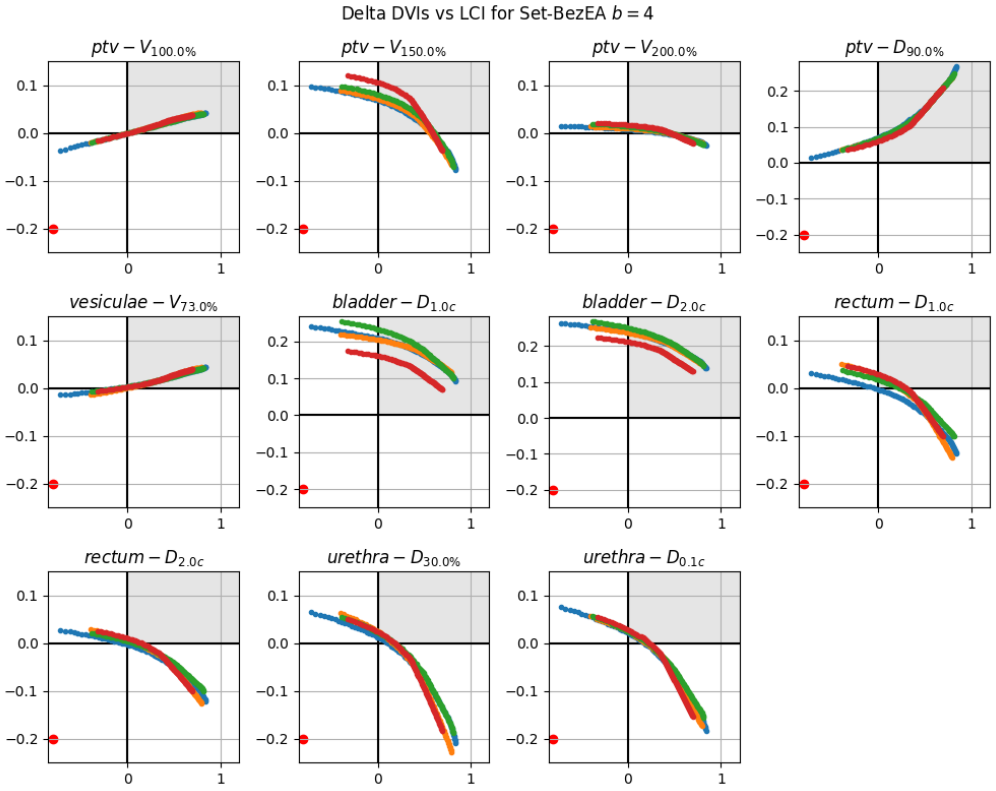


(au) Patient 3 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

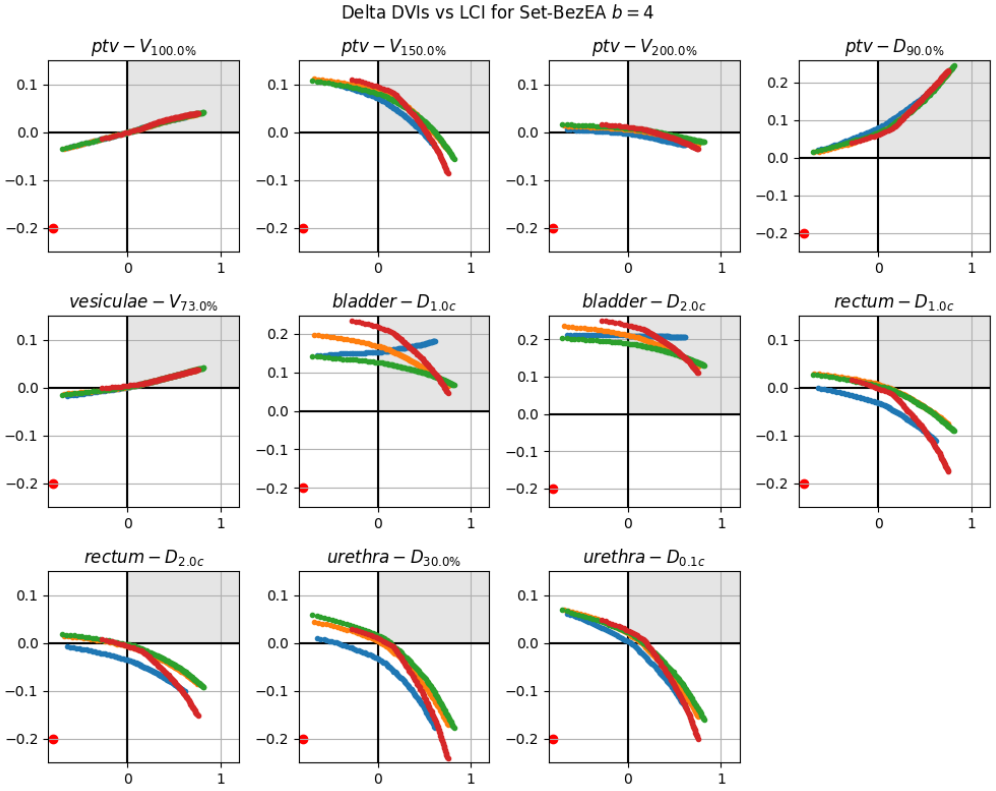


(av) Patient 3 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

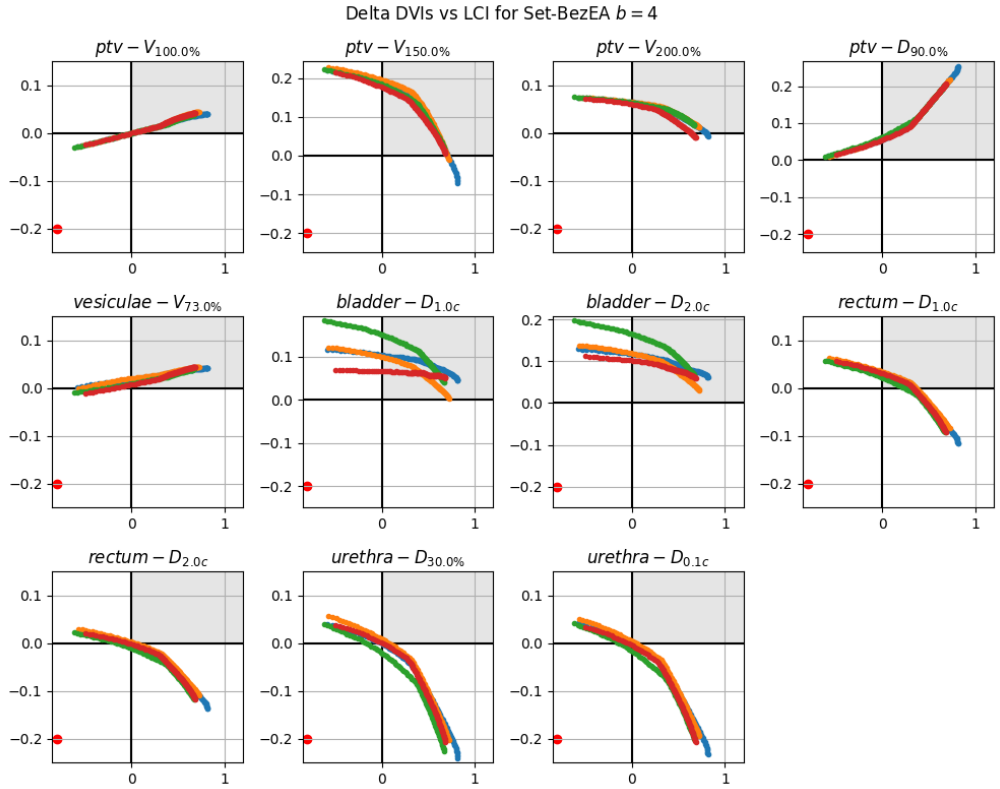


(aw) Patient 4 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

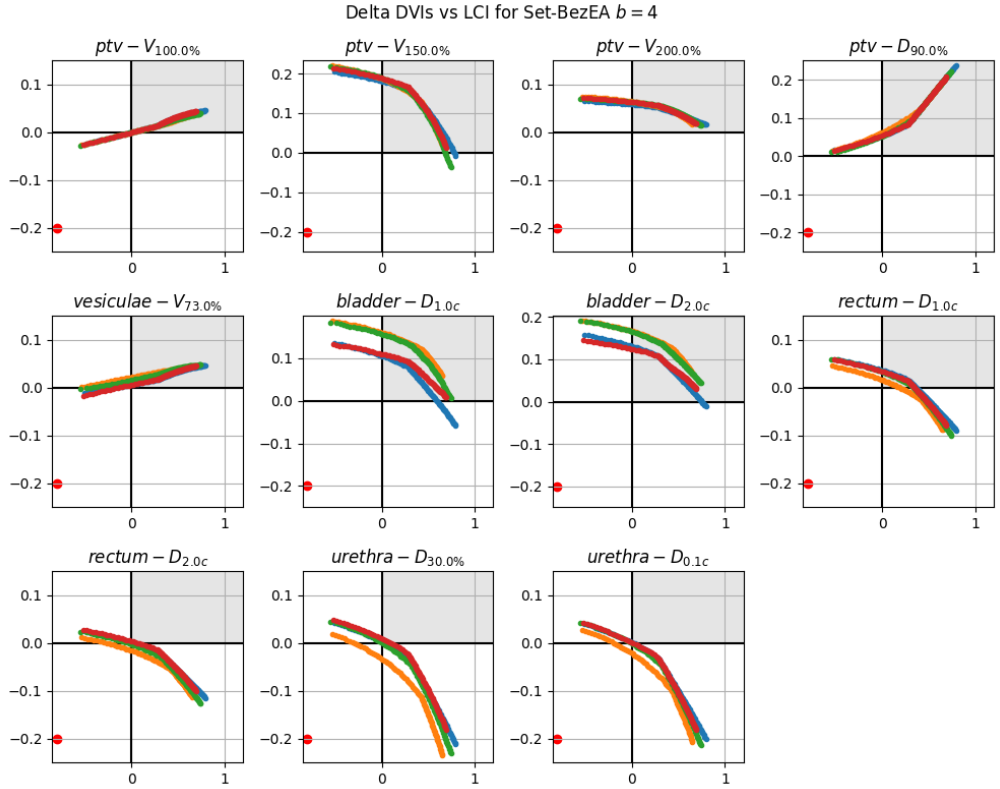


(ax) Patient 4 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

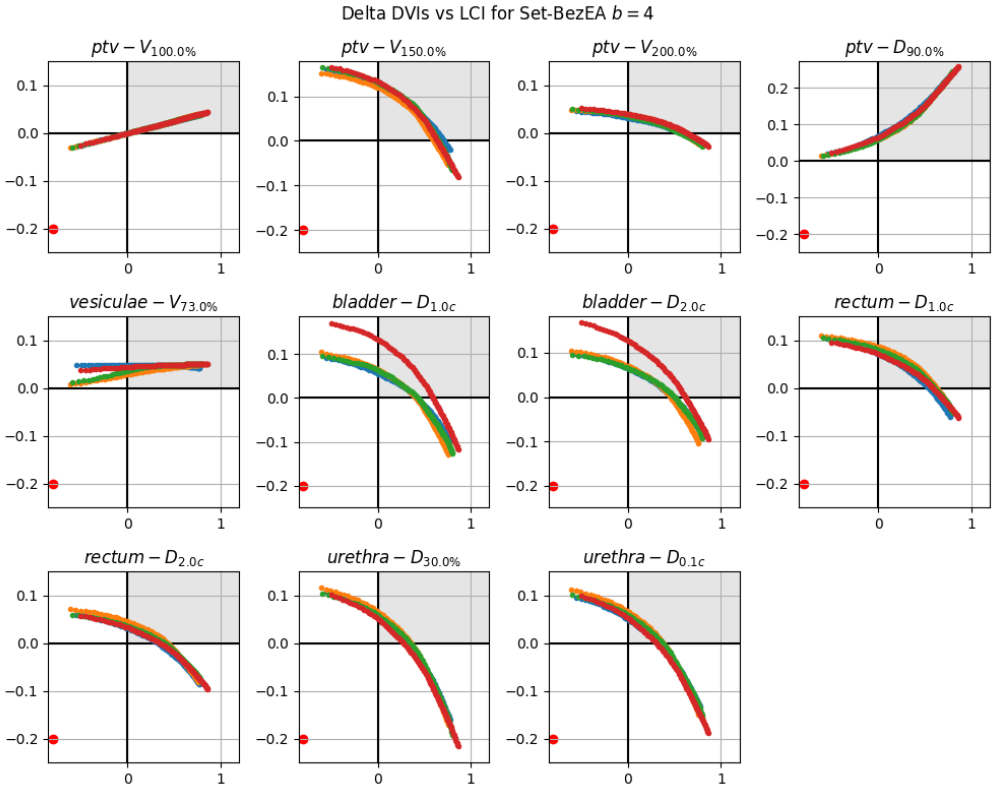


(ay) Patient 5 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

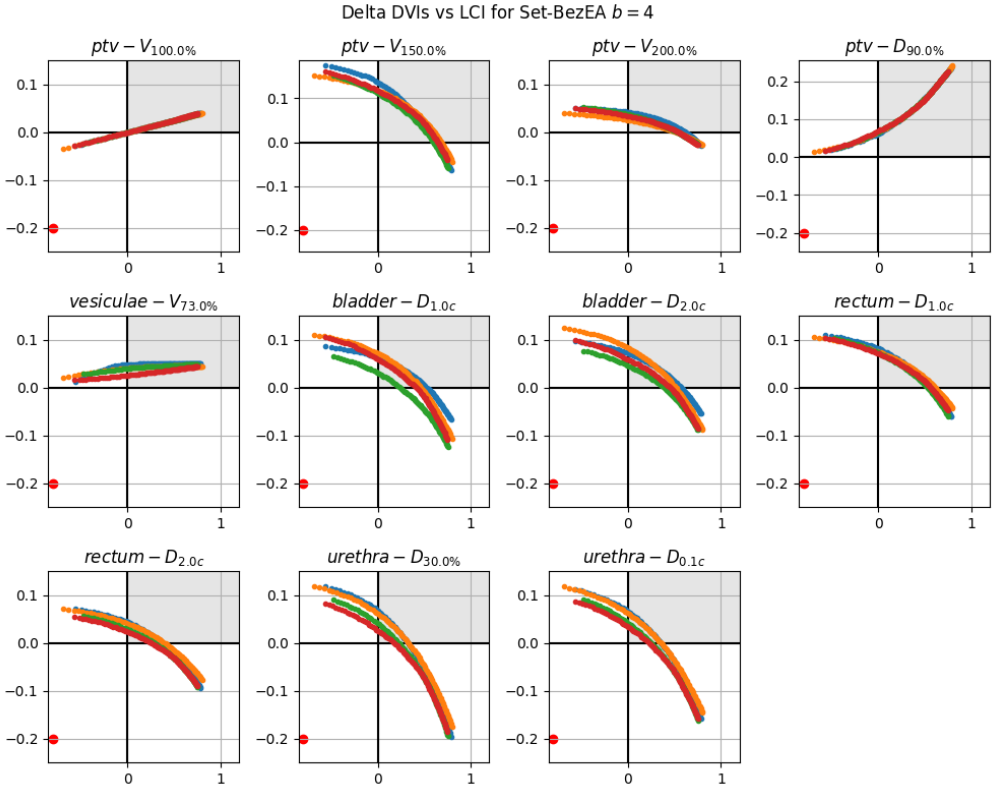


