Many-Objective Evolutionary Algorithms: Objective Reduction, Decomposition and Multi-Modality

January, 2021

A thesis submitted to the *Indian Statistical Institute*

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Computer Science)

Monalisa Pal

Machine Intelligence Unit,

Indian Statistical Institute,

203 Barrackpore Trunk Road,

Kolkata - 700108, India

Under the supervison of

Prof. (Dr.) Sanghamitra Bandyopadhyay

To my family

Acknowledgments

Gratitude and regards go beyond words to *Prof. Sanghamitra Bandyopadhyay*, who has supervised my dissertation work and has kept me motivated in difficult times. I consider myself extremely lucky to have her as my mentor. She has patiently provided the utmost care for my work and promptly responded to my queries throughout the last few years.

Heartfelt thanks are due to Prof. Ujjwal Maulik, for his invaluable advice and help in my research understanding. I am thankful to Dr. Sriparna Saha and Mr. Raunak Sengupta of IIT Patna for their constructive feedback, which formed the stepping stones during my early Ph.D. career.

I sincerely acknowledge Prof. Stéphane Ploix of the G-SCOP lab at the Grenoble Institute of Technology, Grenoble, France, and Dr. Corinne Touati of the LIG lab at INRIA, Grenoble Rhône Alpes, France, for allowing me two academic visits to their labs for learning things from diverse domains. I appreciate the financial assistance provided to me for these visits by the Indian side of the Indo-French Centre for the Promotion of Advanced Research (IFCPAR/CEFIPRA) under the project sanctioned vide DST-INRIA/2015-02/BIDEE/0978. I am also immensely grateful to Dr. Amr Alzouhri Alyafi, Dr. Manar Amayri, Dr. Khadija Tijani, Dr. Mahendra Singh, and Dr. Lisa Scanu for their hospitable welcome and academic suggestions during these visits. A special note of thanks to Prof. Patrick Reignier of the LIG lab at INRIA, Grenoble Rhône Alpes, for spending time listening and commenting on my studies.

I appreciate the research fellowship (Regn. No. JCSK-CC-0094) awarded to me by the Indian Statistical Institute for smoothly conducting my research without any financial constraints. I am also grateful to Prof. Bimal K. Roy and Prof. Sanghamitra Bandyopadhyay, the honorable directors of the Indian Statistical Institute, for providing an unparalleled research infrastructure during the period of my doctoral studies. Heartfelt thanks are due to all the faculties of the Indian Statistical Institute. My work could not have progressed smoothly without their constant support and invaluable guidance. For maintaining a congenial atmosphere throughout my research tenure, I admire my colleagues and labmates. Special thanks go to Ms. Aparajita Khan, Ms. Sampa Misra, Ms. Rishika Sen, Ms. Indrani Ray, Dr. Losiana Nayak, Dr. Debarka Sengupta, Dr. Angana Chakraborty, Ms. Garisha Chowdhary, Mr. Dip Ghosh, Mr. Debajyoti Sinha, Ms. Snehalika Lall, Mr. Sourav Biswas, Ms. Suparna Saha, and Ms. Sucheta Dawn. Missing anyone's name in the above list does not mean denying anyone's help, rather my inability to express gratitude to all others in a confined space. I acknowledge the CSSC for providing computing facilities, the ISI library for providing reference materials, the reprography unit for careful photocopying, the office staff of the Machine Intelligence Unit for providing a friendly environment, and the authorities of ISI for extending various facilities.

I shall forever remain indebted to my beloved parents (Mrs. Mita Pal and Dr. Tapas Kumar Pal) and grandparents (Late Mrs. Gita Palit and Late Mr. Balaram Palit) for guiding me through every step of my life. Their blessings, love, constant encouragement, and unconditional support have helped me face all the challenges in life and complete my research work.

ISI, Kolkata — Monalisa Pal

2021 — (*Monalisa Pal*)

Related Publications by the Author

*** Papers in Peer-reviewed Journals**

J1. M. Pal, S. Saha and S. Bandyopadhyay, (2018) "DECOR: Differential Evolution using Clustering based Objective Reduction for many-objective optimization", *Information Sciences*, Vol. 423, pp. 200–218, Elsevier. ISSN: 0020–0255, DOI: 10.1016/j.ins.2017.09.051.

J2. M. Pal and S. Bandyopadhyay, (2019) "ESOEA: Ensemble of Single Objective Evolutionary Algorithms for Many-Objective Optimization", *Swarm and Evolutionary Computation* (*Special Issue on Differential Evolution*), Vol. 50, 100511, Elsevier. ISSN: 2210–6502, DOI: 10.1016/j.swevo.2019.03.006.

J3. R. Sengupta, M. Pal, S. Saha and S. Bandyopadhyay, (2019) "NAEMO: Neighborhood-sensitive Archived Evolutionary Many-objective Optimization Algorithm", *Swarm and Evolutionary Computation* (*Special Issue on Nature Inspired Optimization Algorithms: Recent Advances and Applications*), Vol. 46, pp. 201–218, Elsevier. ISSN: 2210–6502, DOI:10.1016/j.swevo.2018.12.002.

J4. M. Pal, A. A. Alyafi, S. Ploix, P. Reignier, and S. Bandyopadyay, (2019) "Unmasking the causal relationships latent in the interplay between occupant's actions and indoor ambience: a building energy management outlook", *Applied Energy*, Vol. 238, pp. 1452–1470, Elsevier. ISSN: 0306–2619, DOI: 10.1016/j.apenergy.2019.01/118.

J5. M. Pal and S. Bandyopadhyay, (2021) "Decomposition in Decision and Objective Space for Multi-Modal Multi-Objective Optimization", *Swarm and Evolutionary Computation*, Elsevier. ISSN: 2210–6502. (accepted, in press)

*** Papers in Peer-reviewed Book Chapters**

B1. M. Pal, A. A. Alyafi, S. Bandyopadhyay, S. Ploix and P. Reignier, (2018) "Enhancing Comfort of Occupants in Energy Buildings", in: S. Kar, U. Maulik, X. Li

(Eds.), *Operations Research and Optimization, FOTA 2016, Springer Proceedings in Mathematics & Statistics*, Vol. 225, pp. 133–144, Springer, Singapore. ISBN: 978-981-10-7813-2, DOI: 10.1007/978-981-10-7814-9_10.

B2. R. Sengupta, M. Pal, S. Saha and S. Bandyopadhyay, (2019) "Population Dynamics Indicators for Evolutionary Many-Objective Optimization", in: C. Panigrahi, A. Pujari, S. Misra, B. Pati, K.-C. Li (Eds.) *Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing*, Vol. 714, pp. 261–271, Springer, Singapore. ISBN: 978-981-13-0223-7, DOI: 10.1007/978-981-13-0224-4_24.

B3. M. Pal and S. Bandyopadhyay, (2021) "Multi-Modality of Occupants' Actions for Multi-Objective Building Energy Management", in: S. Bhattacharyya, P. Dutta, K. Datta (Eds.) *Intelligence Enabled Research: DoSIER 2020, Advances in Intelligent Systems and Computing*, pp. 11–19, Springer, Singapore. ISBN: 978-981-15-9289-8, DOI: 10.1007/978-981-15-9290-4_2.

*** Papers in Peer-reviewed Proceedings of International Conferences**

C1. M. Pal and S. Bandyopadhyay, (28-30 January 2016) "Reliability of Convergence Metric and Hypervolume Indicator for Many-Objective Optimization", in: *2^{nd} International Conference on Control, Instrumentation, Energy & Communication* (*CIEC 2016*), Kolkata, India, pp. 511–515. IEEE. DOI: 10.1109/CIEC.2016.7513806.

C2. M. Pal, S. Saha and S. Bandyopadhyay, (24-30 July 2016) "Clustering based Online Automatic Objective Reduction to aid Many-Objective Optimization", in: *IEEE Congress on Evolutionary Computation* (*CEC 2016*), Vancouver, Canada, pp. 1131–1138. IEEE. DOI: 10.1109/CEC.2016.7743915.

C3. A. A. Alyafi, M. Pal, S. Ploix, P. Reignier, and S. Bandyopadyay, (18-20 July 2017) "Differential Explanations for Energy Management in Buildings", in: *IEEE Technically Sponsored - SAI Computing Conference 2017*, London, United Kingdom, pp. 507–516. IEEE. DOI: 10.1109/SAI.2017.8252144.

C4. M. Pal, R. Sengupta, S. Bandyopadhyay, A. A. Alyafi, S. Ploix, P. Reignier and S. Saha, (27-30 December 2017) "Analysis of Optimizers to Regulate Occupant's

Actions for Building Energy Management", in: *Ninth International Conference on Advances in Pattern Recognition* (*ICAPR 2017*), Bangalore, India, pp. 1–6. IEEE. DOI: 10.1109/ICAPR.2017.8593024.

C5. M. Pal and S. Bandyopadhyay, (18-21 November 2018) "Consensus of Subjective Preferences of Multiple Occupants for Building Energy Management", in: *Symposium Series on Computational Intelligence 2018* (*SSCI 2018*), Bangalore, India, pp. 1815–1822. IEEE. DOI: 10.1109/SSCI.2018.8628670.

C6. M. Pal and S. Bandyopadhyay, (13-17 July 2019) "Differential Evolution for Multi-Modal Multi-Objective Problems", in: *Genetic and Evolutionary Computation Conference Companion Proceedings 2019* (*GECCO 2019*), Prague, Czech Republic, pp. 1399–1406. ACM. DOI: 10.1145/3319619.3326862.

Abstract

Evolutionary Algorithms (EAs) for Many-Objective Optimization (MaOO) problems are challenging in nature due to the requirement of large population size, difficulty in maintaining the selection pressure towards global optima and inability of accurate visualization of high-dimensional Pareto-optimal Set (in decision space) and Pareto-Front (in objective space). The quality of the estimated set of Pareto-optimal solutions, resulting from the EAs for MaOO problems, is assessed in terms of proximity to the true surface (convergence) and uniformity and coverage of the estimated set over the true surface (diversity). With more number of objectives, the challenges become more profound. Thus, better strategies have to be devised to formulate novel evolutionary frameworks for ensuring good performance across a wide range of problem characteristics.

In this thesis, the first work adopts the strategy of objective reduction to present the framework of DECOR, which handles MaOO problems through correlation-based clustering by eliminating the less conflicting objectives. While DECOR demonstrates an enhanced convergence, it reveals the necessity of better solution diversity for resembling the true surface. In the second work, ESOEA is presented, which decomposes the objective space for the collaborative optimization of multiple sub-populations. It also adaptively feedbacks the sub-population size to redistribute the solutions for the effective exploration of difficult regions in the fitness landscape. While ESOEA demonstrates enormous improvement in performance over a variety of MaOO problems, lack of theoretical foundations hinders the analysis of its properties. In the third work, the neighborhood property arising out of sub-space formation (in objective space) is recognized and used to present the framework of NAEMO. It not only demonstrates improved performance but also guarantees monotonically improving diversity, theoretically. While such reference vector assisted decomposition-based approaches are useful for good performance in the objective space, it innately neglects the solution distribution in the decision space. This behavior is disadvan-

tageous for multi-modal problems (multiple alternative subsets within the Pareto-optimal Set independently mapping to the entire Pareto-Front). Hence, in the fourth work, the decomposition in objective space is amalgamated with graph Laplacian based clustering in the decision space to present the framework of LORD. Finally, to establish the efficacy on a real-world problem, NAEMO and LORD are customized to address the multi-modal many-objective building energy management problem. Moreover, four decision-making strategies are presented to select one of the Pareto-optimal solutions as the most relevant solution for implementation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Three fundamental questions are asked for any problem: whether a solution exists, how many solutions exist and which is the *best* solution for the given objective(s) [141]. Answering the third question is an optimization problem. Although numerous numerical optimization methods exist, Evolutionary Algorithms (EAs) are more popular due to their benefits [32, 161] such as the ability to provide a decent solution approximation to problems unsolvable by numerical optimization (hard problems and black-box problems), invariance to continuity and convexity of the landscape, and parallel search in multiple directions by intelligent use of a population of solutions.

The problems with multiple conflicting objectives are known as Multi-objective Optimization Problems (MOPs) and the EAs used to address them are known as Multi-Objective Evolutionary Algorithms (MOEAs) [60,85]. Formally, an M-objective minimization[1] problem is defined as follows:

$$\begin{aligned} &\text{Minimize: } \mathbf{F}(\mathbf{X}) = [f_1(\mathbf{X}), f_2(\mathbf{X}), \cdots, f_M(\mathbf{X})] \text{ where } \mathbf{X} = [x_1, \cdots, x_N] \\ &\text{subjected to, } x_i^L \leq x_i \leq x_i^U, \text{ for } i = 1, 2, \cdots, N, \\ &g_j(\mathbf{X}) \geq 0, \text{ for } j = 1, 2, \cdots, K_{ieq} \text{ and } h_k(\mathbf{X}) = 0, \text{ for } k = 1, 2, \cdots, K_{eq}. \end{aligned} \tag{1.1}$$

The N-dimensional search space is the region consisting of the intersection of the constrained regions defined by the K_{ieq} inequality and K_{eq} equality constraints, and the

[1] Without loss of generality, minimization problems are considered throughout this thesis. Also, as a symbolic representation rule in this thesis, bold math variables denote an array/vector/set of scalars, calligraphic math variables denote a matrix/set of arrays and usual math variables denote scalar quantities.

lower (x_i^L) and upper bounds (x_i^U) for all the i^{th} decision variables.

In this class of optimization problems, when number of objectives is four or more (i.e., $M > 3$), several challenges come into play. Hence, this sub-class of problems forms an essential research topic and is called Many-objective Optimization Problems (MaOPs) [85, 106]. Thus, the EAs used for addressing MaOPs are known as Many-Objective Evolutionary Algorithms (MaOEAs). Some practical applications of MaOEAs from varied domains are in nurse scheduling problem [131], factory-shed truss design problem [10], space trajectory design problem [88], pattern recognition problems [31,136], software refactoring problem [126], building energy management problem [133] and cyclone geometry design problem [55].

Rest of this chapter is structured as follows. In Section 1.2, the basic concepts of MOEAs are outlined. In Section 1.3, the various research areas of this domain are briefly described, including the well-explored and scarcely-explored areas. Thereafter, the application domain of building energy management is briefly introduced, which is explored in a later chapter to demonstrate the challenges of a real-world many-objective optimization problem. Finally, the goals and scope of this thesis are stated in Section 1.5.

1.2 Key Concepts

This section gives a brief background of various concepts and terminologies required for overall understanding of this thesis.

1.2.1 A Box-Constrained Multi-objective Optimization Problem

The mathematical formulation of a box-constrained multi-objective minimizations problem presents the mapping from an N-dimensional vector ($\mathbf{X}$) in the decision space ($\mathcal{D}$) to an M-dimensional vector ($\mathbf{F}(\mathbf{X})$) in the objective space [22, 106] as follows:

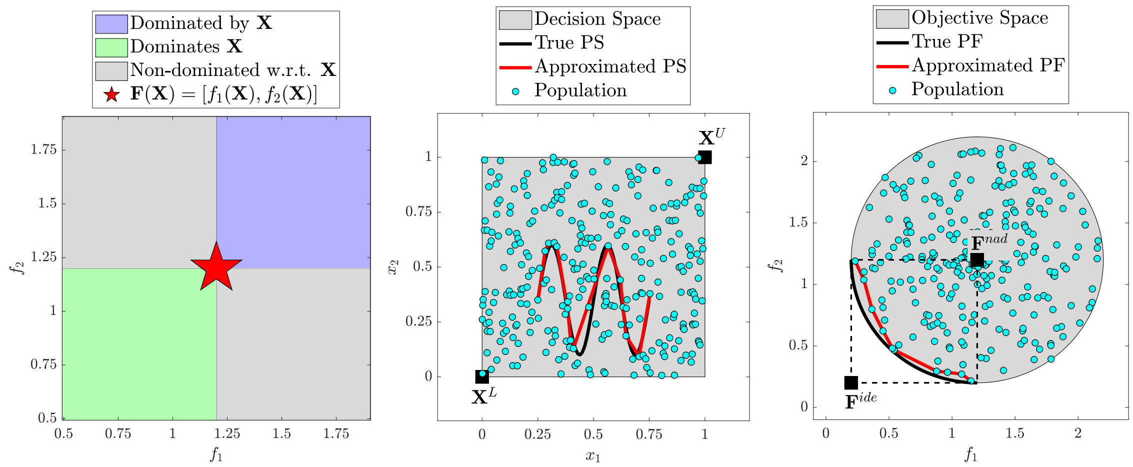
$$\begin{aligned} &\text{Minimize: } \mathbf{F}(\mathbf{X}) = [f_1(\mathbf{X}), f_2(\mathbf{X}), \cdots, f_M(\mathbf{X})] \text{ where, } \mathbf{X} \in \mathcal{D}\left(\subseteq \mathbb{R}^N\right), \\ &\mathbf{F}(\mathbf{X}) : \mathcal{D} \mapsto \mathbb{R}^M \text{ and } \mathcal{D} : x_i^L \le x_i \le x_i^U, \text{ for } i = 1, 2, \cdots, N. \end{aligned} \tag{1.2}$$

1.2.2 Pareto-Dominance Relation

The concept of trade-off (Pareto-optimality) can be formally mentioned as a state where further improvement in one of the objectives only leads to the deterioration in terms of the other objective(s) [10, 85]. Pareto-dominance relation is used to compare two N-dimensional decision vectors. Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be two feasible points, then $\mathbf{X}_1 \prec \mathbf{X}_2$ or $\mathbf{X}_1$ Pareto-dominates $\mathbf{X}_2$ according to the following condition:

$$\forall i \in \{1, \cdots, M\},\ f_i(\mathbf{X}_1) \le f_i(\mathbf{X}_2) \text{ and } \exists j \in \{1, \cdots, M\},\ f_j(\mathbf{X}_1) < f_j(\mathbf{X}_2). \tag{1.3}$$

In other words, $\mathbf{X}_1$ is better than $\mathbf{X}_2$ if $\mathbf{X}_1$ is as good as $\mathbf{X}_2$ in all objectives and at least better in one of the objectives.



(a) Partitions by $\mathbf{F}(\mathbf{X})$ (b) Decision Space (c) Objective Space

Figure 1.1: Illustration of the key concepts of a multi-objective minimization problem with $M = 2$, $N = 2$, $\mathbf{X}^L = [0, 0]$, $\mathbf{X}^U = [1, 1]$.