(az) Patient 5 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

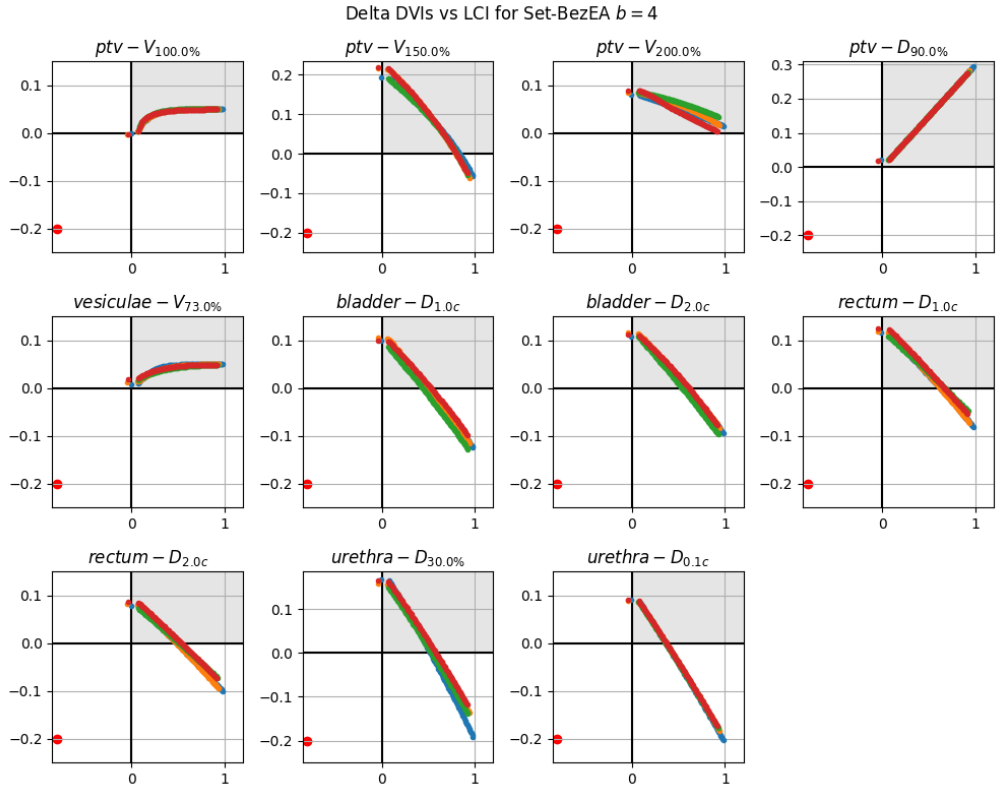


(ba) Patient 6 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.

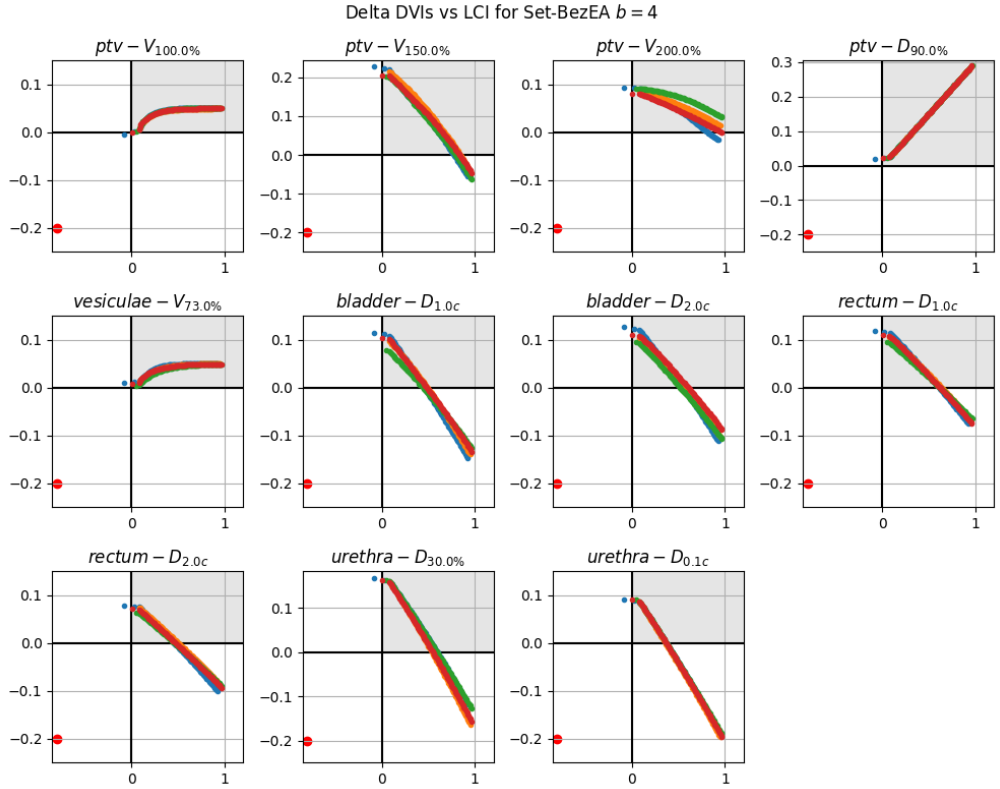


(bb) Patient 6 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.



(bc) Patient 7 Set-BezEA ($b = 4$) approximation front delta DVIs of best run. One color for each of the b fronts.



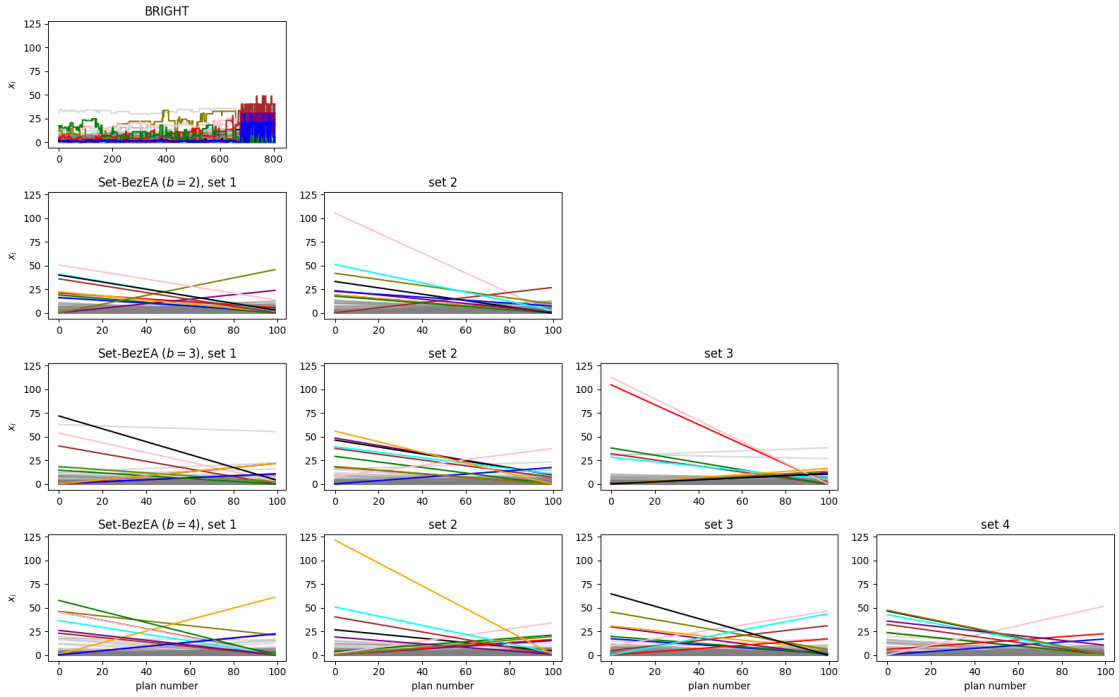
(bd) Patient 7 Set-BezEA ($b = 4$) approximation front delta DVIs of worst run. One color for each of the b fronts.

Figure B.4: Best and worst delta DVIs over all 10 runs.

B.5 Parallel Coordinates Plots

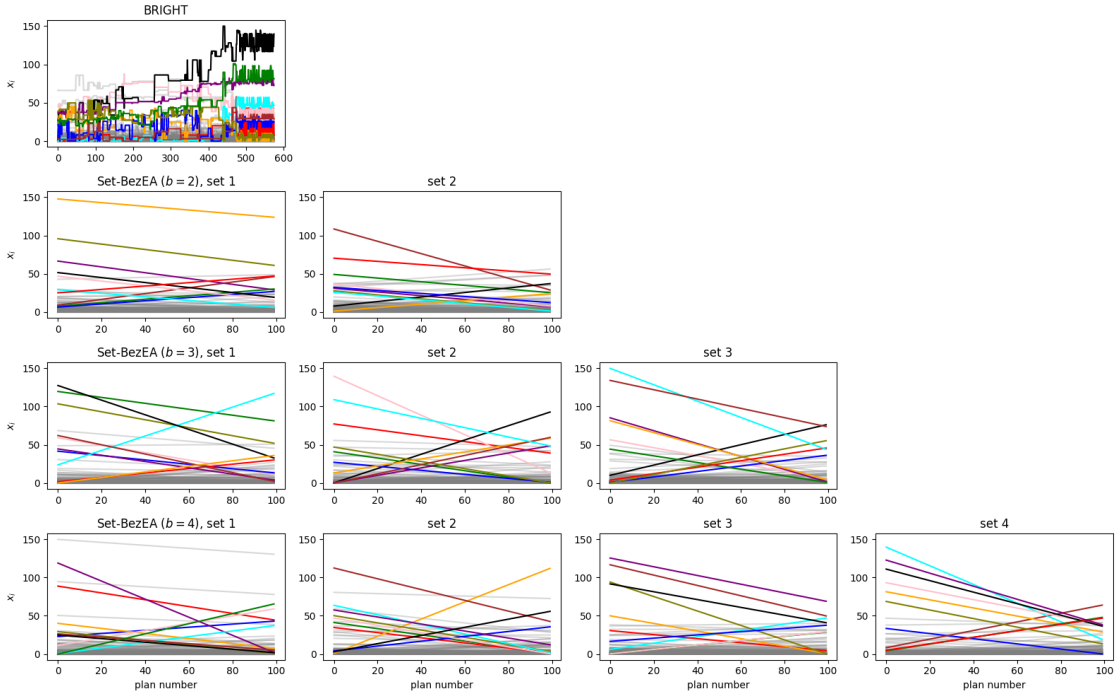
In order to visualize the smoothness of the decision variable values of the approximation sets originating from Set-BezEA, a few parallel coordinate plots will be shown. These display, for each of the decision variable, the value of the decision variable on the y axis and the plan number on the x-axis. Using such a visualization, the difference between the decision variable values for two adjacent treatment plans in the approximation set can be visualized. Here, adjacency in the approximation set is defined as being adjacent in the approximation front based on Pareto dominance.

The parallel coordinates plots show the course of the decision variable values for each and every decision variable used in optimization. But, only the ten decision variables with the highest standard deviation of all decision variables will be shown in color. The rest is given a semi-transparent grey color in order to not clutter the graphs too much in cases with up to 300 decision variables. For each of the 7 patients, the parallel coordinates are shown for the runs that produced the best results in terms of hypervolume.