Based on Pareto-dominance, there can be the following three cases (Fig. 1.1a) when a solution $\mathbf{X}$ is compared with other solutions:

- Some of the other solutions are dominated by $\mathbf{X}$ (blue region).
- Some of the other solutions dominate $\mathbf{X}$ (green region).
- Some of the other solutions are non-dominated with respect to $\mathbf{X}$ (gray regions).

1.2.3 Non-Dominated Solution Set

A solution $\mathbf{X}^\star \in \mathcal{S}_{box}$ (where $\mathcal{S}_{box}$ is a well-defined search space) is said to be a non-dominated solution, if there is no other solution $\mathbf{X} \in \mathcal{S}_{box}$ dominating $\mathbf{X}^\star$. A non-dominated set of solutions is the set of all such $\mathbf{X}^\star$, defined as follows:

$$ndset\,(\mathcal{S}_{box}) = \{\mathbf{X}^\star |\, (\nexists \mathbf{X} \in \mathcal{S}_{box} | \mathbf{X} \prec \mathbf{X}^\star)\}. \tag{1.4}$$

1.2.4 Pareto-optimal Set and Pareto-Front

The Pareto-optimal Set (PS) is the set of solution vectors in the decision space ($\mathcal{D}$) such that there is no other solution that dominates any point of this set, i.e., the non-dominated set of solutions in $\mathcal{S}_{box} = \mathcal{D}$ is the PS. The image of PS in the objective space is known as the Pareto-Front (PF). Thus, PS and PF are defined as follows:

$$\text{PS: } \{\mathbf{X}^\star \in \mathcal{D} \mid (\nexists \mathbf{X} \in \mathcal{D} \mid \mathbf{X} \prec \mathbf{X}^\star)\} = ndset(\mathcal{D}), \tag{1.5}$$

$$\text{PF: } \{\mathbf{Y} \in \mathbb{R}^M \mid \mathbf{Y} = \mathbf{F}(\mathbf{X}), \mathbf{X} \in \text{PS}\}. \tag{1.6}$$

For any given MOPs or MaOPs, an EA yields an approximation of PS and PF at termination as illustrated in Fig. 1.1b and Fig. 1.1c. A better approximation will indicate a better performing ability of the EA for that MOP or MaOP.

1.2.5 Non-Conflicting Objective Set

Among the M objectives, any two objectives $f_i(.)$ and $f_j(.)$ are said to be conflicting in nature, if the following condition [10, 141] is satisfied:

$$\exists\, (\mathbf{X}_1, \mathbf{X}_2) \in \mathcal{D} \times \mathcal{D} \text{ such that } \left(f_i\left(\mathbf{X}_1\right) > f_i\left(\mathbf{X}_2\right)\right) \text{ and } \left(f_j\left(\mathbf{X}_1\right) < f_j\left(\mathbf{X}_2\right)\right). \tag{1.7}$$

The concept of induced Pareto-dominance has been introduced later. Induced Pareto-dominance declares two objective sets $\mathbf{F}_i(.)$ and $\mathbf{F}_j(.)$ to be conflicting if $\preceq_{\mathbf{F}_i} \neq \preceq_{\mathbf{F}_j}$ where $\preceq_{\mathbf{F}_i}$ and $\preceq_{\mathbf{F}_j}$ denote the induced Pareto-dominance by the objective sets $\mathbf{F}_i$ and $\mathbf{F}_j$, respectively. The objective set $\mathbf{F}'$ is called Non-Conflicting Objective Set [18] when $\mathbf{F}' \subseteq \mathbf{F}$, $2 \leq |\mathbf{F}'| \leq |\mathbf{F}|$ and $\preceq_{\mathbf{F}'} = \preceq_{\mathbf{F}}$ so that excluding the objectives $\mathbf{F} - \mathbf{F}'$ does not change the induced Pareto-dominance. The purpose of objective reduction is to find the smallest Non-Conflicting Objective Set.

1.2.6 Ideal and Nadir Objective Vectors

The ideal and nadir points are essential concepts, which are required in several operations to define an algorithmic framework for an M-objective optimization problem.

Let $\mathbf{X}$ be a feasible solution, then the ideal objective vector $\mathbf{F}^{ide}$ is defined as follows:

$$\mathbf{F}^{ide} = \left[f_1^{ide}, f_2^{ide}, \cdots, f_i^{ide}, \cdots, f_M^{ide}\right] \text{ where } f_i^{ide} = \min_{\mathbf{X}\in\mathcal{D}} f_i\left(\mathbf{X}\right). \tag{1.8}$$

Thus, $\mathbf{F}^{ide}$ coincides with the point in the objective space defined by the true optimal values of all objective functions, considered one at a time. The ideal objective vector is not a feasible objective vector.

Let $\mathbf{X}$ be a solution in PS, then the nadir objective vector $\mathbf{F}^{nad}$ is given as follows:

$$\mathbf{F}^{nad} = \left[f_1^{nad}, f_2^{nad}, \cdots, f_i^{nad}, \cdots, f_M^{nad}\right] \text{ where } f_i^{nad} = \max_{\mathbf{X}\in\mathrm{PS}} f_i\left(\mathbf{X}\right). \tag{1.9}$$

Thus, $\mathbf{F}^{nad}$ coincides with the point in the objective space defined by the worst value along each of the objectives in PF. It should be noted that unlike Eq. (1.8), for nadir point, $\mathbf{X}$ is chosen from PS and not from $\mathcal{D}$. Depending on the convexity and the continuity of PF, the nadir objective vector may or may not be a feasible objective vector. An example of the ideal and the nadir objective vectors are illustrated in Fig. 1.1c.

1.3 A Brief Overview of Research Areas

There is an abundance of MOEAs in the literature, some of the notable ones being Non-dominated Sorting Genetic Algorithm (NSGA-II) [47], Strength Pareto EA (SPEA2) [201], Pareto-Envelop based Selection Algorithm (PESA-II) [38] and Differential Evolution for Multi-objective Optimization (DEMO) [153]. However, when these algorithms are applied to MaOPs, several issues appear [32, 109, 142, 197]. These major challenges, along with a few practical caveats, are summarized as follows:

1. With an increase in M, the population gets saturated with non-dominated solutions at very early generations leading to a decrease in selection pressure [10, 25, 32, 85].
2. With an increase in M, the curse of dimensionality [10, 34, 82, 109] appears, i.e., the requirement of a large number of solutions to approximate the PF and to balance the trade-off between convergence and diversity.
3. With an increase in M, visualization issues [22, 32, 109, 138] appear, which makes it difficult to validate the search behavior of the MOEAs, the quality of the approximated PF and the choice of the final solution.

4. While assessment of MOEAs rely on different performance indicators, various indicators capture different representative characteristics of the PF like convergence, diversity, coverage, or a combination of two or more of these attributes.

For tackling these issues of applying EAs to MaOPs and the issues arising out of practical challenges, research studies are broadly categorized into several areas, which are described in the following sub-sections.

1.3.1 Algorithmic Design

There are broadly four classes of algorithms: Pareto-dominance based algorithms, indicator based algorithms, objective reduction based algorithms and reference-vector assisted decomposition based algorithms.

Pareto-dominance based Algorithms

The first class involves modification of the Pareto-dominance relationship to enhance the selection pressure such as ε-dominance [46], θ-dominance [187], favour relation [54], fuzzy Pareto-dominance [69] and grid dominance [184]. A few such tailored Pareto-dominance based MaOEAs are Grid-dominance based EA (GrEA) [184], θ-Dominance based EA (θ-DEA) [187], Approximation-Guided Evolutionary algorithm (AGE-II) [179] and Knee-point driven EA (KnEA) [192].

Indicator based Algorithms

The second class considers convergence and diversity indicators as selection criteria. Common indicator based algorithms are Indicator-Based EA (IBEA) [200], S-Metric Selection based Evolutionary Multi-Objective Algorithm (SMS-EMOA) [12], Generational Distance and ε-dominance based Multi-Objective EA (GDE-MOEA) [124], improved version of Many-Objective Metaheuristic Based on R2-Indicator (MOMBI-II) [61] and algorithm based on Hypervolume Estimation (HypE) [9]. Hypervolume indicator has gained immense attention due to its success, though its computational complexity increases exponentially with the number of objectives. There has been some effort towards generalization of hypervolume for MaOPs such as by using Monte-Carlo simulation [9] or weakly Pareto-compliant Sharpe-Ratio indicator [62].

Objective Reduction based Algorithms

The third class transforms MaOPs into simpler problems by reducing the number of objectives so that the induced PS remains invariant [10, 142]. Hence, it combines dimensionality reduction techniques like principal component analysis [157], clustering based approaches [10, 142], feature selection [86, 87], and many more, in a framework, to deal with MaOPs. The intuition behind such approaches is to reduce the problem complexities such that the MaOPs could efficiently be handled by existing MOEAs. However, investigating the optimal objective subset is tedious, albeit essential for every new problem.

Reference-vector Assisted Decomposition based Algorithms

The fourth class involves decomposition of MOPs or MaOPs into multiple scalar optimization sub-problems which collaborate with each other to be optimized. Some notable algorithms of this class are Multi-Objective EA based on Decomposition (MOEA/D) [150], Evolutionary Dynamic Weighted Aggregation (EDWA) [97], second version of Multiple Single Objective Pareto Sampling algorithm (MSOPS-II) [75] and Multi-Objective Genetic Local Search (MOGLS) [80]. Two recent and successful approaches of this class are MOEA/D-M2M (transforms MOPs into simpler multi-objective sub-problems) [115, 117] and the third version of NSGA (NSGA-III, uses reference points to enhance diversity) [45].

Decomposition based algorithms have shown promising performance in addressing MaOPs. Moreover, such approaches neither suffer from the reduced selection pressure in high dimensional objective space like Pareto-dominance based approaches nor require the extreme computational effort for hypervolume evaluation.

1.3.2 Benchmark Test Problems

For establishing the efficacy of MOEAs, the performance of the algorithms is compared on various benchmark functions which try to simulate real-world problem difficulties. These MOPs differ in problem characteristics [161] such as:

- *Geometry*: Shape of the PF can be convex, concave, linear, mixed, degenerate,
- *Parameter Dependencies*: Separable objectives, non-separable objectives (capability to determine ideal points by considering only one objective at a time),
- *Bias*: Presence of bias (like variable density of solutions) while mapping solutions

from decision space to fitness functions in objective space,

- *Many-to-One mappings*: Pareto one-to-one, Pareto many-to-one, flat regions, isolated optima and
- *Modality*: Uni-modal, multi-modal.

Several benchmark functions are available in the literature [50, 74, 76, 112, 115, 191] and new benchmark functions with added difficulties are also being developed. The definitions of various test problems used in this thesis are provided in Appendix A.

1.3.3 Performance Indices

The objective space cannot be directly visualized when the number of objectives is greater than three. Thus, the performance of the optimization algorithm has to be assessed in terms of the performance indicators. Two criteria which are looked into while assessing the effectiveness of the approximated PF are convergence [24, 47] and diversity [9, 10]. Convergence is the proximity of the approximated PF to the true PF, while diversity is the uniformity in the spread of the solutions in the approximated PF over the true PF.

Several popular evaluation metrics, available in literature, are convergence metric [10, 47], Inverted Generational Distance (IGD) [14] and its variants [78, 198], hypervolume indicator [9, 10], R2 indicator [16], purity metric [11], crowding distance [47], minimal-spacing [10, 11], minimum of sum of objective values (denoting convergence towards the center of true PF) [85], sum of minimum objective values attained along each dimension (denoting convergence along the edges of the true PF) [85] and range of objective values along each dimension (denoting diversity) [85]. These performance metrics operate in the objective space. For assessing the performance of MaOEAs in the decision space, only a handful of metrics are available in literature which includes IGD in decision space [169,188], pareto-set proximity [188] and mixture of IGD in objective and decision space (IGDM) [118]. For the different works presented in this thesis, one or more of those performance measures are considered which are described in the next few paragraphs.

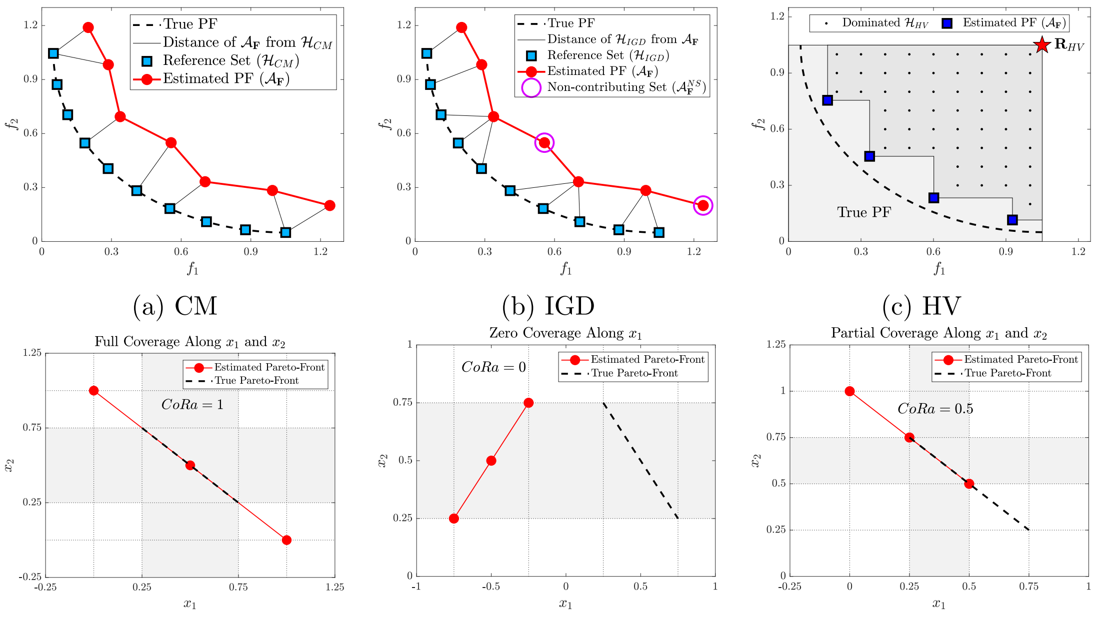
Convergence Metric

Convergence metric (CM) or generational distance [10, 47] indicates the convergence of the approximated PF and is given as follows:

$$CM(\mathcal{A}_{\mathbf{F}}, \mathcal{H}_{CM}) = \frac{1}{|\mathcal{A}_{\mathbf{F}}|} \sum_{i=1}^{|\mathcal{A}_{\mathbf{F}}|} \left(\min_{j=1}^{|\mathcal{H}_{CM}|} \left(D_E \left(\mathbf{F}\left(\mathbf{X}_i\right), \mathbf{H}_j \right) \right) \right), \tag{1.10}$$

where $\mathbf{F}(\mathbf{X}_i) \in \mathcal{A}_{\mathbf{F}}$ and $\mathbf{H}_j \in \mathcal{H}_{CM}$.

In Eq. (1.10), the non-dominated set of solutions approximating the PF is denoted by $\mathcal{A}_{\mathbf{F}}$. For evaluating CM, the knowledge of the true PF is required. To represent the true PF, either several points are sampled uniformly across the surface of the true PF or the intersection points are chosen where the true PF and the reference-vectors (defined by [40]) coincide. This set of points representing the true PF is denoted by $\mathcal{H}_{CM}$. As illustrated in Fig. 1.2a, convergence metric (CM) is estimated as the sample mean of the minimum Euclidean distance $D_E(.)$ of the objective vectors ($\mathbf{F}(\mathbf{X}_i)$) constituting $\mathcal{A}_{\mathbf{F}}$ from the points ($\mathbf{H}_j$) in $\mathcal{H}_{CM}$, over the number of solutions in $\mathcal{A}_{\mathbf{F}}$. Given the same $\mathcal{H}_{CM}$ with the same $N_{CM} = |\mathcal{H}_{CM}|$, among two approximated PFs, the one having a smaller value of convergence metric has a better convergence to the true PF.



(a) CM (b) IGD (c) HV

(d) Full Coverage (e) No Coverage (f) Partial Coverage

Figure 1.2: Illustration for evaluating some performance indices.

Although the convergence metric is a vital performance measure, it suffers from the following two drawbacks [134]:

- For evaluating convergence by Eq. (1.10), defining $\mathcal{H}_{CM}$ requires the knowledge of the true PF which is unavailable for practical problems.
- The value of the convergence metric lies in the range $[0, \infty)$. Thus, without field

knowledge or unless compared with the convergence value of another solution set, it becomes difficult to assert how far the approximated PF is from the true PF.

Inverted Generational Distance

Inverted Generational Distance (IGD) [14, 109] gives an indication of the convergence as well as the diversity of the approximated PF and is obtained as follows:

$$IGD(\mathcal{A}_{\mathbf{F}}, \mathcal{H}_{IGD}) = \frac{1}{|\mathcal{H}_{IGD}|} \sum_{j=1}^{|\mathcal{H}_{IGD}|} \left(\min_{i=1}^{|\mathcal{A}_{\mathbf{F}}|} \left(D_E \left(\mathbf{F}\left(\mathbf{X}_i\right), \mathbf{H}_j \right) \right) \right), \tag{1.11}$$
$$\text{where } \mathbf{F}\left(\mathbf{X}_i\right) \in \mathcal{A}_{\mathbf{F}} \text{ and } \mathbf{H}_j \in \mathcal{H}_{IGD}.$$

In Eq. (1.11), the non-dominated set of solutions approximating the PF is denoted by $\mathcal{A}_{\mathbf{F}}$ and those approximating the PS is denoted by $\mathcal{A}_{\mathbf{X}}$. Similar to the evaluation of convergence metric, IGD also requires a set $\mathcal{H}_{IGD}$ with representative points from the true PF (if evaluated in the objective space). As illustrated in Fig. 1.2b, IGD is estimated as the sample mean of the minimum Euclidean distance $D_E(.)$ of the points ($\mathbf{H}_j$) in $\mathcal{H}_{IGD}$ from $\mathbf{F}(\mathbf{X}_i) \in \mathcal{A}_{\mathbf{F}}$, over the number of solutions in $\mathcal{H}_{IGD}$. If IGD is evaluated in the decision space $\mathcal{H}_{IGD}$ is a representation of the true PS and instead of $\mathbf{F}(\mathbf{X}_i) \in \mathcal{A}_{\mathbf{F}}$, $\mathbf{X}_i \in \mathcal{A}_{\mathbf{X}}$ is considered in Eq. (1.11). Given the same $\mathcal{H}_{IGD}$ with the same $N_{IGD} = |\mathcal{H}_{IGD}|$, among two approximated PFs, the one having a smaller value of IGD has a better convergence, or a better diversity or both with respect to the true PF.