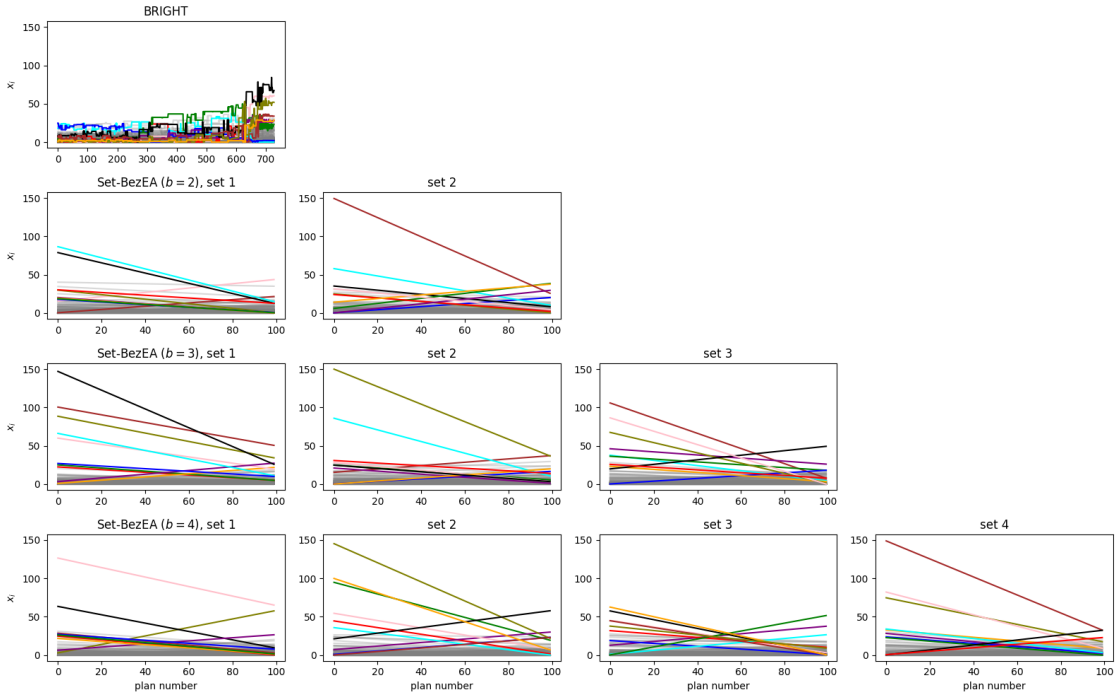


(a) Patient 1 parallel coordinates plot

Figure B.5: Parallel coordinates plots for best runs per patient per algorithm

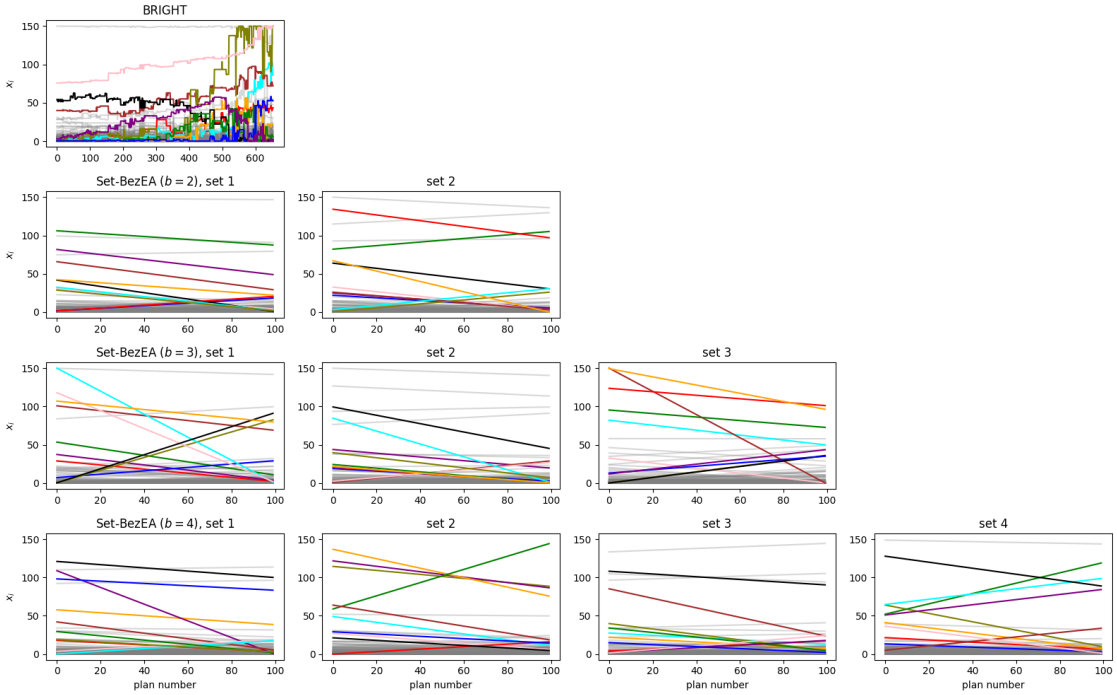


(b) Patient 2 parallel coordinates plot

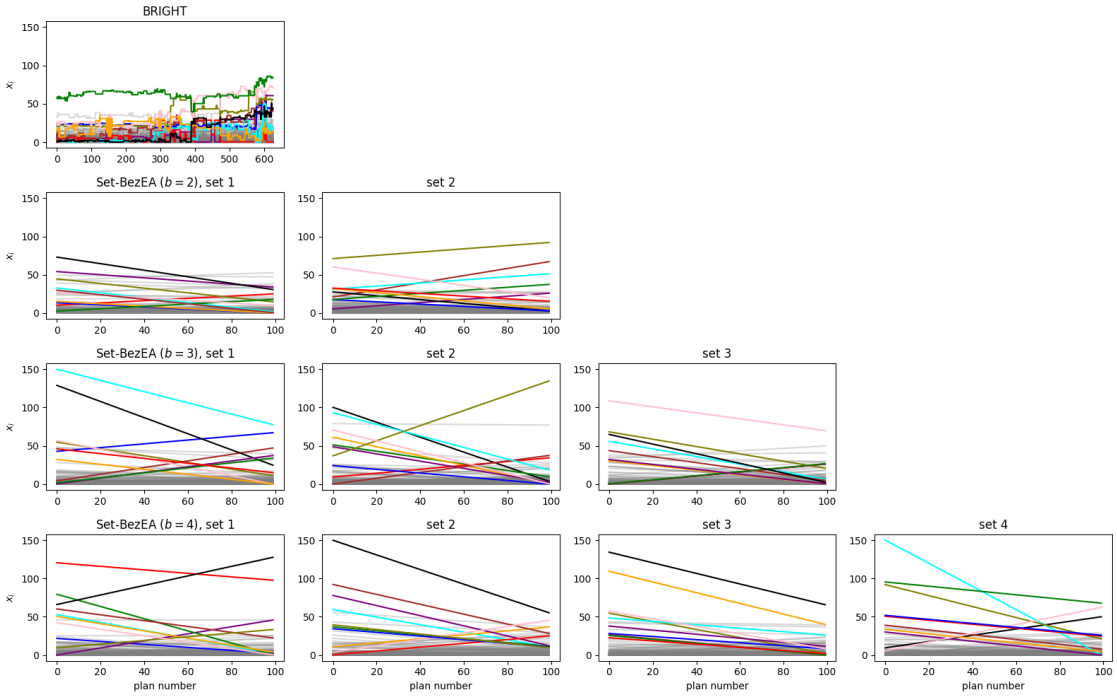


(c) Patient 3 parallel coordinates plot

Figure B.5: Parallel coordinates plots for best runs per patient per algorithm

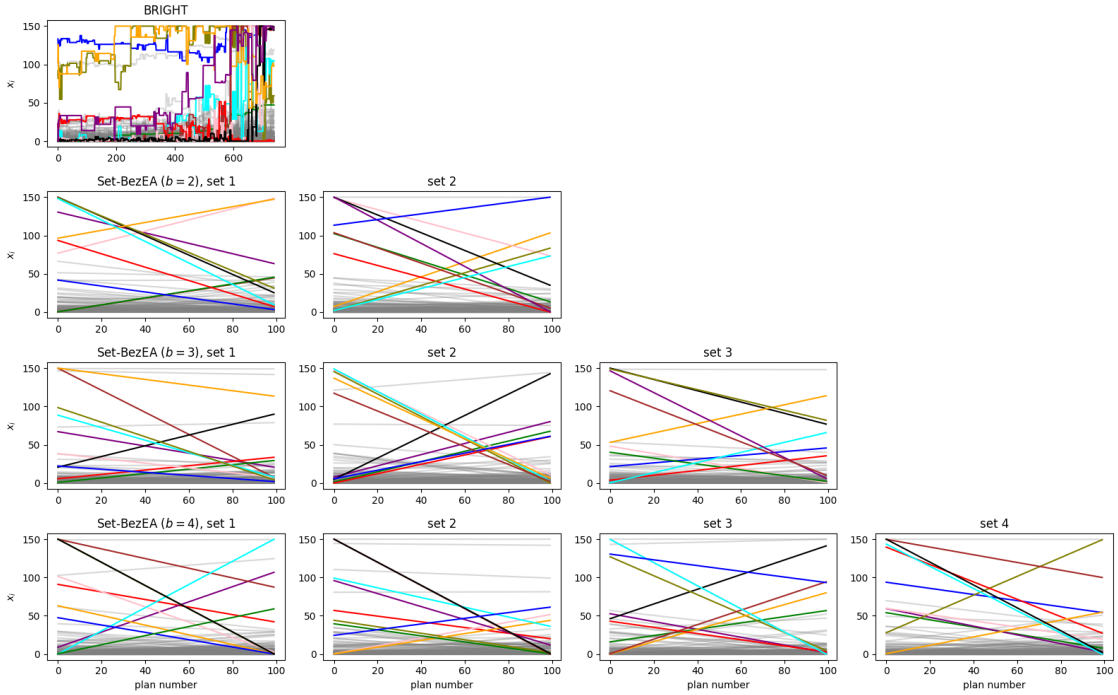


(d) Patient 4 parallel coordinates plot

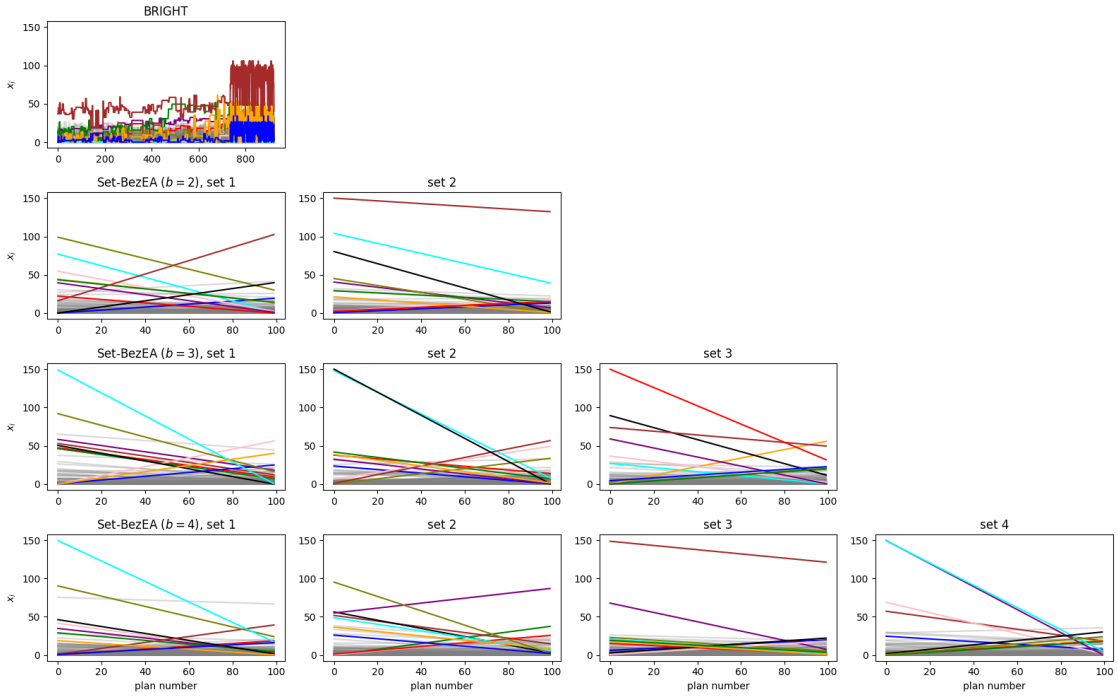


(e) Patient 5 parallel coordinates plot

Figure B.5: Parallel coordinates plots for best runs per patient per algorithm



(f) Patient 6 parallel coordinates plot



(g) Patient 7 parallel coordinates plot

Figure B.5: Parallel coordinates plots for best runs per patient per algorithm

Appendix C

White Paper

A large part of the research performed in this thesis has also been submitted for review in the form of a white paper to the PPSN XVII: The seventeenth International Conference on Parallel Problem Solving from Nature, September 10–14, 2022, Dortmund, Germany. This is a bi-annual conference that has been held since 1990, with the proceedings historically being published by Springer in the Lecture Notes in Computer Science series. Topics covered by the conference concern, amongst others, genetic algorithms, evolutionary multiobjective optimization, and real world applications. As this closely relates to the topics covered in this thesis, the paper has been submitted to this conference. The application of the algorithms regarding the HDR brachytherapy problem have been left out of the paper, it only covers the theory and general experiments regarding MM-BezEA of Chapter 6.

A copy of the paper can be found on the pages below.

Obtaining Smoothly Navigable Approximation Sets in Bi-Objective Multi-Modal Optimization

Renzo J. Scholman[1,3][0000−0003−2813−015X], Anton Bouter[1][0000−0003−4599−0733], Leah R.M. Dickhoff[2][0000−0001−6720−4380] ⋆, Tanja Alderliesten[2][0000−0003−4261−7511], and Peter A.N. Bosman[1,3][0000−0002−4186−6666]

[1] Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
{Renzo.Scholman,Anton.Bouter,Peter.Bosman}@cwi.nl
[2] Leiden University Medical Center, Leiden, The Netherlands
{L.R.M.Dickhoff,T.Alderliesten}@lumc.nl
[3] Delft University of Technology, Delft, The Netherlands

Abstract. Even if a Multi-modal Multi-Objective Evolutionary Algorithm (MMOEA) is designed to find solutions well spread over all locally optimal approximation sets of a Multi-modal Multi-objective Optimization Problem (MMOP), there is a risk that the found set of solutions is not smoothly navigable because the solutions belong to various niches, reducing the insight for decision makers. To tackle this issue, a new MMOEAs is proposed: the Multi-Modal Bézier Evolutionary Algorithm (MM-BezEA), which produces approximation sets that cover individual niches and exhibit inherent decision-space smoothness as they are parameterized by Bézier curves. MM-BezEA combines the concepts behind the recently introduced BezEA and MO-HillVallEA to find all locally optimal approximation sets. When benchmarked against the MMOEAs MO_Ring_PSO_SCD and MO-HillVallEA on MMOPs with linear Pareto sets, MM-BezEA was found to perform best in terms of best hypervolume.

Keywords: Evolutionary algorithms · multi-modal multi-objective optimization · niching · Bézier curve estimation.

1 Introduction

Many real-world optimization problems have multiple conflicting objectives, whereby improvement in one objective often results in the deterioration of another. Multi-Objective Evolutionary Algorithms (MOEAs), like NSGA-II [9], MOEA/D [36], and MO-CMA-ES [17], are widely accepted to be well-suited to solve such Multi-objective Optimization Problems (MOPs) [11]. The aim is to obtain a set of solutions, called the approximation set, such that all solutions are non-dominated and the set itself is close to the set of Pareto-optimal solutions.

⋆ Leah R.M. Dickhoff was supported by the Dutch Cancer Society (KWF kankerbestrijding, Project N.12183) and Elekta.

Here, a solution $\boldsymbol{x}_0$ dominates $\boldsymbol{x}_1$ ($\boldsymbol{x}_0 \succ \boldsymbol{x}_1$) in an MOP with m objectives if $\forall i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) \leq f_i(\boldsymbol{x}_1)$ and $\exists i \in \{0, 1, ..., m-1\} : f_i(\boldsymbol{x}_0) < f_i(\boldsymbol{x}_1)$. The Pareto Set (PS) is $\mathcal{P}_S = \{\boldsymbol{x}_i | \neg \exists \boldsymbol{x}_j : \boldsymbol{x}_j \succ \boldsymbol{x}_i\}$ and the Pareto Front (PF) is $\mathcal{P}_F = \{(f_0(\boldsymbol{x}), \cdots, f_{m-1}(\boldsymbol{x})) \,|\boldsymbol{x} \in \mathcal{P}_S\}$.

A more complex type of MOPs is that of Multi-modal MOPs (MMOPs), where the goal is not to find one, but multiple, if not all, (local) PSs. In MMOPs, each of the PSs pertains to a *niche*, a subset of the search space, where a single mode resides, i.e., with one local PS. The PSs may, however, map to the same PF in objective space, similar to having multiple (locally) optimal solutions of the same quality in a single-objective problem, e.g., the sine function. Here we consider MMOPs in the case of real-valued parameters, or continuous optimization. This field has recently gotten more traction, with reviews [31], proposed formal definitions [14] and new visualization techniques [30].