As this indicator is computationally similar to the convergence metric, it shares the same drawbacks as those of the convergence metric along with the following ones:

- IGD is known to yield Pareto non-compliant results [78]. However, this drawback has been eliminated in the weakly Pareto-compliant, IGD+ metric [78].
- IGD is hugely influenced by the size of the reference set ($N_{IGD} = |\mathcal{H}_{IGD}|$).

In this thesis, IGD represents the performance in objective space. However, when IGD is also used to assess the performance in decision space, for distinction IGDF is used to represent IGD in objective space and IGDX is used to represent IGD in decision space.

Hypervolume Indicator

Hypervolume indicator [9, 10] can also represent both convergence and diversity information using a single value and also its evaluation is independent of the knowledge of the true PF. For its evaluation, a hyper-rectangle is considered between a reference point ($\mathbf{R}_{HV} = [r_{HV,1}, \cdots, r_{HV,M}]$) and the origin of the objective space. The hypervolume ($HV = volume\left(\cup_{\mathbf{F}\in\mathcal{A}_\mathbf{F}} [f_1, r_{HV,1}] \times \cdots \times [f_M, r_{HV,M}]\right)$ with $\mathbf{F} = [f_1, \cdots, f_M]$) indicates the dominated region of this hyper-rectangle. For fair comparison, the following approaches [81, 134] guides the placement of the reference point ($\mathbf{R}_{HV}$) for constructing the hyper-rectangle:

- The most naïve approach considers $\mathbf{R}_{HV}$ to be user-defined over the different estimations of PF [10].
- To account for approximated PF with extreme points, a point just beyond the nadir vector (Eq. (1.9)) can also act as $\mathbf{R}_{HV}$ [77]. If two or more PFs are compared, a point little beyond the maximum of the nadir vectors is chosen as the final $\mathbf{R}_{HV}$.
- The location of $\mathbf{R}_{HV}$ can also be pre-fixed [200] (e.g., at $[1.1, \overset{M}{\cdots}, 1.1]$) and all the approximated PFs can be normalized or scaled within a specified range (e.g., $[0, 1]$).

Due to high computational complexity of exact HV calculation, the hypervolume is often approximated. A set of points ($\mathcal{H}_{HV}$) is randomly sampled in this hyper-rectangle using Monte-Carlo simulation. Hypervolume (HV) of the hyper-rectangle is approximated by the fraction of the points in $\mathcal{H}_{HV}$ dominated by the estimated PF $\mathcal{A}_\mathbf{F}$ (Fig. 1.2c) as follows:

$$
\begin{aligned}
HV\left(\mathcal{A}_\mathbf{F}, \mathcal{H}_{HV}\right) &= \frac{1}{|\mathcal{H}_{HV}|} \sum_{j=1}^{|\mathcal{H}_{HV}|} \alpha_{HV}\left(\mathbf{H}_j, \mathcal{A}_\mathbf{F}\right), \text{ where } \mathbf{H}_j \in \mathcal{H}_{HV} \text{ and} \\
\alpha_{HV}\left(\mathbf{H}_j, \mathcal{A}_\mathbf{F}\right) &= \begin{cases} 1, & \text{if } \exists \mathbf{F}\left(\mathbf{X}\right) \in \mathcal{A}_\mathbf{F} \text{ with } \mathbf{F}\left(\mathbf{X}\right) \prec \mathbf{H}_j \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
\tag{1.12}
$$

For its evaluation, attainment function ($\alpha_{HV}(.)$) is defined which returns 1 when a point $\mathbf{H}_j \in \mathcal{H}_{HV}$ is dominated by any solution ($\mathbf{F}\left(\mathbf{X}\right) \in \mathcal{A}_\mathbf{F}$). Hypervolume indicator is given by the average of the values returned by the attainment function over the set of points belonging to $\mathcal{H}_{HV}$. Given the same $\mathcal{H}_{HV}$ with the same $N_{HV} = |\mathcal{H}_{HV}|$, among two approximated PFs, the one having the larger value of HV has better convergence, or better diversity or both with respect to the true PF.

Hypervolume indicator does not suffer from the drawbacks of the previous two indicators [134], i.e., HV being a ratio is bounded in the range $[0,1]$ and the evaluation of Eq. (1.12) does not require information on true PF. However, two major concerns for evaluating hypervolume are the huge computational complexity ($\mathcal{O}\left(M \cdot |\mathcal{A}_{\mathbf{F}}| \cdot |\mathcal{H}_{HV}|\right)$) and its high sensitivity towards the location of the reference point ($\mathbf{R}_{HV}$) for defining the hyper-rectangle and hence, the reference set $\mathcal{H}_{HV}$ [134]. Literature consists of methods to reduce these disadvantages [9, 199]. Nonetheless, the advantages of HV outweigh its drawbacks and has been a popular choice of performance metric in this domain.

Purity Metric

Purity metric is used for comparison of two or more approximations of the PF [10, 11]. Hence, this metric could be used to compare the results of two or more algorithms by unifying their approximated PFs and evaluating the proportion of non-dominated solutions contributed by each of the solution set towards a unified set ($\mathcal{A}^{\star}_{\mathbf{F}}$). Thus, for comparison of K_{PF} solution sets ($\mathcal{A}_{\mathbf{F},1}, \mathcal{A}_{\mathbf{F},2}, \cdots, \mathcal{A}_{\mathbf{F},K_{PF}}$), the unified approximation of the PF ($\mathcal{A}^{\star}_{\mathbf{F}}$) is given by the non-dominated set of the union of these K_{PF} solution sets as follows:

$$\mathcal{A}^{\star}_{\mathbf{F}} = ndset\left(\cup_{i=1}^{K_{PF}} \mathcal{A}_{\mathbf{F},i}\right), \text{ where } ndset(.) \text{ is given by Eq. (1.4).} \tag{1.13}$$

For the i^{th} approximation of the PF, the intersection of $\mathcal{A}_{\mathbf{F},i}$ and the unified set $\mathcal{A}^{\star}_{\mathbf{F}}$ is used to evaluate the purity metric (PM) is evaluated as follows:

$$PM\left(\mathcal{A}_{\mathbf{F},i}, \mathcal{A}^{\star}_{\mathbf{F}}\right) = \frac{|\mathcal{A}_{\mathbf{F},i} \cap \mathcal{A}^{\star}_{\mathbf{F}}|}{|\mathcal{A}_{\mathbf{F},i}|}, \text{ for } i = 1, 2, \cdots, K_{PF}. \tag{1.14}$$

The purity metric is bounded and can be equal to 1 for all the solution sets, as $\sum_{i=1}^{K_{PF}} PM\left(\mathcal{A}_{\mathbf{F},i}, \mathcal{A}^{\star}_{\mathbf{F}}\right) \neq 1$. Among two approximated PFs, the one having a larger purity value is a better approximation of the true PF. However, as Eq. (1.13) estimates $ndset(.)$, the potency of purity metric decreases when $M > 10$ due to dominance resistance [10, 69].

Pareto-Set Proximity

Pareto-Set Proximity (PSP) [188] evaluates the similarity between the approximated PS ($\mathcal{A}_{\mathbf{X}}$) and the true PS (whose sampled version is $\mathcal{H}_{IGD}$) as follows:

$$PSP\left(\mathcal{A}_{\mathbf{X}}, \mathcal{H}_{IGD}\right) = \frac{CoRa}{IGD\left(\mathcal{A}_{\mathbf{X}}, \mathcal{H}_{IGD}\right)}, \text{ where } CoRa = \left(\prod_{i=1}^{N} \gamma_i\right)^{\frac{1}{2N}} \text{ and}$$

$$\gamma_i = \begin{cases} 1, & \text{if } x_i^{MAX} = x_i^{MIN} \\ 0, & \text{if } x_i^{min} \geq x_i^{MAX} \vee x_i^{max} \leq x_i^{MIN} \\ \left(\frac{min\left(x_i^{max}, x_i^{MAX}\right) - max\left(x_i^{min}, x_i^{MIN}\right)}{x_i^{MAX} - x_i^{MIN}}\right)^2, & \text{otherwise.} \end{cases} \tag{1.15}$$

In Eq. (1.15), $CoRa$ represents the cover rate (overlap ratio of the approximated PS to the true PS) and $IGD\left(\mathcal{A}_{\mathbf{X}}, \mathcal{H}_{IGD}\right)$ represents IGDX. Also, x_i^{min} and x_i^{max} represent the minimum and maximum along the i^{th} decision variable over the approximated PS, respectively. Similarly, x_i^{MIN} and x_i^{MAX} represent the minimum and maximum along the i^{th} decision variable over the true PS, respectively. While $CoRa$ indicates overlap, $IGD\left(\mathcal{A}_{\mathbf{X}}, \mathcal{H}_{IGD}\right)$ represents convergence and diversity of the approximated PS with respect to true PS. Thus, PSP quantifies an overall quality of the approximated PS.

Indicators associated with Non-Contributing Solutions

In the decision space, IGDX (Eq. (1.11)) involves the term $\min_{\mathbf{X} \in \mathcal{A}_{\mathbf{X}}} \left(D_E\left(\mathbf{X}, \mathbf{H}\right)\right)$ where $\mathbf{H} \in \mathcal{H}_{IGD}$. A solution $\mathbf{X}^{NS} \in \mathcal{A}_{\mathbf{X}}$ is called a non-contributing solution, if for a given representation of the true PS ($\mathcal{H}_{IGD}$) and the approximated PS ($\mathcal{A}_{\mathbf{X}}$), the following condition [173, 174] is satisfied:

$$\nexists \mathbf{H} \in \mathcal{H}_{IGD} : D_E\left(\mathbf{X}^{NS}, \mathbf{H}_j\right) = \min_{\mathbf{X} \in \mathcal{A}_{\mathbf{X}}} \left(D_E\left(\mathbf{X}, \mathbf{H}\right)\right). \tag{1.16}$$

The notion of a non-contributing solution is shown in Fig. 1.2b for solutions in the objective space. Let the subset of the non-dominated solution set consisting of all such non-contributing solutions be $\mathcal{A}_{\mathbf{X}}^{NS}$. The proportion of non-contributing solutions in the non-dominated solution set is given by $NSX = \left|\mathcal{A}_{\mathbf{X}}^{NS}\right| / \left|\mathcal{A}_{\mathbf{X}}\right|$. This proportion NSX reflects the amount of outliers, i.e., how many non-dominated solutions of the final population are not the nearest neighbors of any point in $\mathcal{H}_{IGD}$ (the set representing the true PS).

Removing $\mathcal{A}_{\mathbf{X}}^{NS}$ from $\mathcal{A}_{\mathbf{X}}$ does not change IGDX, i.e., $IGD(\mathcal{A}_{\mathbf{X}}, \mathcal{H}_{IGD}) = IGD(\mathcal{A}_{\mathbf{X}} - \mathcal{A}_{\mathbf{X}}^{NS}, \mathcal{H}_{IGD})$. However, to note how far the outliers are from the surface of true PS, the convergence metric of $\mathcal{A}_{\mathbf{X}}^{NS}$ can be obtained in the decision space with respect to $\mathcal{H}_{CM} = \mathcal{H}_{IGD}$, i.e., $CM\_NSX = CM\left(\mathcal{A}_{\mathbf{X}}^{NS}, \mathcal{H}_{CM}\right)$. If both NSX and CM_NSX are large, it implies that a large number of solutions are far away from the true PS.

1.3.4 Visualization Methods

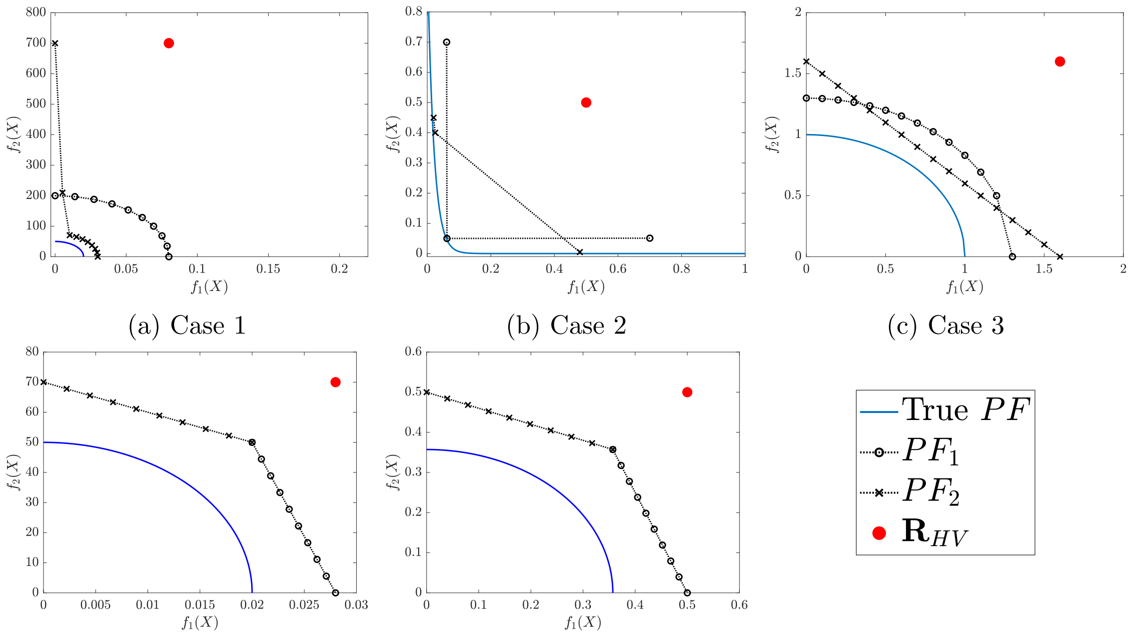
The need to visualize the approximated PF is emphasized in [134] using several case-studies, which show that there can be conflicting assessments based on the different performance indicators [93, 134]. For example, the quality of an approximated PF does not always agree in terms of visualization, convergence metric and hypervolume. Such anomalous results [134] are demonstrated through the following four test scenarios (synthesized for 2-objective minimization problems):

- *Sensitivity to Solutions Far Away from True Pareto-Front (Case 1)*: Among the two Pareto-Front approximations (PF_1 and PF_2), let PF_2 be much closer to the true Pareto-Front (True PF) than PF_1 but there is one very distant point in PF_2. The convergence metric, being sensitive to the distance of all the points, indicates PF_1 to be better. This result is contradicted by the hypervolume indicator, which indicates PF_2 to be a better approximation of the True PF. This case is illustrated in Fig. 1.3a and the performance is mentioned in Table 1.1.

- *Variation in Distribution of Points Near the True Pareto-Front (Case 2)*: Similar to Case-1, let both the Pareto-Fronts (PF_1 and PF_2) be intertwined. However, in contrast to Case-1, let PF_2 have all its points very close to True PF. Based on the convergence metric, PF_2 is a better approximation, whereas based on the hypervolume indicator PF_1 is better. This discrepancy is because the reference point builds the hyper-rectangle in such a manner that the only solution of PF_1 which lies within hyper-rectangle is present at a location near the vertex (origin for this case) which is diagonally opposite to the reference point. Thus, for conflict resolution, this case requires a third metric (such as the number of points of the approximated PF enclosed by the hyper-rectangle). The respective scenario is illustrated in Fig. 1.3b and the other details are mentioned in Table 1.1.

- *Shape of the Pareto-Front (Case 3)*: Among the two Pareto-Fronts (PF_1 and PF_2), PF_1 better approximates the shape of the True PF than PF_2. However, the performance values (convergence metric and hypervolume indicator) are only slightly different between the two PFs concluding that the PFs are nearly equivalent. However, none of the metrics has captured the information on the shape of the PF. The respective scenario is illustrated in Fig. 1.3c and the other details are mentioned in Table 1.1.

- *Normalizing the Pareto-Fronts (Case 4)*: Among the two Pareto-Fronts (PF_1 and PF_2), PF_1 is much closer to the True PF as indicated by the convergence metric (Case 4(a)). Both the PFs have the same hypervolume, indicating that these are of the same quality. The discrepancy is because the scale of the second objective (f_2) is larger than the scale of the first objective (f_1). This case is shown in Fig. 1.3d. After scaling (Fig. 1.3e), this discrepancy disappears (Case 4(b)). Both the objectives of True PF, PF_1 and PF_2 are linearly mapped in the interval $[0, 0.5]$. The performance of these cases is given in Table 1.1.



(a) Case 1 (b) Case 2 (c) Case 3

(d) Case 4(a) (e) Case 4(b) (f) Legend

Figure 1.3: Different scenarios [134] that convergence metric and hypervolume indicator fail to resolve: (a) Case 1 for studying the sensitivity to outliers, (b) Case 2 for studying the effect of the number of points in the PF, (c) Case 3 for studying the capability to preserve the shape of PF, (d-e) Case 4(a-b) for studying the effects of normalizing the PF on convergence metric and (f) graph legend for different cases.

Table 1.1: Conflicting values of convergence metric and hypervolume indicator for various test-cases [134] evaluated with the specified parameters (location of reference point $\mathbf{R}_{HV}$, size of reference sets for convergence metric $|\mathcal{H}_{CM}|$ and hypervolume indicator $|\mathcal{H}_{HV}|$).