In order to have MOEAs solve MMOPs, they need additional tools that prevent their convergence to a single niche in the landscape [22]. Niching [19] is one of such diversity-preserving tools used by Multi-modal MOEAs (MMOEAs) to effectively and simultaneously search for solutions near the (local) PS in each niche. Niching has been successfully applied to established MMOEAs in the form of the multi-objective particle swarm optimization using ring topology and special crowding distance (MO_Ring_PSO_SCD) algorithm [35] and Omni-optimizer [12] among others.

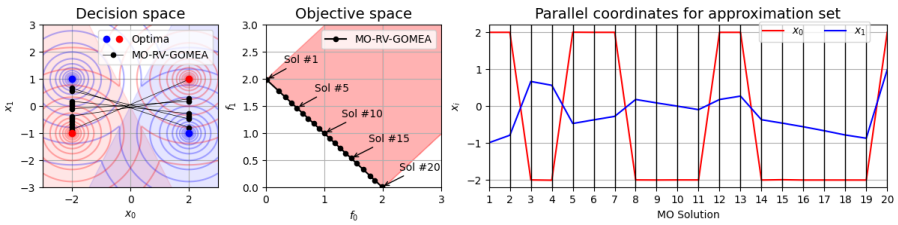


Fig. 1: Approximation set and front with parallel coordinates plot as produced by MO-RV-GOMEA on the MinDist problem: $f_0(\boldsymbol{x}) = min(||\boldsymbol{x} - [1, -1]||, ||\boldsymbol{x} - [-1, 1]||)$ and $f_1(\boldsymbol{x}) = min(||\boldsymbol{x} - [1, 1]||, ||\boldsymbol{x} - [-1, -1]||)$. Shaded blue and red regions correspond to niches with global PSs.

Most MMOEAs do not explicitly model multiple approximation sets, but include diversity preserving techniques to ensure that solutions from multiple niches are maintained. The result of these MMOEAs is usually given in the form of a single approximation front, often derived from (a subset of) the elitist archive. A decision maker can then investigate this front by traversing the solutions for desired trade-offs. However, the underlying solutions are taken from several distinct niches, which could result in observing a counterintuitive change in decision variable values when navigating the approximation front. Decision makers then might have to investigate all solutions before a correct choice can be made [21, 24]. Figure 1 shows such an approximation front and set that contains solutions from both modes on the MinDist problem [23] and demonstrates the counterintuitive changes in decision variable values in the parallel coordi-

nates plot (i.e., if one were to navigate the front by traversing and inspecting the solutions from one extreme to the other). It shows that in objective space a front is found that looks to have approximated the PF to (near) optimality, but the solutions jump around throughout decision space as seen in the parallel coordinates plot.

The issue of counterintuitive navigation along the approximation front has also been explored in recent work, which introduced a new indicator-based MOEA for bi-objective optimization called BezEA [24]. A new problem formulation for population-based MOEAs was introduced whereby they parameterized approximation sets as Bézier curves. This formulation ensures the navigational smoothness of an approximation set, whilst still being able to find good approximation sets when using the HyperVolume indicator (HV) [37]. By design, BezEA disallowes curves to dominate parts of themselves to ensure that the approximation set constitutes a single niche in the landscape.

Recent work that included the concept of niching showed promising results in maintaining multiple approximation sets in a population-based MMOEA called the MO-HillVallEA [23]. The authors extended the concept of Hill-Valley Clustering (HVC) [26] for MOPs to Multi-Objective HVC (MO-HVC) for MMOPs. MO-HillVallEA was found to be capable of finding and preserving approximation sets, one for each niche, in parallel over time by considering Pareto domination per niche. However, MO-HillVallEA produces approximation sets that are not inherently smooth due to slight oscillations around the PS.

In this work, the notions of niching through HVC and Bézier curve parameterizations are combined. The use of niching allows to effectively search the multi-modal landscape. The use of Bézier curve parameterizations not only enforces the smooth and intuitive navigability that is desired by decision makers, but also enforces each approximation set to be within a single niche. Furthermore, the use of the HV indicator allows closer convergence to the PS as compared to the Pareto dominance-based algorithms [4]. The new algorithm that we propose is called Multi-Modal Bézier Evolutionary Algorithm (MM-BezEA). The purpose of MM-BezEA is to find all approximation sets for a given MMOP, where each approximation set consists of solutions from a single mode.

In order to combine the techniques of Bézier curve parameterizations and HVC into the proposed algorithm, several contributions are made. First, the problem of how to niche approximation sets in the form of Bézier curve parameterizations is resolved. Second, initialization of approximation sets within a single niche is enabled, as otherwise, clustering becomes ambiguous if these approximation sets span multiple niches.

2 Bézier parameterizations

One of the key features of the newly proposed algorithm is that Bézier parameterizations are used as approximation sets for bi-objective optimization [24]. This allows the algorithm to model the approximation set as a smooth curve in decision space.

2.1 Definition of solution set

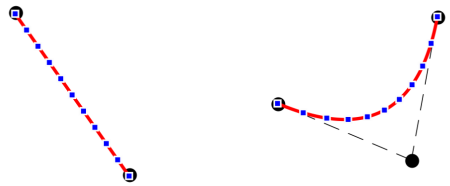
An ℓ-dimensional Bézier curve $\mathbf{B}(t; C_q)$ can be defined using $q \geq 2$ control points c_i in an ordered set $C_q = \{\boldsymbol{c}_1, ..., \boldsymbol{c}_q\}$, where ℓ is the problem dimensionality and $\boldsymbol{c}_i \in \mathbb{R}^\ell$. The full notation is:

$$\mathbf{B}(t; C_q) = \sum_{i=1}^{q} \binom{q-1}{i-1} (1-t)^{q-i-1} t^{i-1} \boldsymbol{c}_i \text{ for } 0 \leq t \leq 1 \tag{1}$$

The endpoints of the Bézier curve are always defined by the first and last control points, whilst the other control points are normally not located on the curve. A solution set of given size p, $S_{p,q}(C_q) = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_p\}$ with $\boldsymbol{x}_i \in \mathbb{R}^\ell$, can now be parameterized by a Bézier curve by selecting an evenly spread set of p points $\boldsymbol{x}_i$. Figure 2 visualizes two solution sets $S_{p,q}(C_q)$ parameterized by Bézier curves. The solution set $S_{p,q}(C_q)$ is formally defined as:

$$S_{p,q}\left(C_q\right) = \left\{ \mathbf{B}\left(\frac{0}{p-1}; C_q\right), \mathbf{B}\left(\frac{1}{p-1}; C_q\right), ..., \mathbf{B}\left(\frac{p-1}{p-1}; C_q\right) \right\} \tag{2}$$

$S_{p,q}(C_q)$ is parameterized for (M)MOEAs by taking the concatenation of the decision variables in the set C_q as a solution [24, 5]. This results in a solution being of the form $[\boldsymbol{c}_1, ..., \boldsymbol{c}_q] \in \mathbb{R}^{q*\ell}$.



$\mathbf{B}(t; \mathcal{C}_2) = (1-t)\mathbf{c}_1 + t\mathbf{c}_2$ $\mathbf{B}(t; \mathcal{C}_3) = (1-t)^2\mathbf{c}_1 + 2(1-t)t\mathbf{c}_2 + t^2\mathbf{c}_3$

Fig. 2: Bézier curves with $q \in \{2, 3\}$ control points in black. Interpolated curve in red with the $p = 11$ points in blue evenly spread in the domain of t along the curve [24].

2.2 Evaluation

To evaluate a solution set $S_{p,q}(C_q)$, a number of new functions were previously introduced [24]. These functions are briefly explained in the following paragraphs. Figure 3 illustrates these functions to give the reader a more graphical indication.

A new function $A^{nb}\left(S_{p,q}\right)$ has been introduced that calculates a navigational Bézier (nb) order o_{nb}. This order is defined as starting from the best solution for objective f_0 to the best solution in f_1. All solutions that are dominated by other solutions on the curve, are omitted from the subset that defines the navigational order. An approximation set $\mathcal{A}_{p,q,o_{nb}}$ is the resulting subset of $S_{p,q}(C_q)$, with only the solution indices as specified in o_{nb}. The quality of the approximation set $\mathcal{A}_{p,q,o_{nb}}$ can now be evaluated, e.g., with the HV indicator [37].

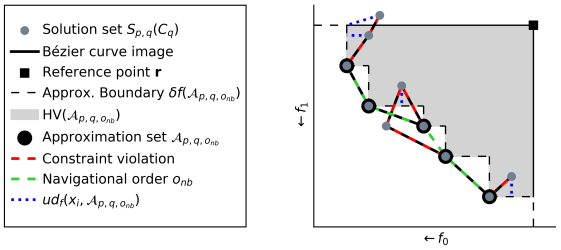


Fig. 3: Evaluation of Bézier parameterizations [24]

A new constraint function $C(\mathcal{S}_{p,q}, o_{nb})$ was also introduced. It is employed in order to not only push all dominated solutions on the curve towards the undominated region of the search space, but also to prevent the curve from intersecting itself in objective space. This may for instance happen if a curve stretches across two local PSs, which is not preferential. The constraint function uses the uncrowded distance metric $ud_f(\boldsymbol{x}_i, \mathcal{A})$ [33], which measures the Euclidean distance from a dominated point $\boldsymbol{x}_i$ to the approximation boundary $\delta f(\mathcal{A}_{p,q,o_{nb}})$ in objective space. Furthermore, to further increase pressure towards the unfolding of Bézier curves in objective space, all dominated solutions and those not in $\mathcal{A}_{p,q,o_{nb}}$ are pulled towards their neighbouring solutions on the Bézier curve by taking the Euclidean distance in objective space between these solution and their neighbours as an additional constraint value. All dominated solutions from $S_{p,q}(C_q)$ now have their uncrowded distance values and the Euclidean distances in objective space to neighbours of those not in $\mathcal{A}_{p,q,o_{nb}}$ summed up as a constraint for the total solution set. In combination with constraint domination [10], this constraint pushes all solutions along the Bézier curve towards the undominated region.

3 Niching methods

To enable the algorithm proposed in this paper, i.e., MM-BezEA, to effectively search the multi-modal landscape, several previously introduced niching methods are used and combined. These are employed in order to extend the uni-modal search that is originally performed by BezEA. As the number of modes is usually unidentified beforehand, the algorithm needs to be able to adapt to the number of modes present in an MMOP.

3.1 HVC and MO HVC

HVC is a so-called two-stage niching approach that clusters and evolves the population for multi-modal single-objective optimization problems. In each generation, the first stage is used to locate each of the distinct niches, for each of which a core search algorithm is initialized in the second stage.

At the heart of the HVC approach is the Hill-Valley Test (HVT) [34], which can be utilized to determine whether two solutions reside in the same niche. This test first determines an edge between two solutions $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the search space. Along this edge, N_t evenly spread points are evaluated, determined by the distance between the two solutions divided by the expected edge length. If any

of these N_t test points have a fitness that is worse than that of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, the test detects that there is a *hill* in between them. Consequently, the two solutions are to be put in separate clusters. On the other hand, if all N_t points have equal or better fitness values than both $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, these two solutions belong to the same *valley* and are to be clustered together. In order to determine in which order the solutions are to be clustered (i.e., undergo the HVT), the concept of the nearest better tree [27] is employed.