Case	**Convergence Metric**	**Hypervolume Indicator**	**Parameters**
Case 1	PF_1: 80.8449 PF_2: 85.3234	PF_1: 0.7674 PF_2: 0.8977	$\mathbf{R}_{HV} = max(PF_1^{nad}, PF_2^{nad})$, $\|\mathcal{H}_{CM}\| = 158$, $\|\mathcal{H}_{HV}\| = 20000$
Case 2	PF_1: 0.0357 PF_2: 0.0062	PF_1: 0.7860 PF_2: 0.2222	$\mathbf{R}_{HV} = (0.5, 0.5)$, $\|\mathcal{H}_{CM}\| = 1001$, $\|\mathcal{H}_{HV}\| = 10000$
Case 3	PF_1: 0.3000 PF_2: 0.3178	PF_1: 0.4710 PF_2: 0.4740	$\mathbf{R}_{HV} = max(PF_1^{nad}, PF_2^{nad})$, $\|\mathcal{H}_{CM}\| = 101$, $\|\mathcal{H}_{HV}\| = 1000$
Case 4	PF_1: 0.0521 (a) PF_2: 10.0020 (a) PF_1: 0.1230 (b) PF_2: 0.1229 (b)	PF_1: 0.1692 (a) PF_2: 0.1686 (a) PF_1: 0.1717 (b) PF_2: 0.1714 (b)	For both cases (a) and (b) $\mathbf{R}_{HV} = max(PF_1^{nad}, PF_2^{nad})$, $\|\mathcal{H}_{CM}\| = 158$, $\|\mathcal{H}_{HV}\| = 20000$

Thus, for ease of assessment and proper representation, it is essential not to rely only on the performance measures but also to visualize the high-dimensional objective space through methods like Buddle chart [67], parallel coordinate plots [22,67], heatmaps [22,67], polar coordinate plots [22, 67, 68] and Self-Organizing Maps [67, 166].

1.3.5 Special Types of Optimization Problems

There are several other types of optimization problems, which are more challenging than the fundamental box-constrained multi-objective optimization problems. The real-world problems can present multiple such challenges at the same time. Hence, it is essential to analyze which algorithmic strategies are beneficial for which problem attributes. However, for studying a novel algorithmic framework for each of these particular types of problems, specific benchmark test functions and performance indicators are barely available, which currently limits extensive studies in these areas.

Constrained Multi-objective Optimization Problems

A multi-objective optimization problem is often subjected to equality and inequality constraints (Eq. (1.1)), which determine the feasible regions of the search space for an optimization problem [33, 96]. The simplest strategy [47] to deal with such constrained optimization problems involve modifying the selection stage of an EA where there are three cases: (i) when the comparison is between two feasible solutions, candidate selection is performed in a way similar to the approaches for box-constrained optimization problems, (ii) when the comparison is between a feasible solution and an infeasible solution, candidate selection favors the feasible solution, and (iii) when the comparison is between two infeasible solutions, candidate selection favors the less constraint-violating solution.

Problems with Expensive Fitness Evaluations

From a practical standpoint, often the computation of fitness of a single feasible candidate can be expensive [32] such as when the computation takes minutes or hours, involves a financial cost (such as assembling cost, reagent cost, and other expenses), or involves simulation (black-box scenario, i.e., the algebraic form of a fitness function is unknown) [94]. For addressing such problems, one of the popular approaches is to introduce approximations, especially function approximations. Computational models for functional approximations are often known as surrogates and EAs using objective values estimated by surrogates are often referred to as Surrogate-Assisted EAs (SAEAs) [94]. A surrogate (synonymous to a metamodel) helps to replace (fully or partly) the computationally expensive objective functions.

Five crucial challenges [29,144] of using surrogates in MOEAs are: (i) choosing the surrogate model (e.g., Kriging, neural networks, polynomial approximation), (ii) determining which quantity to approximate, (iii) maintaining a substantially smaller training cost of surrogate than evaluation cost of expensive objectives, (iv) deciding how to update the surrogate model (selecting representative solutions for training), and (v) deciding when to update the surrogate model (based on the surrogate's accuracy).

Large-scale Multi-objective Optimization Problems

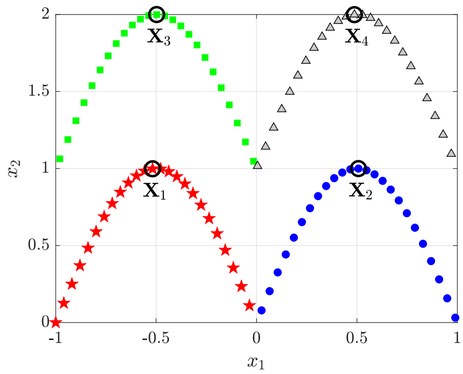
Large-scale optimization problems have more than 100 decision variables (Eq. (1.1) with $N \geq 100$) [27,95]. Optimizing a whole large-scale problem at once is difficult. For tackling such problems, *co-operative co-evolution* is mostly used. It involves the parallel and collaborative evolution (symbiosis) of sub-problems defined by variable groupings (species). Thus, the overhead of dealing with such large-scale multi-objective optimization problems are optimal separation of decision variables (or defining the sub-problems) and efficient collaboration (information exchange) between multiple sub-problems [5].

Multi-Modal Multi-objective Optimization Problems

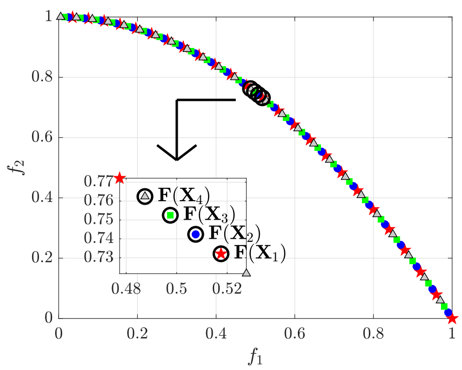
The notion of Multi-Modal Multi-Objective Problem (MMMOP) [171] arises when a set of k_{PS} (≥ 2) distinct decision vectors ($\mathcal{A}_M = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_{k_{PS}}\}$) maps to *almost same* objective vectors as follows:

$$\forall\left(\mathbf{X}_i, \mathbf{X}_j\right) \in \mathcal{A}_M \times \mathcal{A}_M \text{ and } i \neq j : \left\|\mathbf{F}\left(\mathbf{X}_i\right) - \mathbf{F}\left(\mathbf{X}_j\right)\right\| < \epsilon_{PF}, \quad \text{where } \epsilon_{PF} \text{ is a small number.} \tag{1.17}$$

An example is illustrated in Fig. 1.4 for a benchmark multi-modal multi-objective test problem (MMF4 [112]). Thus, the PS can consist of multiple subsets of non-dominated solutions, where each subset can independently generate the entire PF.



(a) PS of MMF4



(b) PF of MMF4

Figure 1.4: Four solution vectors ($\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ and $\mathbf{X}_4$) mapping to *almost same* objective vectors ($\mathbf{F}(\mathbf{X}_1), \mathbf{F}(\mathbf{X}_2), \mathbf{F}(\mathbf{X}_3)$ and $\mathbf{F}(\mathbf{X}_4)$) for a benchmark test problem (MMF4) [140].

The motivation to study MMMOPs arises due to those decision maker's preferences which cannot be mathematically formulated and introduced in the MMMOPs. Thus, providing a diverse set of nearly equivalent solutions help the users to make an informed decision. Another advantage of studying MMMOPs is if the practical implementation of a solution is hindered, an equivalent alternative is readily available. Some practical MMMOPs are seen in rocket engine design [103], feature selection problem [189] and path-planning problem [90]. To optimize such MMMOPs, an EA faces the following challenges:

1. Maintaining diversity in the decision space, i.e., representing and maintaining diversity within each of the multiple solution sets which independently maps to a diverse approximation of the PF.
2. Necessity of large population to efficiently represent MMMOPs. For example, if k_{PF}

points (e.g., 100) represent a 2-objective PF and k_{PS} decision vectors (e.g., 4 for MMF4 problem [112, 137]) map to each point of the PF, then the final population size required is $k_{PF} \times k_{PS}$ (e.g., $100 \times 4 = 400$).

Problems having Dynamic Fitness Landscape

This class of problems have their fitness landscape as a variable of time (like in a noisy environment) [98,154]. Thus, algorithms designed for problems with such dynamic landscapes are characterized by fast adaptation, sensitive to these changes. Along with specialized benchmark functions [58], specific performance measures [70] are also required for assessing the efficacy of an algorithm for problems with dynamic landscapes.

Interval-valued Multi-objective Optimization Problems

Often in real-world MOPs [114, 193], there are some uncertainties or tolerances involved with the values of the objectives. When such uncertainties in objectives can be defined to be uniform within a real-valued interval $\left[f_i^L\left(\mathbf{X}\right), f_i^U\left(\mathbf{X}\right)\right]$, the MOP becomes an Interval-Valued MOP (IVMOP). Mathematical formulation of such an IVMOP is similar to Eq. (1.1) except that the i^{th} objective function is further characterized as follows:

$$f_i\left(\mathbf{X}\right) = \{p_{iv} | f_i^L\left(\mathbf{X}\right) \leq p_{iv} \leq f_i^U\left(\mathbf{X}\right), p_{iv} \in \mathbb{R}\}. \tag{1.18}$$

For extending MOEAs to address IVMOPs [193], strategies of conventional individual comparison, population diversity and population evolution need to be modified to adapt to interval-valued environments.

1.3.6 Multi-Criteria Decision-Making

After termination of MOEAs, the approximated PS and PF are obtained, both of which are sets of solutions. Such a set of possible trade-offs are essential to make an informed decision. However, an application problem can implement only one solution. This selection of a Pareto-optimal solution from the approximations of PS and PF is dictated by multiple criteria (preferences of the decision-makers) and the methods dealing with this selection composes the domain of Multi-Criteria Decision-Making (MCDM) [152]. In this regard, the most prominent work deals with the selection of knee-point [192], which considers nearly equal compromise in all objectives. However, a knee point has several disadvantages for irregular PF and more recent methods exists to deal with such challenges [65, 186]. While MCDM for box-constrained MOPs is a widely-studied domain, it is relatively scarce for those particular types of optimization problems which are discussed in Section 1.3.5.

1.3.7 Theoretical Studies on Population Dynamics

While the domain of metaheuristics evolves by the introduction of novel search strategies, there are limited theoretical studies to prove the scope (such as convergence, stability, bounds of hyper-parameters) of these strategies for MOEAs. In the absence of such theoretical evidence, it is not apparent why and when a particular strategy works. Thus, most of the studies rely on empirical results. Even the hyper-parameter values are set based on sensitivity studies. Similarly, EAs are accompanied by scalability studies. However, experiments on the convergence behavior or the population dynamics of the MOEAs are often overlooked, which are essential to investigate the effect of various strategies. Motivated by this research gap, some population dynamics indicators have been recently developed in [161] to aid in the analysis of MOEAs and MaOEAs.

Overall, the various research areas for studying MaOEAs are summarized in Fig. 1.5, where the well-explored to scarcely-explored areas are color-coded.

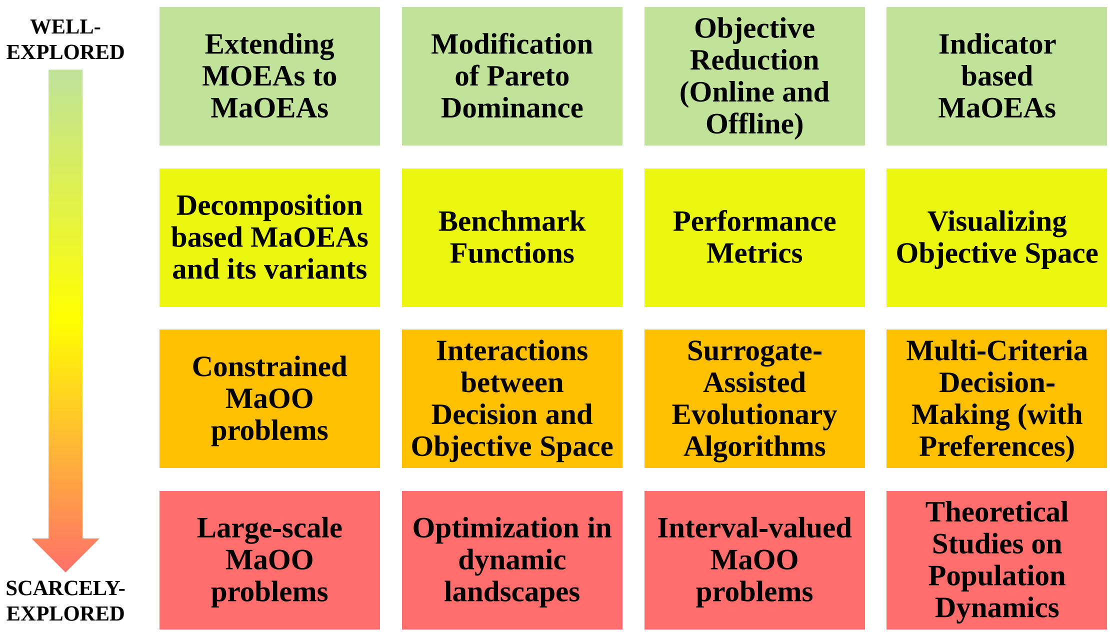


Figure 1.5: Research areas in the domain of Many-Objective Optimization (MaOO).

As mentioned in Section 1.1, there are several real-world applications where MaOO

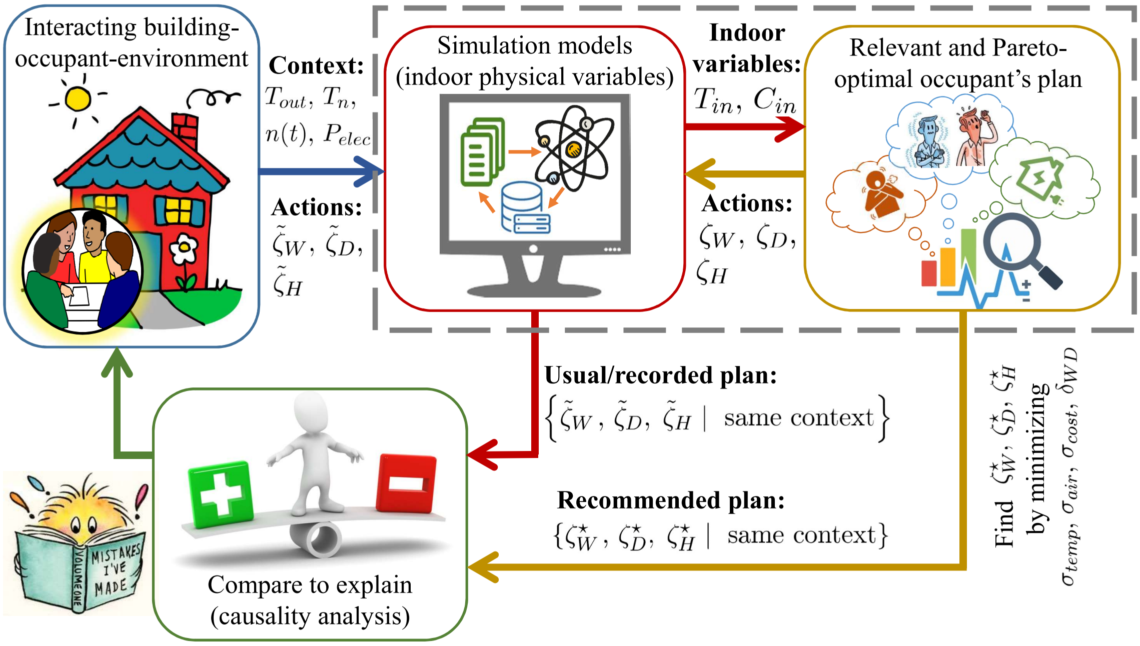


Figure 1.6: Building energy management framework [139] where the optimization module (dashed box) aims at estimating the relevant and Pareto-optimal schedule of occupants' actions (opening/closing of doors $\zeta^\star_D$, windows $\zeta^\star_W$ and turning on/off heater $\zeta^\star_H$) such that indoor thermal discomfort (σ_{temp}), aeraulic discomfort (σ_{air}), heater-related energy-cost (σ_{cost}) and number of changes in recommendations (δ_{WD}) are minimized.

plays a central role. The use of MaOEAs in real-world problems help to identify various open research areas and thus, provides the necessary research motivation. Inspired by this practice, the efficacy of some strategies presented in different chapters of this thesis are also established on a real-world problem of building energy management.

1.4 Many-Objective Building Energy Management Problem

Building energy management has been a trending topic over the past decade as nearly 40% of the global energy consumption is from the buildings sector [72]. One of the prevailing strategies for building energy management, even applicable to existing non-green buildings, is regulating occupants' actions to attain the finest indoor ambience [2]. This optimal schedule of occupants' actions helps to generate cause-and-effect explanations such that the occupants can learn to modify their actions towards an energy-efficient routine [133].

The overall building energy management framework [133] is outlined in Fig. 1.6. The loop begins with sensor-fitted rooms and creation of a database ($\mathcal{H}_{DB}$) to store usual occupants' actions (opening/closing of windows $\tilde{\zeta}_W$ and doors $\tilde{\zeta}_D$ and turning on/off the

Table 1.2: Mathematical formulation of the four optimization objectives.

Objective	Parameters
$\sigma^k_{temp} = \begin{cases} \frac{294.15-T^k_{in}}{294.15-291.15}, & \text{if } T^k_{in} < 294.15 \text{ and } n^k > 0 \\ 0, & \text{if } 294.15 \leq T^k_{in} \leq 296.15 \text{ or } n^k = 0 \\ \frac{T^k_{in}-296.15}{299.15-296.15}, & \text{if } T^k_{in} > 296.15 \text{ and } n^k > 0 \end{cases}$	simulated T^k_{in} (in K) and occupancy (n^k) at the k^{th} hour
$\sigma^k_{air}\left(C^k_{in}\right) = \begin{cases} 0, & \text{if } C^k_{in} \leq 400 \text{ or } n^k = 0 \\ \frac{C^k_{in}-400}{1500-400}, & \text{if } C^k_{in} > 400 \text{ and } n^k > 0 \end{cases}$	simulated C^k_{in} (in ppm) and occupancy (n^k) at the k^{th} hour
$\sigma^k_{cost}\left(P^k_{elec}, P^k_{fuel}\right) = \begin{cases} \frac{P^k_{elec}E_{elec}+P^k_{fuel}E_{fuel}}{1000}, & \text{if } n^k > 0 \\ 0, & \text{if } n^k = 0 \end{cases}$	simulated P^k_{fuel} (in W); and recorded P^k_{elec} (in W), E_{elec} and E_{fuel} (in Euros per kWh) at the k^{th} hour in $\mathcal{H}_{DB}$
$\delta^k_{WD}\left(\zeta^k_{pair}\right) = \begin{cases} \delta^{k-1}_{WD}\left(\zeta^{k-1}_{pair}\right)+1, & \text{if } \zeta^k_{pair} \neq \zeta^{k-1}_{pair} \\ \delta^{k-1}_{WD}\left(\zeta^{k-1}_{pair}\right)+0, & \text{if } \zeta^k_{pair} = \zeta^{k-1}_{pair} \end{cases}$ where $\zeta^k_{pair} = \left(\zeta^k_W, \zeta^k_D\right)$ and $\delta^0_{WD}\left(\zeta^0_{pair}\right) = 0$	schedule of occupants' actions $\mathbf{X}_B$

room heater $\tilde{\zeta}_H$) and contextual variables (outdoor temperature T_{out}, corridor temperature T_n, occupancy n, plug load energy consumption P_{elec}, fuel cost E_{fuel} and electricity cost E_{elec}). For the k^{th} hour, this data is used by the simulation models [133] such that the indoor physical variables (temperature T^k_{in}, CO_2 concentration C^k_{in} and heater energy consumption P^k_{fuel}) can be evaluated for hypothetical actions (ζ^k_W, ζ^k_D and ζ^k_H). Hence, a 24-hour action schedule is denoted by a 72-dimensional solution vector ($\mathbf{X}_B$) as follows:

$$\mathbf{X}_B = [x_{B,1}, \cdots, x_{B,72}] = \left[\zeta^0_W, \cdots, \zeta^{23}_W, \zeta^0_D, \cdots, \zeta^{23}_D, \zeta^0_H, \cdots, \zeta^{23}_H\right]. \tag{1.19}$$

These variables (input and simulated) are used to minimize thermal discomfort σ_{temp}, aeraulic discomfort σ_{air}, heater associated cost indicator σ_{cost} and the number of changes in a schedule δ_{WD}. Thus, using the formulation from Table 1.2, the optimization objective vector $\mathbf{F}_B$ is given as follows:

$$\mathbf{F}_B = [f_{B,1}, f_{B,2}, f_{B,3}, f_{B,4}] = \frac{1}{24}\left[\sum_{k=0}^{23}\sigma^k_{temp}, \sum_{k=0}^{23}\sigma^k_{air}, \sum_{k=0}^{23}\sigma^k_{cost}, \sum_{k=1}^{23}\delta^k_{WD}\right]. \tag{1.20}$$

Thus, the minimization of $\mathbf{F}_B(\mathbf{X}_B)$ estimates the optimal hourly actions $\zeta^\star_W$, $\zeta^\star_D$ and $\zeta^\star_H$ under the same recorded context used by the simulation models (fetched from the database $\mathcal{H}_{DB}$). These recommended actions can be compared with the usual actions for generating causal explanations, from which the occupants can learn by themselves the impact of their actions [2, 133]. In this regard, several case-studies [1, 2, 132, 133, 135, 143] are conducted for an office at Grenoble Institute of Technology, France, leading to the prototyping of an human-computer interface for implementing the framework of Fig. 1.6.

1.5 Organization of the Thesis

This thesis is a comprehensive attempt to present EAs with improved performance for tackling a wide range of MaOPs having different characteristics. The thesis also aims at demonstrating the use of MaOEAs for tackling a real-world problem where decision-making (post-optimization) also plays a significant role in presenting the final problem solution. The current chapter deals with the basics of multi- and many-objective optimization algorithms and introduces the optimization problem for building energy management. The next five chapters constitute the contributory part of the entire thesis, followed by a concluding chapter. The content of the chapters are outlined below.

- **Chapter 2:** The unsuitability of MOEAs for solving MaOPs, due to reduced selection pressure with an increased number of objectives, is often tackled using objective reduction approaches [141]. Chapter 2 of this thesis presents Differential Evolution using Clustering based Objective Reduction (DECOR) [142]. Correlation distance based clustering of objectives from the approximated PF, followed by elimination of all but the centroid constituent of the most compact cluster (with special care to singleton cluster), yields the reduced objective set. The objective set is periodically switched between full and reduced size to ensure both global and local exploration. For finer clustering, the number of clusters is eventually increased until it is equal to the remaining number of objectives. DECOR is applied on 10- and 20-objective DTLZ problems which demonstrate its superior performance in terms of convergence and equivalence in terms of diversity as compared to state-of-the-art MOEAs.

- **Chapter 3:** Enhanced diversity of solutions in the objective space can be attained using reference vector based decomposition algorithms. For achieving better solution diversity, this chapter presents Ensemble of Single Objective Evolutionary Algorithms (ESOEA) [138]. It adopts the reference-direction based approach to decompose the population, followed by scalarization to transform the MaOP into several single objective sub-problems which further enhances the selection pressure. Additionally, with a feedback strategy, ESOEA explores the directions along difficult regions and thus, improves the search capabilities along those directions. For experimental validation, ESOEA is executed on several benchmark problems from the DTLZ, WFG, IMB and CEC 2009 competition test suites. For assessing the efficacy of ESOEA, its performance is compared with numerous other MOEAs and adaptive MOEAs such as MOEA/D, MOEA/D-M2M, NSGA-II, AR-MOEA, MOEA/D-DRA ENS-MOEA/D, HypE, DEMO, α-DEMO-revised, DECOR, NSGA-III, A-NSGA-III, RVEA, RVEA* and MOEA/DD.

- **Chapter 4:** While reference direction based decomposition of the objective space is one of the prominent strategies to address MaOPs, literature severely lacks formal mathematical analysis to establish the reason behind the superior performance of such methods. In this chapter, the neighborhood property of the MaOPs is recognized. It is used to present Neighborhood-sensitive Archived Evolutionary Many-objective Optimizer (NAEMO) [160], where mating occurs within a local neighborhood and every reference direction continues to retain at least one associated candidate solution. Such preservation of candidate solutions leads to a monotonic improvement in diversity, as proven using a novel diversity indicator (D_metric [161]). This characteristic of NAEMO is also supported by experimental evidence. Moreover, to keep the archive size under control, periodic filtering modules are integrated with the NAEMO framework. Experimental results reveal that NAEMO outperforms several state-of-the-art algorithms such as NSGA-III, MOEA/D, θ-DEA, MOEA/DD, GrEA, HypE, MOPSO and dMOPSO. It is also competitive to MOEA/D-M2M on IMB test problems.

- **Chapter 5:** Multi-Modal Multi-objective Optimization Problems (MMMOPs) have multiple subsets within the PS, each independently mapping to the same PF [137]. The existing MOEAs are incapable of ensuring that the multiple subsets in PS are represented in the solution set. Moreover, the solution diversity in the PF, obtained by the handful of EAs designed for MMMOPs, are inferior to those obtained by MOEAs. This chapter highlights the disadvantage of using crowding distance in the decision space. It presents two evolutionary frameworks (LORD and LORD-II) which use decomposition in both objective and decision space for dealing with MMMOPs and multi-modal many-objective problems, respectively [140]. Its efficacy is established by comparing its performance on test instances obtained from the CEC 2019 test suite and polygon problems. These EAs (LORD and LORD-II) not only improve the diversity of PS over EAs for MMMOPs but also improve the performance in objective space over MOEAs.

- **Chapter 6:** While several effective and versatile EAs for MaOO problems have been presented in this thesis, often real-world scenarios call for application-specific customizations. In this chapter, the many-objective building energy management problem is considered, which aims to minimize the occupants' discomfort (thermal and aeraulic) [143], heater-related energy expenses [132] and the number of recommended changes [133]. This study demonstrates the real-world applicability of NAEMO (Chapter 4), whose performance is also compared with other state-of-the-art EAs. However, such search procedures for Pareto-optimal occupants' actions overlook the possibility of multiple action schedules for similar comfort trade-offs [139]. To address the decision space attributes, like binary-encoding, multi-view and multi-modal nature of occupants' actions, the algorithm of LORD (Chapter 5) is further customized and its performance is noted for the concerned optimization problem [139].

  Furthermore, this chapter discusses four strategies for the selection of the compromise of interest (from both objective and decision space). The first strategy deals with decision-making in the absence of any preference. The second strategy considers the occupants' interactions with the system to set a realistic preference in the objective space by learning about the Pareto-Front. However, when the subjective preferences of multiple occupants are considered with equal priority, the third approach of obtaining a fair consensus solution is developed [135]. Finally, the fourth strategy discusses how the preference of action schedule (in decision space) can be amalgamated with any of the above-mentioned approaches to generate a context-relevant yet Pareto-optimal schedule of occupants' actions [139].

  For generating awareness among building occupants, causal explanations are obtained from the difference between the recommended Pareto-optimal scenario and the usual scenario of the occupants [2, 133]. Such explanations help the occupants in embracing the recommended energy-efficient routine.

- **Chapter 7:** In the final chapter, the various studies presented in this thesis are briefly summarized. The limitations of these studies are highlighted. Alongside, several open areas are enumerated, which present the scope of further extending the computational strategies discussed in this thesis.

The development of approaches over the various chapters is summarized in Fig. 1.7.

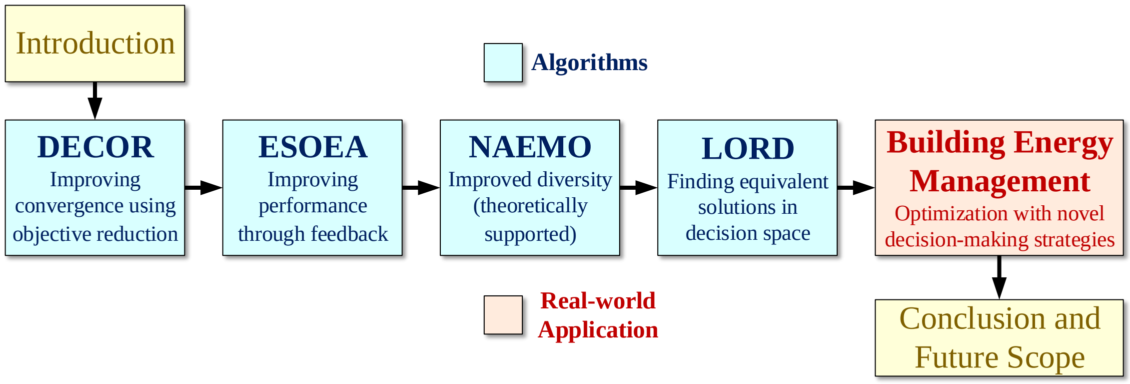


Figure 1.7: Graphical summarization of the thesis.

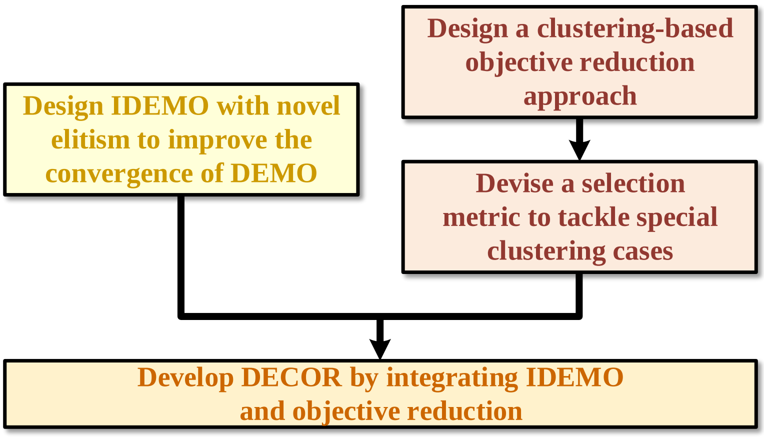
Chapter 2

DECOR: Differential Evolution using Clustering based Objective Reduction for Many-Objective Optimization [142]

Outline

Objective: *To develop an optimization algorithm which uses objective reduction in the background of a multi-objective optimization algorithm for addressing many-objective optimization problems.*

Workflow:



2.1 Introduction

In order to deal with the challenges of applying Multi-Objective Evolutionary Algorithms (MOEAs) to Many-Objective Optimization (MaOO) problems (Eq. (1.2)), objective reduction approaches are often adopted. In these approaches, the most conflicting m objectives out of M objectives ($m \leq M$) are chosen. The size of the full and reduced objective sets are denoted as M and m, respectively. If $m \leq 3$ is achieved, the MaOO problem reduces to a Multi-Objective Optimization (MOO) problem and can be solved using MOEAs. Even otherwise, i.e., achieving $4 \leq m < M$ helps to minimize the computational cost [10,87,157]. Thus, objective reduction is an efficient way to deal with MaOO problems.

This chapter presents an evolutionary algorithm known as Differential Evolution using Clustering based Objective Reduction (DECOR) [142]. Correlation distance based clustering of objectives from the approximated Pareto-Front (PF), followed by elimination of all but the centroid constituent of the most compact cluster (with special care to singleton cluster), yields the reduced objective set. During optimization, the objective set periodically toggles between full and reduced size to ensure both global and local exploration. For finer clustering, the number of clusters is eventually increased until it is equal to the remaining number of objectives. DECOR is integrated with an Improved Differential Evolution for Multi-objective Optimization (IDEMO), which uses a novel elitist selection and ranking strategy to solve MaOO problems. DECOR is applied to some DTLZ problems for 10 and 20 objectives. These experiments demonstrate the superior convergence of DECOR in comparison to several state-of-the-art algorithms.

Outline of the rest of the chapter is as follows. Previous works, related to the objective reduction, are briefly described in Section 2.2 while highlighting the primary contributions of DECOR. The modifications of the base optimization algorithm (DEMO) to yield Improved DEMO (or IDEMO) are presented in Section 2.3 and the developed objective reduction based optimization approach (i.e., DECOR) is described in Section 2.4. The performance and the statistical significance of DECOR are analyzed in Section 2.5, and the major observations are further discussed in Section 2.6. Finally, Section 2.7 concludes the chapter, summarizing the overall observations of the various experiments to perform optimization using objective reduction.

2.2 Motivation for the Work

This section briefly describes the shortcomings of the existing works to explain the motivation behind developing DECOR. Subsequently, the novelties of DECOR are highlighted.

2.2.1 Related Studies on Objective Reduction

The objective reduction algorithms can be divided into two sets: (i) m is specified by the user [17, 86], and (ii) optimal m is automatically determined [17, 86]. Some of the notable objective reduction algorithms are Brockhoff and Zitler's $\underline{\delta}$-Minimum Objective Sub-Set (δ-MOSS) and $\underline{k}$-sized Error Minimizing Objective Sub-Set(k-EMOSS) [17], Deb and Saxena's Principal Component Analysis NSGA-II (PCA-NSGA-II) [48], Coello and Lopez's $\underline{k}$-sized Objective Sub-Set Algorithm (kOSSA) and mixed search scheme of kOSSA [86], Bandyopadhyay and Mukherjee's α-DEMO and α-DEMO-revised [10]. All these methods (except δ-MOSS and k-EMOSS [17]) use correlation among objectives to determine conflicting objectives. The major drawbacks of these approaches are as follows:

- In Brockhoff and Zitler's approach (δ-MOSS and k-EMOSS [17]), a greedy approach is followed where the minimal alteration in induced Pareto-dominance relation is searched by removing one objective in every turn. The high time complexity of these algorithms limits their practical usages [17, 87].

- Online objective reduction [86], which performs objective reduction during the search, can help in speeding up the process. However, removing one objective at a time (like in the mixed search scheme of kOSSA [86]) can still be slow. Hence, the provision for removal of multiple objectives at a time is adopted in recent years, like in α-DEMO and α-DEMO-revised [10] and in the approach of [141].

- For k-EMOSS [17], kOSSA [86] and α-DEMO [10], allowing user to choose m can be disadvantageous [141]. Firstly, the optimal m, which results in the same PF as with M, cannot be pre-determined. Secondly, for desirable performance, the algorithms are repeatedly evaluated by varying m. This approach is not user-friendly.

Considering all these disadvantages, an objective reduction approach has been developed in [142], which is fast (online and has provision for elimination of multiple objectives at a time) and automatically finds m.

2.2.2 Novel Characteristics of DECOR

The algorithm (DECOR) [142] is an extension of the work reported in [141]. Experimental results demonstrate the superior convergence performance of DECOR [142], while in terms of diversity, its performance is equivalent to that of other MaOO algorithms.

On the one hand, DECOR [142] has the following two key features which are similar to some existing algorithms:

- It uses Differential Evolution as the underlying optimizer similar to the approaches in [10, 141].
- It uses a similar principle of correlation-based clustering for online objective reduction as done in [86, 141]. Hence, the algorithm is called Differential Evolution using Clustering based Objective Reduction (DECOR) [142].

On the other hand, DECOR [142] differs from the approach in [141], from which it has been extended, in the following aspects:

- DECOR presents both the versions of objective reduction (automatic determination of m and user-specified m).
- DECOR avoids premature termination of objective reduction due to appearance of a singleton cluster, by determining whether it is relatively close to the nearest cluster.
- DECOR uses an Improved DEMO (IDEMO) with a novel elitist selection strategy to avoid early saturation of the population by non-dominated solutions.
- DECOR uses IDEMO with a novel ranking scheme which combines the distance of a solution from the ideal point with the crowding distance to account for both convergence and diversity during online objective reduction.

Thus, the contributions are two-fold. Firstly, a new MaOO technique (IDEMO) is introduced with an enhanced ranking strategy which utilizes a regulated elitist approach and a new selection operator based on the crowding distance and the distance from the ideal point. Secondly, a new objective reduction technique is presented, which is further integrated with IDEMO to yield the MaOO algorithm called DECOR. Such an optimization algorithm, having all the above-mentioned features, has not been developed before DECOR. Thus, these features highlight the novelty of DECOR [142].

2.3 Underlying Optimization Algorithm - IDEMO

A new ranking strategy and a new elitist operation are introduced in the classic Differential Evolution based Multi-objective Optimization (DEMO) to develop the Improved DEMO (or IDEMO) such that its selection process is more pertinent for solving MaOO problems. This section briefly describes the different steps of IDEMO.

2.3.1 Differential Evolution for Multi-objective Optimization (DEMO)

DEMO [153] is an extension of the single-objective version of Differential Evolution (DE) [164, 168]. DEMO has four steps: Initialization, Mutation, Recombination and Selection, which are described in the following paragraphs.