The MO-HillVallEA algorithm [23] expands on the previous HVC approach in the form of MO-HVC. It uses the same concept of the HVT, but now performs clustering for each of the m objective functions separately, which results in m cluster sets of the population. To obtain a single cluster set, the intersection of all clustering sets is taken, similar to the colored regions of Figure 1.

3.2 Restart scheme with elitist archive

Various algorithms implemented a form of a restart scheme whereby the population size is increased over time. Examples of such schemes are the interleaved multistart scheme [8, 16] and the restart-Covariance Matrix Adaptation Evolution Strategy with Increased Population (IPOP-CMA-ES) algorithm [3]. In HillVallEA [26], an elitist archive is combined with a restart scheme, where the population size is doubled after each restart as in IPOP-CMA-ES [3]. By employing the HVT to check if a solution resides in another niche, the elitist archive of HillVallEA is capable of holding the elites for each of the modes, despite it being developed for single objective problems.

To prevent HillVallEA from revisiting already searched modes, it makes use of the elitist archive, which is inspired by the repelling subpopulations (RS-CMSA) algorithm [1] that defines taboo regions close to elites. The steps taken to discard the regions of the search space, for which an elite was already found in one of the earlier populations, start with adding the elites to the population of the current restart. Then, all solutions are clustered using HVC, followed by discarding all clusters that have one of the elites as their best solution. As a result of discarding these regions of the search space, more attention is given to undiscovered parts of the search space after each restart.

4 Multi Modal-Bézier Evolutionary Algorithm

In this section MM-BezEA is described. MM-BezEA is comprised of a combination and modification of techniques described in the previous sections. The most notable of the modifications are the adjustments implemented in HVC in order to apply it to Bézier curve parameterizations, as well as the initialization of approximation sets within niches.

4.1 Clustering approximation sets

The Bézier curves are evaluated using the uncrowded HV measure [25]. Since this is a scalar, the HVC approach seems to intuitively allow the clustering of

single-objective problems. However, the approximation set $\mathcal{A}_{p,q,o_{nb}}$ that is used in the HV calculation only considers the undominated indices of the Bézier solution set $S_{p,q}(C_q)$ as defined in o_{nb}. Hence, the objective value of a solution set $S_{p,q}(C_q)$ seems highly dependant on how many dominated solutions there are on the Bézier curve due to its orientation and length in decision space.

To enable the clustering of Bézier solution sets $S_{p,q}(C_q)$, the idea behind MO-HVC can be used on the set of control points C_q, as each of these is a single solution as normally defined in MO optimization. Also, since a solution set is defined to be deteriorating in f_0 and improving in f_1 according to o_{nb}, the order of the control points is inverted if $f_0(\boldsymbol{c}_1) < f_0(\boldsymbol{c}_q)$ does not hold [24]. Accordingly, the i-th Bézier solution can be designated to be in the same niche as the j-th Bézier solution if their control points $\boldsymbol{c}_l^i \in C_q^i$ and $\boldsymbol{c}_l^j \in C_q^j$ for $l = \{1, ..., q\}$ are in the same niche. In a general sense, the same HVC approach as used in HillVallEA is used, but inspiration has been taken from the MO-HVC approach to produce a new test for Bézier solution sets, which is shown in Algorithm 1.

Algorithm 1: $[B]$ = Bezier-HillValleyTest($\mathbf{S_i}$, $\mathbf{S_j}$, N_t, f)

Input: Solutions sets $\mathbf{S_i}$, $\mathbf{S_j}$, int N_t, objective functions $f_0, ..., f_{m-1}$
Output : Whether $\mathbf{S_i}$ and $\mathbf{S_j}$ belong to the same niche
for $l = 1, ..., q$ **do**
  $\boldsymbol{c}_{i,l}, \boldsymbol{c}_{j,l} \leftarrow$ control point l of $\mathbf{S_i}$, control point l of $\mathbf{S_j}$
  `// Check if` $\boldsymbol{c}_{i,l}$ `and` $\boldsymbol{c}_{j,l}$ `are in same niche for all` m `objectives`
  for $k = 0, ..., m-1$ **do**
    if *HillValleyTest*($\boldsymbol{c}_{i,l}$, $\boldsymbol{c}_{j,l}$, N_t, f_k) **then return** false
return true

4.2 Initialization within niches

The original BezEA algorithm initializes all solution sets by sampling from a uniform distribution over the search space. As it is an MOEA that was not designed for multi-modal optimization, the uniform initialization allows solution sets to be initialized within or in between any niche(s). Clustering these solutions with the newly introduced Bézier HVT will result in finding a large number of separate niches as each control point has to be in the same mode. To prevent this, a new initialization method for Bezier solution sets is proposed to enforce their initialization within a niche. First, an iteration of MO-HVC is run on a set of $q * N$ solutions, N being the population size, that is sampled from a uniform distribution over the search space, where the resulting clusters include all of the $\boldsymbol{x}_{test}$ solutions resulting from applying the Hill-Valley Test. Second, selection is performed on the clusters proportional to their size in order to reduce their combined size, with test solutions, down to $q * N$. Lastly, Bézier solution sets $S_{p,q}(C_q)$ are initialized by randomly choosing q solutions as the control

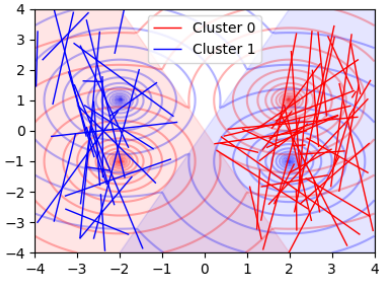


Fig. 4: Initialization of Bézier solutions ($q = 2$) for MinDist

points from one single cluster $\mathcal{C}$ as produced by MO-HVC. The result can be seen in Figure 4 in the case of the example problem MinDist.

4.3 Algorithm overview

MM-BezEA has a similar structure as the restart scheme in HillVallEA [26] that is described in section 3.2. Every iteration, the combination of initialization of Bézier curves and dismissal of previously optimized clusters with an elite as their best solution takes place. For each of the resulting niches, a core search algorithm is run for one generation, which in the case of MM-BezEA is the RV-GOMEA algorithm [7] that is also used in BezEA. At the end of each generation, the Bézier HVT of Algorithm 1 is used in the HVC step. This step takes all solutions originating from all clusters and clusters them again for the next generation. In between generations, the notion of cluster registration [6] is used on the cluster mean closest in decision space to transfer the parameters for RV-GOMEA between the clusters of each generation. An overview of the algorithm in the form of pseudocode is given in Algorithm 2.

Algorithm 2: $[\mathbb{E}]$ = MM-BezEA(...)

Input: MO function f, popsize N, test points p, control points q, budget
Output : Elitist archive $\mathbb{E} = [\mathcal{E}_0, \mathcal{E}_1, ...]$

```
E = {}
while budget remaining do
    P_mo = UniformSampling(q · N, f)
    C_mo = MO-HillValleyClustering(P_mo, f)
    C = InitializeBezierSolutions(C_mo, q, p, f)
    C = RemoveElitesFrom(C)
    while budget remaining do
        P = E
        for C_i ∈ C do
            O_i = core_search_algorithm(C_i)
            P = P ∪ O_i
        C_prev = C
        C = BezierHillValleyClustering(P, f)
        E = ConstructElitistArchive(C, E)
        C = RemoveElitesFrom(C)
        C = ClusterRegistration(C, C_prev)
```

5 Experiments

MM-BezEA is empirically benchmarked on several test problems. The results are compared to MO-HillVallEA [23], MO-RV-GOMEA [8], and MO_Ring_PSO_SCD [35]. MO_Ring_PSO_SCD is implemented through the PlatEMO framework [32], together with a manual implementation of the used metrics and problems. For the other algorithms, original C++ implementations are used.

5.1 Test problems

Several test problems are employed. First of these is the MinDist problem [23] that was described in the introduction, where linear PSs are to be found. The other employed functions are frequently used in literature, namely OmniTest [12], Two on One [28], and Sympart {1,2,3} [29]. Lastly, several problems are

taken from the Multi-modal Multi-objective test Function (MMF) benchmark suite [20] in the form of MMF {1, 2, 12, 14, 15}. A mix of PS and PF shapes have been chosen to determine the capabilities of MM-BezEA on different problem types. Table 1 shows some of the important characteristics for each of the problems. For all problems with a configurable number of PSs, it is set to 2, likewise the problem dimension ℓ is fixed to 2. In order to determine the values of the performance indicators, the reference PSs will be made using 5000 points that adhere to the analytical formulas describing the PSs. In the case of Two on One, a very close approximation is used [28].

Table 1: Bi-objective problem characteristics.

Problem	ℓ	PS	PS Shape	PF Shape
MinDist	$[2, \infty) \in \mathbb{Z}$	n	Linear	Convex
Omni Test	$[2, \infty) \in \mathbb{Z}$	3^ℓ	Linear	Convex
Two on One	2	2	Linear	Convex
Sympart 1, 2	2	9	Linear	Convex
Sympart 3	2	9	Non-linear	Convex
MMF 1, 2	2	2	Non-linear	Convex
MMF 12	$[2, \infty) \in \mathbb{Z}$	n	Linear	Disconnected
MMF 14, 15	$[2, \infty) \in \mathbb{Z}$	n	Linear	Concave

5.2 Benchmark setup

In order to get a fair comparison, each of the algorithms will be given an equally sized budget of $200,000$ function evaluations for each of the problems. This removes the influence of the used programming languages, as the computation time is not limited. The parameters of MO-RV-GOMEA, MO_Ring_PSO_SCD, and MO-HillVallEA are set to the values reported in relevant literature. Furthermore, for each problem and metric, the average over 31 runs will be taken.

The elitist archives sizes $N_{\mathbb{E}}$ are set to be 1250 for MO-RV-GOMEA and MO-HillVallEA. The population size N is set to 96 for MO-RV-GOMEA and 250 for MO-HillVallEA [23]. MO-RV-GOMEA uses a linkage tree as its linkage model, with a total of 5 clusters. For MO_Ring_PSO_SCD the population size is 800 [35]. For the MM-BezEA algorithm, the number of control points q for each approximation set is set to 2. Just like for the original BezEA algorithm, MM-BezEA is given population sizes of 76 [24]. The number of test points p is set to 7.

5.3 Performance indicators

The HV indicator [37] is used to see how well the algorithms perform in getting close the PF. As a result of the use of test points in MM-BezEA, the Bézier solutions sets will have a limited amount of points in the approximation set that can be used to calculate the HV values. Therefore, a subset of the approximation set will be taken for the other algorithms to allow a fair comparison based on the HV indicator. Specifically, the same number of test points is selected for a fair comparison by means of greedy Hypervolume Subset Selection (gHSS) [15].

We further use a relatively new performance indicator for multi-modal multi objective optimization, named Pareto Set Proximity (PSP) [35]. It is an indicator that determines how well all PSs are approximated by taking the Cover Rate

(CR), that shows how well the extremes of all PSs are captured, divided by the Inverted Generational Distance in decision space (IGDX) [36], which can be used to determine how close the approximation sets are to the PSs. For the IGDX measure, the approximation sets as produced by MM-BezEA are interpolated by taking 1000 intermediate points before determining the IGDX value. This can be performed relatively easily as interpolating these parameterizations does not require any extra fitness evaluations.