Initialization

The initial population ($\mathcal{A}^{parent}_{G=0}$) is a randomly initialized matrix of order $n_{pop} \times N$, where there are n_{pop} number of N-dimensional decision vectors. Using the lower and upper bounds of the j^{th} decision variable (x^L_j and x^U_j) which define the search space ($\mathcal{D}$), the i^{th} decision vector ($\mathbf{X}_{i,0} \in \mathcal{A}^{parent}_{G=0}$) is initialized as follows:

$$\begin{gathered} \mathbf{X}_{i,0} = \left[x_{i1,0}, x_{i2,0}, \cdots, x_{iN,0}\right], \text{ where } x_{ij,0} = x^L_j + rand(0,1) \times \left(x^U_j - x^L_j\right) \\ \text{for } i = 1, \cdots, n_{pop} \text{ and } j = 1, \cdots, N. \end{gathered} \tag{2.1}$$

Random Mutation with One Difference Vector

For the i-th candidate ($\mathbf{X}_{i,G}$) at generation G, three distinct indices r_1, r_2 and r_3 of decision vectors are randomly chosen and the mutant vector $\mathbf{X}^{mut}_{i,G}$ is obtained as follows:

$$\begin{gathered} \mathbf{X}^{mut}_{i,G} = \mathbf{X}_{r_1,G} + F^{DE} \times \left(\mathbf{X}_{r_2,G} - \mathbf{X}_{r_3,G}\right), \text{ for } i = 1, \cdots, n_{pop} \\ \text{where } F^{DE} \in [0,2] \text{ is the scale factor.} \end{gathered} \tag{2.2}$$

Binomial Crossover

The trial vector for the next generation ($\mathbf{X}^{trial}_{i,G+1}$) is formed by combining variables from $\mathbf{X}^{mut}_{i,G}$ with a probability higher than the crossover rate (CR) and from $\mathbf{X}_{i,G}$ with a probability lower than CR. Also, for forming $\mathbf{X}^{trial}_{i,G+1}$ a decision variable corresponding to a random index (I_{rand}) is always chosen from $\mathbf{X}^{mut}_{i,G}$ so that $\mathbf{X}^{trial}_{i,G+1}$ is different from $\mathbf{X}_{i,G}$.

Thus, the binomial crossover [42] for the j^{th} decision variable is given as follows:

$$x_{ij,G+1}^{trial} = \begin{cases} x_{ij,G}^{mut}, & \text{if } rand(0,1) \leq CR \text{ or } j = I_{rand} \\ x_{ij,G}, & \text{if } rand(0,1) > CR \text{ and } j \neq I_{rand} \end{cases} \tag{2.3}$$

$$\text{where } i = 1, \cdots, n_{pop} \text{ and } j = 1, \cdots, N.$$

Selection

The population for the next generation is obtained as $\mathcal{A}_{G+1}^{parent} = \cup_{i=1}^{n_{pop}} \mathbf{X}_{i,G+1}$ where $\mathbf{X}_{i,G+1}$ is chosen from $\mathbf{X}_{i,G}$ and $\mathbf{X}_{i,G+1}^{trial}$ using Pareto-dominance relation (Eq. (1.3)) as follows:

$$\mathbf{X}_{i,G+1} = \begin{cases} \mathbf{X}_{i,G}, & \text{if } \mathbf{X}_{i,G} \prec \mathbf{X}_{i,G+1}^{trial} \\ \mathbf{X}_{i,G+1}^{trial}, & \text{otherwise} \end{cases} \quad \text{for } i = 1, \cdots, n_{pop}. \tag{2.4}$$

2.3.2 Improved Elitist Strategy at the Selection Stage

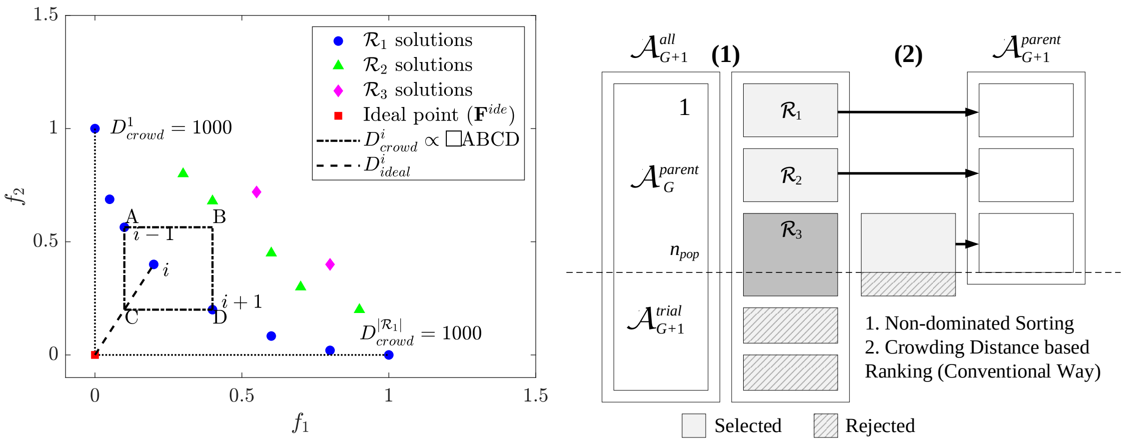
In the elitist framework, at the selection stage, the trial solutions ($\mathbf{X}_{i,G+1}^{trial}$) form the set of new candidates ($\mathcal{A}_{G+1}^{trial}$) which are combined with the parent population to form population pool for next generation ($\mathcal{A}_{G+1}^{all} = \mathcal{A}_{G}^{parent} \cup \mathcal{A}_{G+1}^{trial}$). Selection of candidates from this population pool ($\mathcal{A}_{G+1}^{all}$) helps in retaining the good solutions over generations.

Conventional Elitist Strategy

The size of the population pool ($\mathcal{A}_{G+1}^{all}$) will exceed n_{pop}. For generating a population ($\mathcal{A}_{G+1}^{parent}$) of size n_{pop}, the population pool ($\mathcal{A}_{G+1}^{all}$) has to be trimmed by selecting the candidates which will go to $\mathcal{A}_{G+1}^{parent}$. For this elitist selection, non-dominated sorting [10, 47, 141] is the first step where the population pool $\mathcal{A}_{G+1}^{all}$ is partitioned into several ranks ($\{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_l, \cdots\}$) such that the following properties are satisfied:

- Solutions within each of the l^{th} rank (i.e., $(\mathbf{X}_i, \mathbf{X}_j) \in \mathcal{R}_l \times \mathcal{R}_l$) are non-dominated with respect to each other.
- Each solution in $\mathcal{R}_l$ is dominated by at least one solution in $\mathcal{R}_{l'}$ where $l' < l$.
- Each solution in $\mathcal{R}_l$ dominates at least one of the remaining solutions from $\mathcal{A}_{i,G+1}^{all} - \{\mathcal{R}_1 \cup \mathcal{R}_2 \cup \cdots \cup \mathcal{R}_l\}$.

The concept of non-dominated sorting is demonstrated in Fig. 2.1a where the population pool $\mathcal{A}_{G+1}^{all}$ has three ranks of solutions ($\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$).



(a) (b)

Figure 2.1: (a) Ranks of solutions, crowding distance on $\mathcal{R}_1$ solutions (along f_1) and distance from ideal point [142], (b) conventional elitist framework [142].

The conventional next step is to fill the final population by starting from rank-one ($\mathcal{R}_1$) candidates, allowing the lower-ranked candidates to fill the population until the population size reaches n_{pop}. The candidates of the last essential rank (which might not be fully accommodated) are sorted by crowding distance (D_{crowd}) [10, 47] and the lesser crowded candidates are allowed to enter the population until the population size reaches n_{pop}. This method is shown in Fig. 2.1b and it performs satisfactorily for MOO problems.

Problem with the Existing Approach

In several studies [10,69], it is shown that the population pool ($\mathcal{A}_{G+1}^{all}$) gets saturated with non-dominated solutions ($\mathcal{R}_1$) towards the early generations of MOEAs for problems with 10 or higher objectives. Thus, conventional trimming of the population pool leads to a higher amount of non-dominated solutions. However, as classical DE operators cannot avoid local optima [42, 164], retaining dominated (non-optimal) candidates can help to steer the search in other directions.

Alteration to Conventional Elitist Strategy

For avoiding such a sub-optimal scenario, the second step of the elitist strategy is modified. A novel ranking strategy (described in Section 2.3.3) is used to rearrange the solutions within each rank. Following this, a mixture of mostly rank-one solutions ($\mathcal{R}_1$) and a few solutions from remaining ranks ($\mathcal{R}_{rest}$) are used to create $\mathcal{A}_{G+1}^{parent}$ whose size is n_{pop}. This step is performed after the non-dominated sorting of $\mathcal{A}_{G+1}^{all}$ and is outlined in Fig. 2.2a. The proportion of $\mathcal{R}_1$ and $\mathcal{R}_{rest}$ solutions are regulated by β (in the range $[0, 100]$) which should usually be high to prefer the non-dominated solutions. For DECOR [142], β is chosen as 75. This elitist selection strategy overcomes the saturation problem and hence, is applicable for MaOO problems with higher objectives.

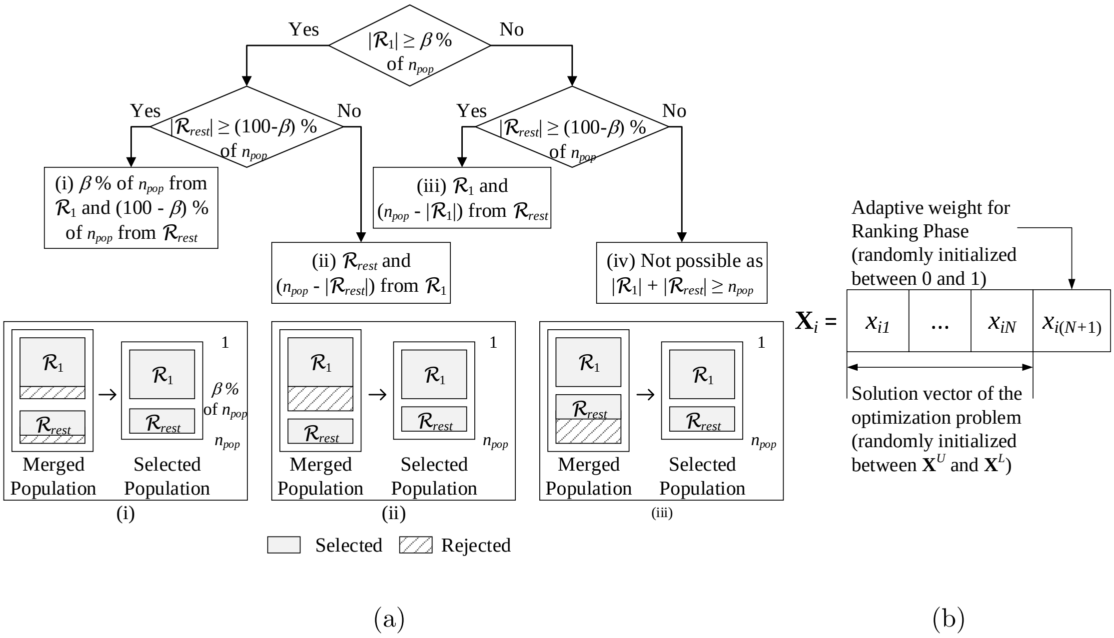


Figure 2.2: (a) Selection step of IDEMO [142] to create a population of size n_{pop}, (b) candidate representation for IDEMO [142].

2.3.3 Improved Ranking Strategy at the Selection Stage

For single-objective optimization, the candidates can be rearranged in ascending/descending order of fitness values for minimization/maximization problem. Ranking of solutions is not that simple for MaOO as Pareto-dominance (Eq. (1.3)) is not a total order relation.

Conventional Ranking Strategy

Primary sorting using non-dominated ranking and secondary sorting using crowding-distance is the conventional ranking strategy, as demonstrated in Fig. 2.1b. For the i^{th} candidate within the last essential rank l of solutions ($\forall \mathbf{X}_i \in \mathcal{R}_l$), crowding distance ($D_{crowd}^i$) is evaluated along the j^{th} objective (f_j) as follows:

1. The candidates having maximum (f_j^{max}) and minimum (f_j^{min}) objective values are assigned a crowding distance of 1000.

2. For the remaining candidates, the crowding distance is proportional to the perimeter of the hyper-rectangle formed by the normalized objectives of the candidates that precede and succeed the corresponding candidate ($\mathbf{X}_i$) in terms of f_j.

Finally, the distances are summed across all the objectives (of full or reduced set) to give the crowding distance corresponding to a candidate (D^i_{crowd}) as follows:

$$D^i_{crowd} = D_{crowd}(\mathbf{X}_i) = \sum_{j=1}^{a} \left(D_{crowd}(\mathbf{X}_i|f_j)\right), \text{ where } a = M \text{ or } a = m, \text{ and}$$
$$D_{crowd}(\mathbf{X}_i|f_j) = \begin{cases} 1000, & \text{if } f_j(\mathbf{X}_i) = f_j^{max} \text{ or } f_j(\mathbf{X}_i) = f_j^{min} \\ \left|\frac{f_j(\mathbf{X}_{i-1}) - f_j(\mathbf{X}_{i+1})}{f_j^{max} - f_j^{min}}\right|, & \text{otherwise.} \end{cases} \quad (2.5)$$

In Eq. (2.5), $\mathbf{X}_{i-1}$, $\mathbf{X}_i$ and $\mathbf{X}_{i+1}$ are assumed to be consecutive when $\mathcal{R}_l$ is sorted with respect to the j^{th} objective, $f_j(.)$. Thus, within a frontier (solutions of a particular rank), a higher perimeter means the neighbors of a candidate are far away and hence, lesser is the crowding of the candidate and its surrounding areas (implying better diversity of the frontier). This concept of crowding distance is illustrated in Fig. 2.1a.

Problem with Existing Approach

This approach heavily weighs the boundary solutions of each objective. Hence, with an increased number of objectives, the population ends up having a higher number of candidates representing only the bordering points of the estimated PF. While these candidates are essential for estimating PF, a non-uniform solution distribution will hamper the diversity if the majority of the population consists of these bordering solutions.

Alteration to Conventional Ranking Strategy

For retaining the solutions towards the center of the estimated PF, after the non-dominated sorting, the candidates can be ranked in ascending order in terms of the distance (D^i_{ideal}) between the objective vector ($\mathbf{F}(\mathbf{X}_i)$) and the ideal point (Eq. (1.8)). The concept of D_{ideal} is illustrated in Fig. 2.1a and is mathematically given as follows:

$$D^i_{ideal} = D_{ideal}(\mathbf{X}_i) = \sqrt{\sum_{j=1}^{a} \left(f_j(\mathbf{X}_i) - f_j^{ide}\right)^2}, \text{ where } a = m \text{ or } a = M. \quad (2.6)$$

For the test-suite (DTLZ) under consideration, $\mathbf{F}^{ide}$ is the origin of the objective space. The idea to rank the population $\mathcal{A}_{G+1}^{all}$ according to D_{ideal} is that given a frontier, shorter distance to ideal point implies better convergence of the candidate in objective space.

This ranking strategy considers equal preference among all the objectives. If the preference varies, weighted Euclidean distance or some other distance metric could be used which is suitable to the application under consideration. However, the problem with this method is that it ignores the information about the diversity of the approximated PF.

As a trade-off between the two ranking strategies (ranking based on D_{crowd} and ranking based on D_{ideal}), the weighted combination (D^i_{comb}) of the crowding distance (D^i_{crowd}) and distance from ideal point (D^i_{ideal}) is considered as follows:

$$D^i_{comb} = w^i \times \frac{1}{D^i_{crowd}} + (1 - w^i) \times D^i_{ideal} \text{ where } w^i = x_{i(N+1)}. \tag{2.7}$$

In Eq. (2.7), the weight (w^i) is adaptively selected between 0 and 1 by optimizing it as the last element of the candidate vector, i.e., $x_{i(N+1)}$. Hence, the dimension of the candidate vector increases from N to $(N + 1)$, as explained in Fig. 2.2b. It should be mentioned that a $(N+1)$-dimensional candidate (including the last element) goes through all the stages of IDEMO, except during the selection stage, the N-dimensional sub-vector is considered for computing the objectives of the test problem.

Having described IDEMO, the next section presents the objective reduction approach and its integration with IDEMO to form the MaOO algorithm of DECOR.

2.4 Objective Reduction based Optimization - DECOR

This section presents DECOR, which uses correlation-based online objective reduction and IDEMO (described in Section 2.3).

2.4.1 Correlation Distance

Linear correlation coefficient has been used to measure the degree of conflict among the objectives in several existing objective reduction algorithms [10,86,141]. In DECOR [142], correlation distance ($D_C(.) \in [0, 2]$) [87] is used which is defined between two vectors,

$\mathbf{Y} = [y_1, \cdots, y_n]$ and $\mathbf{Z} = [z_1, \cdots, z_n]$, as follows:

$$D_C(\mathbf{Y}, \mathbf{Z}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(z_i - \bar{z})^2}}. \tag{2.8}$$

2.4.2 Objective Reduction Principle

For implementing objective reduction, the conflicting objectives are to be identified. From the current population, the rank-one ($\mathcal{R}_1$) solutions represent the estimated PF. Distances ($D_C(.)$) between all the objective pairs are noted, where the i^{th} objective ($f_i(.)$) is estimated as shown in Fig. 2.3a using the candidates from $\mathcal{R}_1$ solutions as follows:

$$f_i(.) = \left[f_i(\mathbf{X}_1), \cdots, f_i(\mathbf{X}_{|\mathcal{R}_1|})\right], \text{ where } i = 1, \cdots, a. \tag{2.9}$$

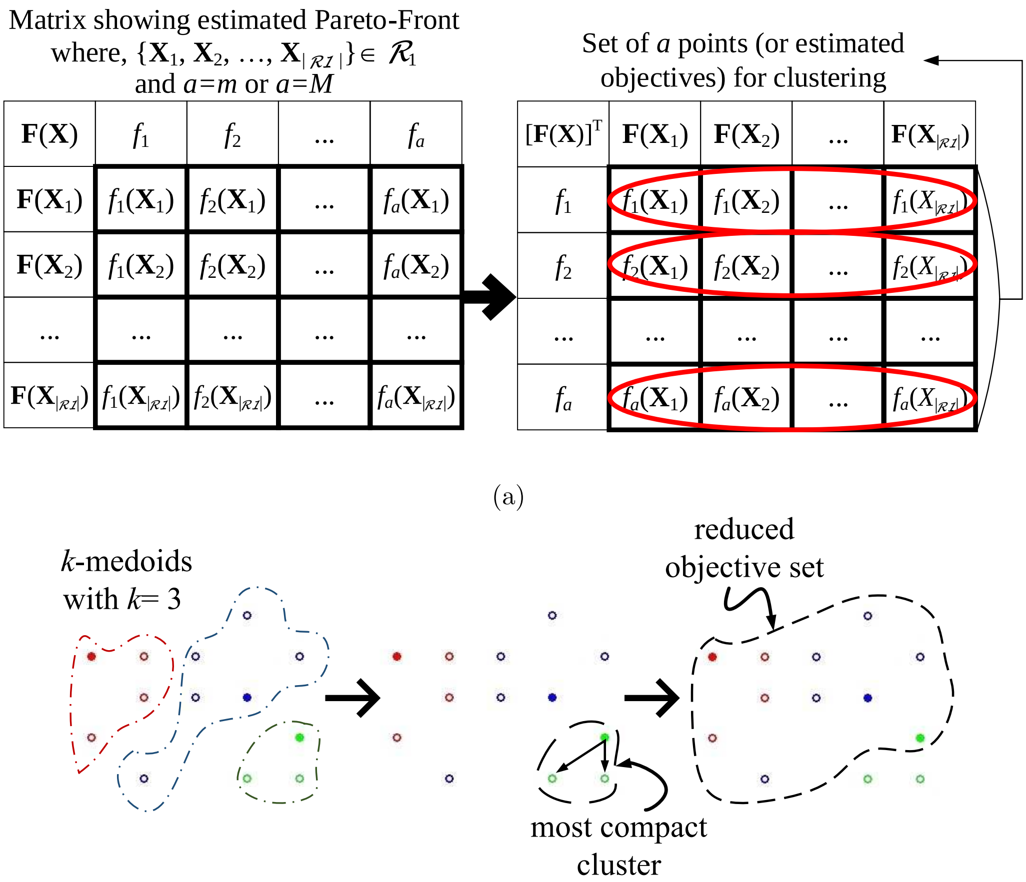
As more number of solutions contribute to form the $\mathcal{R}_1$ solutions, better is the estimate of the objectives (f_1 to f_a, with $a = M$ for full objective set and $a = m$ for reduced objective set). Closer the objectives in terms of $D_C(.)$ (Eq. (2.8)), more is the correlation between these objectives and hence, these are less conflicting than other objective pairs.

The central idea for objective reduction is to cluster the estimated objectives and eliminate all the neighbors of the cluster center from the *most compact cluster* while retaining its center. This approach allows the elimination of multiple objectives at a time. Thus, less conflicting objectives can be eliminated to yield the Non-Conflicting Objective Set (defined in Section 1.2.5). An example is shown in Fig. 2.3b.

2.4.3 Specifications of Clustering

For clustering, k-medoids is used which selects a constituent (real) objective vector of a cluster as the cluster center. It is implemented using Partitioning Around Medoid (PAM) [100] with correlation distance representing the similarity among data points for clustering. This clustering step is characterized by the following specifications:

- Clustering partitions a ($= M$ or m) objectives $\{f_1, f_2, \cdots, f_a\}$ in to k clusters $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_k\}$
- $\exists j$ such that $f_i \in \mathcal{C}_j$, and $\nexists(j_1, j_2)$ such that $f_i \in \mathcal{C}_{j_1}$ and $f_i \in \mathcal{C}_{j_2}$, where $i = 1, 2, \cdots, a$ (hard clustering, non-overlapping clusters)
- $\sum_{j=1}^{k} |\mathcal{C}_j| = a$ and $1 \leq |\mathcal{C}_j| \leq (a - k - 1)$, where $j = 1, 2, \cdots, k$



(a)

(b)

Figure 2.3: (a) Construction of the data-points (representatives of each of the objectives encircled in red) for clustering [142], (b) objective reduction principle to illustrate elimination of non-medoids from the most-compact cluster where filled circles are cluster medoids and empty circles are non-medoids [142].

- $\forall f_i \in \mathcal{C}_j$, either it is the medoid ($f_i = \mathcal{C}_j^{med}$) or it belongs to the non-medoid set ($f_i \in \mathcal{C}_j^{nmed}$)

2.4.4 Concept of Most Compact Cluster

The next step after clustering is to choose the most compact cluster ($\mathcal{C}_{com}$). For each cluster $\mathcal{C}_j$, the sum of medoid to non-medoid correlation distance is calculated. Then, $\mathcal{C}_{com}$ is the cluster having the minimum value of this sum over the current estimate of PF. Using Eq. (2.8) for $D_C(.)$, the most compact cluster ($\mathcal{C}_{com}$) is given as follows:

$$\mathcal{C}_{com} = \arg\min_{j=1}^{k} \sum_{i=1}^{|\mathcal{C}_j^{nmed}|} D_C(\mathcal{C}_j^{med}, f_i) \text{ where } f_i \in \mathcal{C}_j^{nmed}. \tag{2.10}$$

2.4.5 Problems with Singleton Clusters

A special case needs to be considered during the determination of $\mathcal{C}_{com}$ by Eq. (2.10). This special case arises when clustering results in one or more singleton clusters (i.e., $|\mathcal{C}_j| = 1$) which trivially are the most compact clusters. As there is no neighbor in a singleton cluster, no objective reduction occurs. If such a singleton cluster keeps on appearing in successive stages, the objective reduction gets stuck.

A possible way to avoid this problem of singleton cluster is to consider the most compact non-singleton cluster for objective reduction. However, a scenario may arise when M objectives are clustered into k clusters such that there are $(k-1)$ singleton clusters and one non-singleton cluster with $(M-k-1)$ objectives. Objective reduction from such a non-singleton cluster will yield k objectives in the reduced set. As incrementing k will not be possible any further, the objective reduction procedure will terminate prematurely.

An illustration to demonstrate both these extreme scenarios arising in the presence of singleton clusters is given in Fig. 2.4. Hence, a trade-off approach has to be adopted to compromise between these two extreme cases in the objective reduction stage.

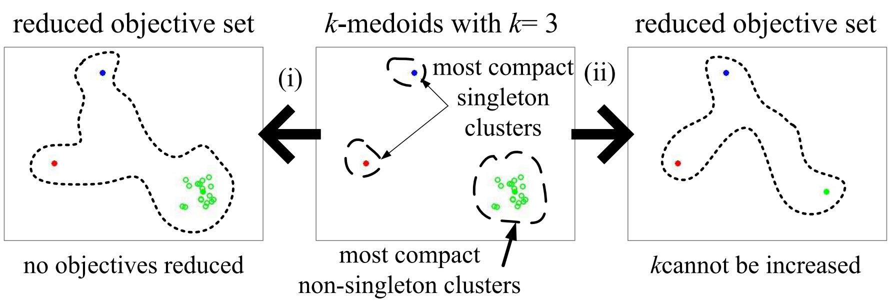


Figure 2.4: Issues in objective reduction due to the presence of singleton clusters [142]: (i) scenario when a singleton cluster is considered as the most compact cluster, (ii) scenario when a non-singleton cluster is considered as the most compact cluster but there are $(k-1)$ singleton clusters and only one non-singleton cluster.

2.4.6 A Solution to Handle Singleton Clusters

There can be two possible scenarios, while singleton clusters are encountered. These scenarios and the strategies to tackle these scenarios are as follows:

1. *Case-1:* When the singleton cluster is relatively far away from its nearest cluster, it forms a crucial objective and is conflicting with other clusters of objectives. Hence, it is directly added to the reduced set of objectives, and the next most compact cluster is analyzed, immediately. Thus, objective reduction continues without getting stuck.

2. *Case-2:* When the singleton cluster is comparatively closer to its nearest cluster, there can be a possibility that the singleton cluster is resultant because of the current value of k (number of clusters) and/or the present estimate of PF. In such a scenario, the singleton cluster itself is considered as the most compact cluster so that no objective reduction occurs and k is increased in successive turns for finer clustering.

In order to differentiate whether the singleton cluster ($\mathcal{C}_j$) is relatively far away or closer to its nearest cluster, an indicator (D^j_{ratio}) is defined as follows:

$$D^j_{ratio} = \frac{D^j_{near}}{D_{neigh}}, \text{ where} \tag{2.11}$$

$$D^j_{near} = \min_{j'=1,j'\neq j}^{k} D_C(\mathcal{C}^{med}_j, \mathcal{C}^{med}_{j'}), \text{ with } \mathcal{C}^{med}_j \text{ as a singleton cluster, and} \tag{2.12}$$

$$D_{neigh} = \max_{j=1}^{k} D_C(\mathcal{C}^{med}_j, f_i), \text{ where } f_i \in \mathcal{C}^{nmed}_j. \tag{2.13}$$

In Eq. (2.11), D^j_{near} is the correlation distance between the singleton cluster ($\mathcal{C}_j$) and its nearest cluster's medoid, and D_{neigh} is the maximum of intra-cluster medoid to non-medoid correlation distance over all the clusters at the present state.

If D^j_{ratio} is at least greater than some threshold (th), implying that the singleton cluster ($\mathcal{C}_j$) is relatively far away from its nearest cluster, it is considered as a conflicting objective and is directly added to the reduced objective set (by Case-1). On the other hand, if D^j_{ratio} is less than the threshold (th) implying the singleton cluster ($\mathcal{C}_j$) is comparatively closer to its nearest cluster, no objective reduction occurs in the present state (by Case-2). For DECOR [142], the value of this threshold (th) is chosen by trial and error.

2.4.7 Selecting the Number of Clusters

Next issue with this clustering approach of objective reduction is choosing the number of clusters (i.e., the value of k). In DECOR [142], the objective reduction procedure starts with a small value of k, and after G_{op} generations of the optimization algorithm, the value of k is increased by 1 at a time. This increment happens periodically until k equals the number of objectives because clustering is not possible with k higher than this value. Through this step, most of the clusters are explored eventually, even if these were not declared as the most compact cluster in earlier stages.

Reasons for Online Objective Reduction

As $\mathcal{R}_1$ is an estimate of the PF, the true correlation between the objectives or the exact groups of conflicting objectives cannot be determined at an early stage. Due to this evolving nature of the correlation structure, the online version of objective reduction is adopted, where the optimization alternates between global exploration (with full objective set) and local exploitation (with reduced objective set). The switching parameter (G_{sw}) regulates the number of generations after which the global and the local search toggles.

Two Different Versions of DECOR

Automatic-DECOR (aDECOR): The flowchart describing the framework (automatic DECOR or aDECOR) obtained by integrating the objective reduction and the optimization algorithm is shown in Fig. 2.5a and Algorithm 2.1, which in turn calls the objective reduction procedure as described in Algorithm 2.2. It is called automatic because the number of objectives of the reduced set is automatically determined by the algorithm and does not involve any input from the user.

Using aDECOR is problematic when for sufficiently large M (number of objectives) and very small k (number of clusters), there is a very high number of objectives in each of the clusters. In such a situation, objective reduction from the most compact cluster leads to the elimination of a high number of objectives. For example, when $k = 2$ and M/k objectives are there in each cluster, nearly half of the objectives are eliminated. Such removal is undesirable as the estimation of the conflicting objectives is poor at early stages.

Fixed-DECOR (fDECOR): For avoiding the problem of aDECOR, the fixed version of DECOR (fixed DECOR or fDECOR) is developed. It is called fixed because the size of the reduced objective set is lower-bounded by a fixed user-specified value of k. Clustering yields k clusters and in an extreme case like in (ii) of Fig. 2.3b, the reduced objective set has at least k objectives. So, for MaOO problems with very high M, specifying a high k will not lead to a drastic reduction in the number of objectives. It should be noted that although k is not incremented periodically in fDECOR, the objective reduction module is executed periodically to ensure that the correlation structure is evolved along with the evolution of the objective vectors.

The flowchart describing fDECOR is shown in Fig. 2.5b and Algorithm 2.3, which in turn also calls the objective reduction procedure as described in Algorithm 2.2.

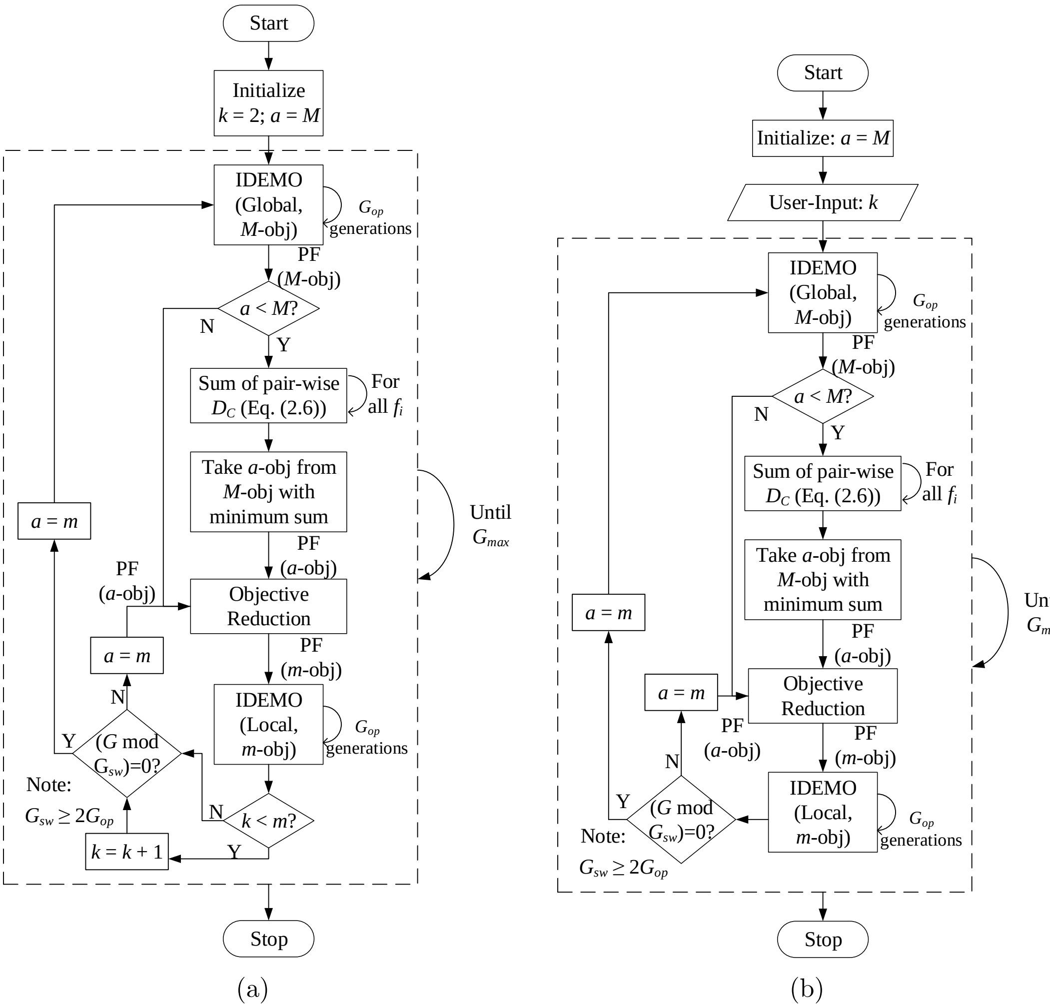


Figure 2.5: Flowchart describing the framework of DECOR [142]: (a) automatic-DECOR (or aDECOR), (b) fixed-DECOR (or fDECOR).

Computational Complexity of the Objective Reduction Module

The time complexity of Algorithm 2.2 (objective reduction) is obtained by considering that line 2 requires $\mathcal{O}(k.(a-k)^2.|\mathcal{R}_1|^2.G_{PAM})$ [100] operations, lines 3 and 14 requires $\mathcal{O}((a-k).|\mathcal{R}_1|)$ operations, lines 4 to 13 require $\mathcal{O}(a+k^2.|\mathcal{R}_1|)$ operations and line 15 requires $\mathcal{O}(a)$ operations. It is equivalent to $\mathcal{O}(k.(a-k)^2.|\mathcal{R}_1|^2.G_{PAM})$ where a is the size of $\mathbf{F}'_a$, k is the number of cluster, $|\mathcal{R}_1|$ is the dimension of each objective representative (Eq. 2.9) for clustering and G_{PAM} is the number of iterations needed by k-medoid.

The designed framework of DECOR is applied on several benchmark problems and the associated results are analyzed in the subsequent sections.

Algorithm 2.1 Complete Framework of aDECOR [142]

Input: $\mathbf{F}$: Objective functions for M-objective MaOO problem
Output: $\mathcal{A}_G^{parent}$: Estimated PS; $\mathcal{A}_{\mathbf{F},G}^{parent}$: Estimated PF

Initialize: $k = 2$; $a = M$; $G = 0$; $flag = 1$;
while $G < G_{max}$ (GLOBAL SEARCH) **do**
  Execute IDEMO on full objective set (of size M) for G_{op} generations, set $flag = 0$ when population has no new candidate, and evolve $\mathcal{A}_G^{parent}$
  $G = G + G_{op}$
  if flag=0 (if no new candidates are found) **then**
    Break;
  end if
  if $a < M$ **then**
    $\forall f_i$, $S_i^{DC} = \sum_{j=1, j\neq i}^{M} D_C(f_i, f_j)$ (Eq. (2.8))
    Sort all f_i based on S_i^{DC} in descending order
    Form $\mathbf{F}'_a$ with top a conflicting objectives
  end if
  while $(G) mod(G_{sw}) \neq 0$ (LOCAL SEARCH) **do**
    $\mathbf{F}'_m = ObjRed(\mathcal{A}_{\mathbf{F},G}^{parent}, \mathbf{F}'_a, k)$ (Algorithm 2.2)
    Execute IDEMO on reduced objective set (of size m) for G_{op} generations, set $flag = 0$ when population has no new candidate, and evolve $\mathcal{A}_G^{parent}$
    $G = G + G_{op}$
    if flag=0 (if no new candidates are found) **then**
      Break;
    end if
    if $k < m$ **then**
      $k = k + 1$;
    end if
    $a = m$;
  end while(END OF LOCAL SEARCH)
  if flag=0 (if no new candidates are found) **then**
    Break;
  end if
end while(END OF GLOBAL SEARCH)

2.5 Results

The performance of aDECOR and fDECOR are presented and analyzed in this section. The algorithms are executed on a computer having 4GB RAM and Intel Core i3 processor @2.30GHz, using the 32-bit version of MATLAB R2012b.