Finally, we use a performance indicator regarding smoothness, for which we follow the definition as introduced in the work on BezEA [24]. It captures how smooth an approximation set can be navigated in terms of decision variables by measuring the detour length in decision space when traversing the approximation set from one solution to the next via an intermediate solution, as compared to going to the next solution directly. The smoothness approaches its maximum value of 1 if all solutions would be colinear in decision space, where the lowest possible value is 0. In cases where multiple approximation sets are explicitely determined, like in MO-HillVallEA and MM-BezEA, the average smoothness over all clusters will be taken. In the other cases the smoothness over the entire approximation set is taken.

5.4 Results

Table 2 shows the results for all problems and algorithms per indicator.

The HV results clearly show that all algorithms are capable of performing nearly equally in obtaining a good approximation front. However, MM-BezEA with $q = 2$ does deteriorate in performance on the MMF1 and 2 problems that have non-linear PSs. The deterioration is inherently caused by the chosen parameterizations that create approximation sets which are linear in shape. Another problem instance where a smaller HV for the new algorithm is obtained, is that of MMF12. Here, despite MM-BezEA obtaining the best PSP values, the approximation sets did not fully approximate the actual PSs and did not cover the endpoints.

The PSP indicator shows similar results, except that MO-RV-GOMEA performs worse as it is not an MMOEA and therefore does not explicitly search for multiple niches. Again promising results for MM-BezEA are shown in cases where linear PSs can be found, as seen in Figure 5a where MM-BezEA approximates all 9 Pareto sets very well. In the problems with non-linear PSs, MO-HillVallEA and MO_Ring_PSO_SCD manage to find better approximations.

The smoothness results show, as intended, that the chosen parameterizations inherently cause smooth approximation sets with a perfect smoothness of 1.0 for MM-BezEA. Other algorithms do not obtain this, except for MO-RV-GOMEA on 2 of the 11 problems. A visualization of the results of MM-BezEA on the MinDist problem is given in Figure 5b. This figure depicts the smooth progression of the decision variables values in the parallel coordinates plot for the rightmost approximation set in decision space. It contrasts sharply to the parallel coordinates plot of Figure 1 when navigating the approximation set as it now shows a smooth course of the decision variable values.

Table 2: Results (avg. (± st.dev.)) per problem and algorithm over 31 runs, bold identifies best result with statistical significance (Wilcoxon rank-sum test with $\alpha = 0.05$ and Holm-Bonferroni correction).

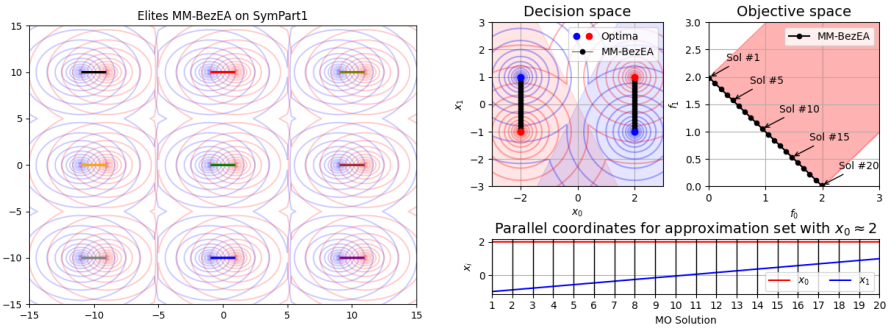
	Problem	MM-BezEA	MO-HillVallEA	MO_Ring_PSO_SCD	MO-RV-GOMEA
HV	MinDist	**1.17e+2 (±9.43e-5)**	1.17e+2 (±1.62e-2)	1.17e+2 (±5.95e-3)	1.17e+2 (±1.36e-3)
	OmniTest	**8.48e+0 (±2.18e-6)**	8.47e+0 (±1.96e-3)	8.47e+0 (±7.34e-4)	8.47e+0 (±4.82e-4)
	Sympart 1	**1.17e+2 (±2.16e-5)**	1.17e+2 (±1.42e-2)	1.17e+2 (±7.64e-3)	1.17e+2 (±1.30e-3)
	Sympart 2	**1.17e+2 (±2.91e-5)**	1.17e+2 (±7.77e-3)	1.17e+2 (±8.75e-3)	1.17e+2 (±4.47e-3)
	Sympart 3	**1.17e+2 (±9.61e-5)**	1.17e+2 (±1.61e-2)	1.17e+2 (±9.21e-3)	1.17e+2 (±4.91e-3)
	TwoOnOne	1.13e+2 (±1.33e-4)	1.13e+2 (±2.39e-4)	**1.13e+2 (±1.82e-4)**	1.13e+2 (±1.10e-4)
	MMF 1	6.04e-1 (±3.86e-2)	**8.05e-1 (±2.37e-4)**	**8.05e-1 (±8.69e-5)**	**8.05e-1 (±6.60e-5)**
	MMF 2	6.34e-1 (±2.01e-4)	8.04e-1 (±6.90e-4)	8.04e-1 (±9.59e-4)	**8.05e-1 (±1.75e-4)**
	MMF 12	1.78e+0 (±2.02e-6)	2.06e+0 (±2.57e-3)	2.06e+0 (±2.05e-3)	**2.06e+0 (±1.49e-4)**
	MMF 14	5.63e+0 (±1.33e-5)	**5.63e+0 (±7.26e-4)**	5.63e+0 (±1.92e-3)	5.63e+0 (±2.23e-4)
	MMF 15	5.56e+0 (±2.03e-2)	**5.57e+0 (±6.52e-4)**	5.56e+0 (±1.54e-3)	**5.57e+0 (±1.79e-4)**
PSP	MinDist	**3.26e+2 (±7.31e+1)**	5.02e+1 (±2.74e+0)	6.97e+1 (±6.73e+0)	1.21e+0 (±1.65e+0)
	OmniTest	**2.36e+2 (±8.81e+1)**	7.13e+1 (±2.76e+0)	6.90e+1 (±9.56e+0)	1.36e-1 (±2.43e-1)
	Sympart 1	**2.67e+2 (±1.21e+2)**	3.56e+1 (±1.57e+0)	2.79e+1 (±3.69e+0)	1.16e-2 (±2.70e-2)
	Sympart 2	**3.09e+2 (±8.46e+1)**	3.60e+1 (±7.42e-1)	2.38e+1 (±2.46e+0)	1.02e-2 (±1.71e-2)
	Sympart 3	6.41e+1 (±7.77e+1)	**4.33e+1 (±2.10e+0)**	2.64e+1 (±5.62e+0)	8.09e-3 (±1.33e-2)
	TwoOnOne	**3.04e+2 (±2.18e+2)**	4.50e+1 (±7.21e-1)	2.45e+1 (±1.03e+1)	2.68e+0 (±1.11e+0)
	MMF 1	7.22e+0 (±2.41e+0)	3.17e+1 (±6.84e-1)	**3.80e+1 (±6.79e+0)**	1.02e+0 (±2.89e-1)
	MMF 2	4.00e+0 (±2.49e+0)	**1.17e+2 (±1.21e+1)**	5.04e+1 (±1.51e+1)	2.18e+0 (±1.41e+0)
	MMF 12	**2.42e+1 (±7.72e+0)**	1.94e+1 (±6.46e+0)	1.53e+1 (±1.46e-1)	8.67e+0 (±1.36e+0)
	MMF 14	**2.68e+3 (±4.82e+2)**	3.70e+2 (±1.03e+1)	2.31e+2 (±2.16e+1)	1.08e+0 (±2.84e+0)
	MMF 15	**2.73e+2 (±4.17e+1)**	2.65e+2 (±4.58e+0)	2.44e+2 (±7.25e+0)	2.24e+1 (±1.40e-2)
Smoothness	MinDist	**1.00e+0 (±0.00e+0)**	8.09e-1 (±3.70e-2)	7.63e-1 (±5.91e-3)	8.94e-1 (±1.81e-1)
	OmniTest	**1.00e+0 (±0.00e+0)**	9.23e-1 (±5.65e-3)	7.28e-1 (±9.76e-3)	7.16e-1 (±1.96e-1)
	Sympart 1	**1.00e+0 (±0.00e+0)**	8.76e-1 (±2.52e-2)	6.84e-1 (±9.76e-3)	7.01e-1 (±2.09e-1)
	Sympart 2	**1.00e+0 (±0.00e+0)**	8.78e-1 (±2.14e-2)	5.73e-1 (±9.30e-3)	7.87e-1 (±1.68e-1)
	Sympart 3	**1.00e+0 (±0.00e+0)**	8.70e-1 (±3.10e-2)	5.29e-1 (±1.18e-2)	8.30e-1 (±1.92e-1)
	TwoOnOne	**1.00e+0 (±0.00e+0)**	7.77e-1 (±1.21e-2)	7.47e-1 (±9.36e-3)	7.51e-1 (±1.59e-1)
	MMF 1	**1.00e+0 (±0.00e+0)**	9.01e-1 (±3.98e-2)	7.82e-1 (±8.49e-3)	5.86e-1 (±9.13e-2)
	MMF 2	**1.00e+0 (±0.00e+0)**	9.38e-1 (±1.26e-2)	5.01e-1 (±8.92e-3)	8.46e-1 (±1.60e-1)
	MMF 12	**1.00e+0 (±0.00e+0)**	8.35e-1 (±2.73e-2)	6.35e-1 (±9.86e-3)	**1.00e+0 (±0.00e+0)**
	MMF 14	**1.00e+0 (±0.00e+0)**	9.31e-1 (±1.17e-2)	8.71e-1 (±1.11e-2)	9.62e-1 (±1.05e-1)
	MMF 15	**1.00e+0 (±0.00e+0)**	9.19e-1 (±1.43e-2)	8.63e-1 (±8.54e-3)	**1.00e+0 (±0.00e+0)**

6 Discussion

MM-BezEA did not cover the endpoints of the PSs in the case of MMF12. This can be caused by the fact that the Bézier fitness function will constrain a solution when one of its control points is dominated in objective space by one of the test points. As the endpoints of each part of the discontinuous PF are close to being dominated, i.e., close to the constraint space, it can lead to not entirely capturing the discontinuous pieces of the PF and thus resulting in a lower HV.

The HV indicator is a Pareto compliant indicator [13, 38], but it does suffer from a downside. In some situations the endpoints of the approximation sets cannot reach the endpoints of the Pareto set because the distribution of points that maximizes the hypervolume does not include the extreme solutions. Even when the number of test points will be set to infinity, the reference point can never be set so that the extremes are captured [2].

Even though the smoothness indicator tries to determine whether an approximation set is smooth by measuring the detour length, it comes down to determining the angle between neighboring solutions. This implies that it only



(a) All approximation sets produced by MM-BezEA on SymPart 1 [29]

(b) Approximation sets and front with parallel coordinates plot for one of the approximation sets produced by MM-BezEA on MinDist [23]

Fig. 5: Visualization of results

considers linear curves to be perfectly smooth, where there is a straight angle between solutions. When the number of solutions in an approximation set increases, the average distance between the solutions in objective space decreases. As a result of the lower distances and ever so slight oscillations around the PF, the angle between solutions decreases due to which the smoothness indicator will report low smoothness values. In cases where the niches can be separated in a good manner, another definition of smoothness that considers the oscillation around the PS might be more useful.

Future work could investigate the further use of the Bézier parameterizations with more control points to approximate non-linear Pareto sets. Note that the definitions given in this paper already allow for this. Furthermore, no limit on the number of approximation sets can currently be set, which degrades the quality of the approximation sets in highly multi-modal problems as the population is then divided over all niches through HVC [23]. Finally, Bézier simplexes [18] might be usable for problems with more than two objectives.

7 Conclusion

We proposed the algorithm MM-BezEA to search for multiple parameterized approximation sets that define smooth curves in the decision space for bi-objective multi-modal optimization problems. The results show that MM-BezEA is competently capable of locating all modes in a multi-modal landscape as exemplified in various benchmark problems and that the smoothness is indeed enforced by the Bézier parameterizations. Furthermore, MM-BezEA significantly outperformed other algorithms in problems with linear Pareto sets, but was outperformed in problems with non-linear Pareto sets. However, only low-order Bézier curves were used in our experiments, and these results may well be different if higher order curves were used, which the definitions in this paper readily allow.

References

1. Ahrari, A., Deb, K., Preuss, M.: Multimodal optimization by covariance matrix self-adaptation evolution strategy with repelling subpopulations. Evolutionary Computation **25**(3), 439–471 (Sep 2017). https://doi.org/10.1162/evco_a_00182
2. Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: Optimal μ-distributions and the choice of the reference point. In: Proceedings of the tenth ACM SIGEVO workshop on Foundations of genertic algorithms. p. 87–102. FOGA '09, Association for Computing Machinery, New York, NY, USA (2009). https://doi.org/10.1145/1527125.1527138
3. Auger, A., Hansen, N.: A restart CMA evolution strategy with increasing population size. In: 2005 IEEE Congress on Evolutionary Computation. vol. 2, pp. 1769–1776 Vol. 2. IEEE, New York, NY, USA (2005). https://doi.org/10.1109/CEC.2005.1554902
4. Berghammer, R., Friedrich, T., Neumann, F.: Convergence of set-based multi-objective optimization, indicators and deteriorative cycles. Theoretical Computer Science **456**, 2–17 (Oct 2012). https://doi.org/10.1016/J.TCS.2012.05.036
5. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: Multi-objective selection based on dominated hypervolume. European Journal of Operational Research **181**(3), 1653–1669 (Sep 2007). https://doi.org/https://doi.org/10.1016/j.ejor.2006.08.008
6. Bosman, P.A.N.: The anticipated mean shift and cluster registration in mixture-based EDAs for multi-objective optimization. In: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation. p. 351–358. GECCO '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1830483.1830549
7. Bouter, A., Alderliesten, T., Witteveen, C., Bosman, P.A.N.: Exploiting linkage information in real-valued optimization with the real-valued gene-pool optimal mixing evolutionary algorithm. In: Proceedings of the Genetic and Evolutionary Computation Conference. p. 705–712. GECCO '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3071178.3071272
8. Bouter, A., Luong, N.H., Witteveen, C., Alderliesten, T., Bosman, P.A.N.: The multi-objective real-valued gene-pool optimal mixing evolutionary algorithm. In: Proceedings of the Genetic and Evolutionary Computation Conference. p. 537–544. GECCO '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3071178.3071274
9. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation **6**, 182–197 (4 2002). https://doi.org/10.1109/4235.996017
10. Deb, K.: An efficient constraint handling method for genetic algorithms. Computer Methods in Applied Mechanics and Engineering **186**(2), 311–338 (Jun 2000). https://doi.org/10.1016/S0045-7825(99)00389-8
11. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. John Wiley & Sons, Inc., USA (2001)
12. Deb, K., Tiwari, S.: Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization. European Journal of Operational Research **185**(3), 1062–1087 (Mar 2008). https://doi.org/https://doi.org/10.1016/j.ejor.2006.06.042
13. Fleischer, M.: The measure of Pareto optima applications to multi-objective metaheuristics. In: Proceedings of the 2nd International Conference on Evolutionary

Multi-Criterion Optimization. p. 519–533. EMO'03, Springer-Verlag, Berlin, Heidelberg (Apr 2003)
14. Grimme, C., Kerschke, P., Aspar, P., Trautmann, H., Preuss, M., Deutz, A.H., Wang, H., Emmerich, M.: Peeking beyond peaks: Challenges and research potentials of continuous multimodal multi-objective optimization. Computers & Operations Research **136**, 105489 (2021). https://doi.org/https://doi.org/10.1016/j.cor.2021.105489
15. Guerreiro, A.P., Fonseca, C.M., Paquete, L.: Greedy hypervolume subset selection in low dimensions. Evolutionary Computation **24**, 521–544 (Sep 2016). https://doi.org/10.1162/EVCO_a_00188
16. Harik, G.R., Lobo, F.G.: A parameter-less genetic algorithm. In: Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation - Volume 1. p. 258–265. GECCO'99, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1999)
17. Igel, C., Hansen, N., Roth, S.: Covariance matrix adaptation for multi-objective optimization. Evolutionary Computation **15**, 1–28 (3 2007). https://doi.org/10.1162/evco.2007.15.1.1
18. Kobayashi, K., Hamada, N., Sannai, A., Tanaka, A., Bannai, K., Sugiyama, M.: Bézier simplex fitting: Describing Pareto fronts of simplicial problems with small samples in multi-objective optimization. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, the 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019. pp. 2304–2313. AAAI press, Palo Alto, CA, USA (Jan 2019)
19. Li, X., Epitropakis, M.G., Deb, K., Engelbrecht, A.: Seeking multiple solutions: An updated survey on niching methods and their applications. IEEE Transactions on Evolutionary Computation **21**(4), 518–538 (Aug 2017). https://doi.org/10.1109/TEVC.2016.2638437
20. Liang, J., Qu, B., Gong, D., Yue, C.: Problem definitions and evaluation criteria for the cec 2019 special session on multimodal multiobjective optimization. Tech. rep. (Nov 2018)
21. Luong, N.H., Alderliesten, T., Bel, A., Niatsetski, Y., Bosman, P.A.N.: Application and benchmarking of multi-objective evolutionary algorithms on high-dose-rate brachytherapy planning for prostate cancer treatment. Swarm and Evolutionary Computation **40**, 37–52 (6 2018). https://doi.org/10.1016/j.swevo.2017.12.003
22. Mahfoud, S.W.: Niching Methods for Genetic Algorithms. Ph.D. thesis, University of Illinois at Urbana-Champaign, USA (1996), uMI Order No. GAX95-43663
23. Maree, S.C., Alderliesten, T., Bosman, P.A.N.: Real-valued evolutionary multi-modal multi-objective optimization by hill-valley clustering. In: Proceedings of the Genetic and Evolutionary Computation Conference. p. 568–576. GECCO '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3321707.3321759
24. Maree, S.C., Alderliesten, T., Bosman, P.A.N.: Ensuring smoothly navigable approximation sets by Bézier curve parameterizations in evolutionary bi-objective optimization. In: Parallel Problem Solving from Nature – PPSN XVI. pp. 215–228. Springer International Publishing, Cham (2020)
25. Maree, S.C., Alderliesten, T., Bosman, P.A.N.: Uncrowded Hypervolume-based Multi-objective Optimization with Gene-pool Optimal Mixing. Evolutionary Computation pp. 1–24 (Dec 2021). https://doi.org/10.1162/evco_a_00303

26. Maree, S.C., Alderliesten, T., Thierens, D., Bosman, P.A.N.: Real-valued evolutionary multi-modal optimization driven by hill-valley clustering. In: Proceedings of the Genetic and Evolutionary Computation Conference. p. 857–864. GECCO '18, Association for Computing Machinery, New York, NY, USA (Jul 2018). https://doi.org/10.1145/3205455.3205477
27. Preuss, M.: Niching the CMA-ES via nearest-better clustering. In: Proceedings of the 12th Annual Conference Companion on Genetic and Evolutionary Computation. p. 1711–1718. GECCO '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1830761.1830793
28. Preuss, M., Naujoks, B., Rudolph, G.: Pareto set and EMOA behavior for simple multimodal multiobjective functions. In: Parallel Problem Solving from Nature – PPSN IX. p. 513–522. PPSN'06, Springer-Verlag, Berlin, Heidelberg (2006). https://doi.org/10.1007/11844297_52
29. Rudolph, G., Naujoks, B., Preuss, M.: Capabilities of EMOA to detect and preserve equivalent pareto subsets. In: Evolutionary Multi-Criterion Optimization, 4th International Conference, EMO 2007, Matsushima, Japan, March 5-8, 2007, Proceedings. Lecture Notes in Computer Science, vol. 4403, pp. 36–50. Springer, Berlin, Heidelberg (2006). https://doi.org/10.1007/978-3-540-70928-2_7
30. Schäpermeier, L., Grimme, C., Kerschke, P.: To boldly show what no one has seen before: A dashboard for visualizing multi-objective landscapes. In: Ishibuchi, H., Zhang, Q., Cheng, R., Li, K., Li, H., Wang, H., Zhou, A. (eds.) Evolutionary Multi-Criterion Optimization. pp. 632–644. Springer International Publishing, Cham (2021)
31. Tanabe, R., Ishibuchi, H.: A review of evolutionary multimodal multiobjective optimization. IEEE Transactions on Evolutionary Computation **24**(1), 193–200 (2020). https://doi.org/10.1109/TEVC.2019.2909744
32. Tian, Y., Cheng, R., Zhang, X., Jin, Y.: PlatEMO: A MATLAB platform for evolutionary multi-objective optimization. IEEE Computational Intelligence Magazine **12**(4), 73–87 (Nov 2017)
33. Touré, C., Hansen, N., Auger, A., Brockhoff, D.: Uncrowded hypervolume improvement: COMO-CMA-ES and the sofomore framework. In: Proceedings of the Genetic and Evolutionary Computation Conference. p. 638–646. GECCO '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3321707.3321852
34. Ursem, R.: Multinational evolutionary algorithms. In: Proceedings of the 1999 Congress on Evolutionary Computation-CEC99. vol. 3, pp. 1633–1640 Vol. 3. IEEE, New York, NY, USA (1999). https://doi.org/10.1109/CEC.1999.785470
35. Yue, C., Qu, B., Liang, J.: A multi-objective particle swarm optimizer using ring topology for solving multimodal multi-objective problems. IEEE Transactions on Evolutionary Computation **22**, 805–817 (Sep 2017). https://doi.org/10.1109/TEVC.2017.2754271
36. Zhang, Q., Li, H.: MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Transactions on Evolutionary Computation **11**, 712–731 (12 2007). https://doi.org/10.1109/TEVC.2007.892759
37. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems. pp. 95–100. International Center for Numerical Methods in Engineering, Athens, Greece (Sep 2001)

38. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G.: Performance assessment of multiobjective optimizers: an analysis and review. IEEE Transactions on Evolutionary Computation **7**(2), 117–132 (May 2003). https://doi.org/10.1109/TEVC.2003.810758