The performance of DECOR is compared with NSGA-II [47], MOEA/D [150], HypE [9], DEMO [153], α-DEMO-revised [10] and the optimization approach of [141]. The reasons for selecting these algorithms for comparison are as follows:

- Performance of some notable MOEAs are considered: NSGA-II [47] (pioneering work for crowding distance-based ranking), HypE [9] (indicator-based ranking) and MOEA/D [150] (decomposition of objective space into sub-spaces).

Algorithm 2.2 Procedure for Objective Reduction to assist DECOR [142]

Input: $\mathcal{A}^{parent}_{\mathbf{F},G}$: Objective vectors of population candidates; $\mathbf{F}'_a$: objective set of size a; k: number of clusters
Output: $\mathbf{F}'_m$ Objective set of size m where $m \leq a$
1: **procedure** ObjRed($\mathcal{A}^{parent}_{\mathbf{F},G}$, $\mathbf{F}'_a$, k)
2: Execute k-medoids on $\mathcal{A}^{parent}_{\mathbf{F}'_a,G}$ using $D_C(.)$ (Eq. (2.8))
3: Find D_{neigh} (Eq. (2.13))
4: **for** $i = 1$ to k (for all clusters) **do**
5: Initialize $Flag_i = 0$
6: Obtain $|\mathcal{C}_i|$
7: **if** $|\mathcal{C}_i| = 1$ (if $\mathcal{C}_i$ is a singleton cluster) **then**
8: Find D^i_{near} (Eq. (2.12)) and D^i_{ratio} (Eq. (2.11))
9: **if** $D^i_{ratio} \geq th$ **then**
10: Set $Flag_i = 1$ (to ignore $\mathcal{C}_i$ in Step 14)
11: **end if**
12: **end if**
13: **end for**
14: Find $\mathcal{C}_{com}$ using Eq. (2.10) (ignore $\mathcal{C}_i$ if $Flag_i = 1$)
15: Construct $\mathbf{F}'_m = \mathbf{F}'_a - \{f_j | f_j \in \mathcal{C}^{nmed}_{com}\}$
16: **return** $\mathbf{F}'_m$
17: **end procedure**

- As DECOR uses IDEMO (an improved version of DEMO), the performance of classic DEMO [153] is considered.
- As DECOR is an extension of the work from [141], the results of [141] are considered.
- Finally, α-DEMO-revised [10] is considered as it is a contemporary objective reduction based MaOO algorithm.

In this experiment, DECOR is tested on DTLZ1, DTLZ2, DTLZ3 and DTLZ4 problems [50] for 10 and 20 objectives. The mean and standard deviation of the performance values (convergence metric [10,134,141] and hypervolume indicator [10,134,141]) are considered as reported in [10,141] for NSGA-II, MOEA/D, HypE, DEMO, α-DEMO-revised and for the approach of [141]. DECOR is implemented using the source code available at `http://decor.droppages.com/`.

2.5.1 Parameter Specifications

For executing DECOR and recording the performance metrics, several parameters have to be set. DECOR [142] is executed with various parameter settings and best results are obtained for the parameters which are specified in Table 2.1.

Algorithm 2.3 Complete Framework of fDECOR [142]

Input: $\mathbf{F}(.)$ Objective functions for M-objective MaOO problem; k: Number of clusters (lower bound of m)
Output: $\mathcal{A}_G^{parent}$: Estimated PS; $\mathcal{A}_{\mathbf{F},G}^{parent}$: Estimated PF

Initialize: $a = M$; $G = 0$; $flag = 1$;
User-Input: k;
while $G < G_{max}$ (GLOBAL SEARCH) **do**
    Execute IDEMO on full objective set (of size M) for G_{op} generations, set $flag = 0$ when population has no new candidate, and evolve $\mathcal{A}_G^{parent}$
    $G = G + G_{op}$
    if flag=0 (if no new candidates are found) **then**
        Break;
    end if
    if $a < M$ **then**
        $\forall f_i$, $S_i^{DC} = \sum_{j=1, j\neq i}^{M} D_C(f_i, f_j)$ (Eq. (2.8))
        Sort all f_i based on S_i^{DC} in descending order
        Form $\mathbf{F}'_a$ with top a conflicting objectives
    end if
    while $(G)mod(G_{sw}) \neq 0$ (LOCAL SEARCH) **do**
        $\mathbf{F}'_m = ObjRed(\mathcal{A}_{\mathbf{F},G}^{parent}, \mathbf{F}'_a, k)$;
        Execute IDEMO on reduced objective set (of size m) for G_{op} generations, set $flag = 0$ when population has no new candidate, and evolve $\mathcal{A}_G^{parent}$
        $G = G + G_{op}$
        if flag=0 (if no new candidates are found) **then**
            Break;
        end if
        $a = m$;
    end while(END OF LOCAL SEARCH)
    if flag=0 (if no new candidates are found) **then**
        Break;
    end if
end while(END OF GLOBAL SEARCH)

Table 2.1: Recommended values of different parameters used for DECOR [142].

Parameters	Explanation	Values
n_{pop}	Population size	100
G_{max}	Maximum generations	2000
CR	Crossover Rate	0.8
G_{op}	Number of generations for which IDEMO runs at a time	20
G_{sw}	Number of generations after which the algorithm switches from reduced to full objective dimension	100
β	Percentage of $\mathcal{R}_1$ solutions	75
th	Threshold on D_{ratio} for considering singleton clusters during objective reduction	Chosen from $\{1.2, 1.5, 2.0\}$
$\lvert\mathcal{H}_{CM}\rvert$	Size of sampled set for Convergence Metric	5000
$\lvert\mathcal{H}_{HV}\rvert$	Size of sampled set for Hypervolume Indicator	10000
$\mathbf{R}_{HV}$	Reference point (objective vector) for hypervolume	$\{3, \overset{M}{\cdots}, 3\}$

Increasing n_{pop} while keeping G_{max} fixed, does not improve performance as comparatively more random initialization of candidates occurs rather than mutation and recombination, which implies a proportionately lesser number of good solutions being propagated. Again, keeping n_{pop} fixed and increasing G_{max}, does not improve the results any further. The parameter CR is kept high in order to generate a solution far from the parent candidate. However, with $CR > 0.8$, the trial vector is mostly independent of the parent candidate, which leads to poor performance. The parameter G_{op} is the minimum number of generations over which some significant change in performance (change in total $D_{ideal} > 10^{-2}$) is observed. Incrementing G_{sw} leads to more local search and thus, poor performance at the global level whereas decreasing G_{sw} slows down objective reduction. When $\beta > 75$, more sub-optimal non-dominated solutions are passed on to next generations, whereas when $\beta < 75$, more number of potential $\mathcal{R}_1$ solutions are not propagated, resulting in poor performance. For setting the reference point ($\mathbf{R}_{HV}$), $|\mathcal{H}_{HV}|$ and $|\mathcal{H}_{CM}|$ for performance metrics, the work in [10] is consulted.

2.5.2 Performance of aDECOR

DECOR [142] involves a thresholding on D_{ratio} for deciding how to process a singleton cluster in objective reduction procedure. As D_{near} is always greater than D_{neigh}, D_{ratio} is greater than 1. But a D_{ratio} higher than 2 implies the singleton cluster is very far away from the nearest cluster to be considered as an essential cluster. Hence, the threshold (th) on D_{ratio} is chosen in the range of $(1, 2]$. For sensitivity study, DECOR samples th from $\{1.2, 1.5, 2.0\}$ and reports those values in the th column in Tables[1] 2.2, 2.3, 2.4 and 2.5 which provide the best performance for aDECOR and fDECOR.

The performance of aDECOR is mentioned in Tables 2.2 and 2.3, in terms of convergence metric and hypervolume indicator, respectively. For evaluating convergence metric, a sampled version of the true PF is obtained by consulting [35]. NSGA-II, MOEA/D, HypE and DEMO are not objective reduction based optimization algorithms. Hence, the results are obtained using the full objective set for these four algorithms, i.e., $M = m$ =10 (or 20). For each problem, the value of m automatically determined by aDECOR is used as an input for α-DEMO-revised [10] and is mentioned in Tables 2.2 and 2.3 in the format DTLZ$type(M \to m)$, where $type$ indicates the type of the DTLZ problem. It should be noted that the approach of [141] and aDECOR cannot be compared directly as both of these automatically determine m, and thus, the final m values are often different.

[1]In this thesis, across all the tables reporting experimental results, the best and second-best performing values are highlighted in dark and light shades of gray, respectively.

Table 2.2: Mean and standard deviation of convergence metric over 50 independent runs for comparing MOEAs with aDECOR [142].

Problem type	*th*	**NSGA-II**	**MOEA/D**	**HypE**	**DEMO**	**α-DEMO-revised**	**aDECOR**
DTLZ1(10→6)	1.2	225.4502 ± 5.9816	2.4800 ± 1.0351	146.3039 ± 2.2147	142.2519 ± 3.1073	1.0291 ± 0.0061	0.3993 ± 0.0042
DTLZ1(20→10)	1.5	176.2357 ± 3.6600	3.2397 ± 1.1651	305.1945 ± 9.7488	143.5408 ± 2.7434	1.2356 ± 0.0210	0.3307 ± 0.0310
DTLZ2(10→6)	1.5	1.4716 ± 0.0317	0.7419 ± 0.0101	1.3979 ± 0.0156	1.3891 ± 0.0161	1.3858 ± 0.0907	0.4088 ± 0.0111
DTLZ2(20→10)	1.5	1.9273 ± 0.0224	1.3116 ± 0.0050	1.9240 ± 0.0144	1.9009 ± 0.0092	1.1112 ± 0.0172	0.4696 ± 0.0177
DTLZ3(10→6)	1.2	1048.0740 ± 39.3631	24.8627 ± 4.5587	409.5137 ± 3.9870	939.7426 ± 9.8824	1.0011 ± 0.0245	0.5256 ± 0.0153
DTLZ3(20→10)	1.5	978.3490 ± 44.9975	37.8409 ± 7.2125	911.8077 ± 5.5582	1024.4046 ± 12.5577	94.6363 ± 1.5260	0.4925 ± 0.0293
DTLZ4(10→6)	1.5	1.1784 ± 0.0264	0.7461 ± 0.0102	0.8914 ± 0.0106	1.2663 ± 0.0347	1.5048 ± 0.0306	0.4768 ± 0.0092
DTLZ4(20→11)	1.5	1.4337 ± 0.0309	1.0818 ± 0.0070	0.9572 ± 0.0077	1.6816 ± 0.0370	1.4034 ± 0.0197	0.4768 ± 0.0307

Table 2.3: Mean and standard deviation of hypervolume indicator over 50 independent runs for comparing MOEAs with aDECOR [142].

Problem type	*th*	**NSGA-II**	**MOEA/D**	**HypE**	**DEMO**	**α-DEMO-revised**	**aDECOR**
DTLZ1(10→6)	1.2	0.0044 ± 0.0061	0.8132 ± 0.0984	0.0000 ± 0.0000	0.0000 ± 0.0000	0.1385 ± 0.0092	0.9915 ± 0.0098
DTLZ1(20→10)	1.5	0.0000 ± 0.0000	0.7233 ± 0.1172	0.0000 ± 0.0000	0.0000 ± 0.0000	0.9779 ± 0.0176	0.9994 ± 0.0206
DTLZ2(10→6)	1.5	0.8399 ± 0.0079	1.0000 ± 0.0000	0.9514 ± 0.0034	0.8863 ± 0.0059	0.9103 ± 0.0675	0.8765 ± 0.0018
DTLZ2(20→10)	1.5	0.8280 ± 0.0070	1.0000 ± 0.0000	0.9372 ± 0.0019	0.8487 ± 0.0059	0.9213 ± 0.0150	0.8016 ± 0.0050
DTLZ3(10→6)	1.2	0.0000 ± 0.0000	0.0235 ± 0.0388	0.0000 ± 0.0000	0.0000 ± 0.0000	0.2967 ± 0.0013	0.9879 ± 0.0105
DTLZ3(20→10)	1.5	0.0000 ± 0.0000	0.0301 ± 0.0391	0.0000 ± 0.0000	0.0000 ± 0.0000	0.3213 ± 0.0055	0.9964 ± 0.0098
DTLZ4(10→6)	1.5	0.9765 ± 0.0056	1.0000 ± 0.0000	0.8741 ± 0.0169	0.9956 ± 0.0012	0.1786 ± 0.0313	0.9488 ± 0.0072
DTLZ4(20→11)	1.5	0.9914 ± 0.0030	1.0000 ± 0.0000	0.8963 ± 0.0103	0.9829 ± 0.0111	0.9018 ± 0.0761	0.9420 ± 0.0155

2.5.3 Performance of fDECOR

The execution of fDECOR is also compared with NSGA-II, MOEA/D, HypE, and DEMO, which are not objective reduction based MOEAs. Hence, these MOEAs are executed on the full objective set, i.e., $M = m = 10$ (or 20). The performance of the objective reduction based MaOO algorithms of [10, 141] are also compared. For each problem, the value of m automatically determined by the approach of [141] is used as an input to fDECOR [142] and α-DEMO-revised [10]. These values are mentioned in Tables 2.4 (reporting convergence metric) and 2.5 (reporting hypervolume indicator) in DTLZ$type(M \to m)$ format.

Table 2.4: Mean and standard deviation of convergence metric over 50 independent runs for comparing MOEAs with fDECOR [142].

Problem type	*th*	**NSGA-II**	**MOEA/D**	**HypE**	**DEMO**	**α-DEMO-revised**	**Approach of [141]**	**fDECOR**
DTLZ1(10→7)	1.2	225.4502 ± 5.9816	2.4800 ± 1.0351	146.3039 ± 2.2147	142.2519 ± 3.1073	1.1718 ± 0.0094	0.3991 ± 0.0017	0.3421 ± 0.0029
DTLZ1(20→14)	1.2	176.2357 ± 3.6600	3.2397 ± 1.1651	305.1945 ± 9.7488	143.5408 ± 2.7434	1.7954 ± 0.0565	1.0095 ± 0.0041	0.4391 ± 0.0046
DTLZ2(10→6)	1.2	1.4716 ± 0.0317	0.7419 ± 0.0101	1.3979 ± 0.0156	1.3891 ± 0.0161	1.3858 ± 0.0907	0.5214 ± 0.0069	0.3843 ± 0.0070
DTLZ2(20→12)	1.5	1.9273 ± 0.0224	1.3116 ± 0.0050	1.9240 ± 0.0144	1.9009 ± 0.0092	1.3915 ± 0.0175	1.1610 ± 0.0076	0.8412 ± 0.0183
DTLZ3(10→6)	1.2	1048.0740 ± 39.3631	24.8627 ± 4.5587	409.5137 ± 3.9870	939.7426 ± 9.8824	1.0011 ± 0.0245	0.5877 ± 0.0014	0.5256 ± 0.0153
DTLZ3(20→13)	1.2	978.3490 ± 44.9975	37.8409 ± 7.2125	911.8077 ± 5.5582	1024.4046 ± 12.5577	1.4153 ± 0.0111	0.8591 ± 0.0068	0.5211 ± 0.0034
DTLZ4(10→7)	1.5	1.1784 ± 0.0264	0.7461 ± 0.0102	0.8914 ± 0.0106	1.2663 ± 0.0347	0.5218 ± 0.0059	0.4780 ± 0.0014	0.3815 ± 0.0067
DTLZ4(20→11)	1.2	1.4337 ± 0.0309	1.0818 ± 0.0070	0.9572 ± 0.0077	1.6816 ± 0.0370	1.4034 ± 0.0197	0.4716 ± 0.0021	0.3111 ± 0.0101

Table 2.5: Mean and standard deviation of hypervolume indicator over 50 independent runs for comparing MOEAs with fDECOR [142].

Problem type	*th*	**NSGA-II**	**MOEA/D**	**HypE**	**DEMO**	**α-DEMO-revised**	**Approach of [141]**	**fDECOR**
DTLZ1(10→7)	1.2	0.0044 ± 0.0061	0.8132 ± 0.0984	0.0000 ± 0.0000	0.0000 ± 0.0000	0.9281 ± 0.0016	0.9544 ± 0.0545	0.9182 ± 0.0087
DTLZ1(20→14)	1.2	0.0000 ± 0.0000	0.7233 ± 0.1172	0.0000 ± 0.0000	0.0000 ± 0.0000	0.8743 ± 0.0081	0.9704 ± 0.0356	0.9791 ± 0.0234
DTLZ2(10→6)	1.2	0.8399 ± 0.0079	1.0000 ± 0.0000	0.9514 ± 0.0034	0.8863 ± 0.0059	0.9103 ± 0.0675	0.6054 ± 0.0037	0.6314 ± 0.0114
DTLZ2(20→12)	1.5	0.8280 ± 0.0070	1.0000 ± 0.0000	0.9372 ± 0.0019	0.8487 ± 0.0059	0.9117 ± 0.0477	0.6444 ± 0.0028	0.9127 ± 0.0167
DTLZ3(10→6)	1.2	0.0000 ± 0.0000	0.0235 ± 0.0388	0.0000 ± 0.0000	0.0000 ± 0.0000	0.2967 ± 0.0013	0.9848 ± 0.0759	0.9525 ± 0.0130
DTLZ3(20→13)	1.2	0.0000 ± 0.0000	0.0301 ± 0.0391	0.0000 ± 0.0000	0.0000 ± 0.0000	0.4830 ± 0.0035	0.9825 ± 0.0468	0.9870 ± 0.0099
DTLZ4(10→7)	1.5	0.9765 ± 0.0056	1.0000 ± 0.0000	0.8741 ± 0.0169	0.9956 ± 0.0012	0.8632 ± 0.054	0.8850 ± 0.0273	0.9095 ± 0.0301
DTLZ4(20→11)	1.2	0.9914 ± 0.0030	1.0000 ± 0.0000	0.8963 ± 0.0103	0.9829 ± 0.0111	0.9018 ± 0.0761	0.9608 ± 0.0383	0.9053 ± 0.0083

2.5.4 Statistical Analysis

For statistical validation of the results, Friedman Test [52], McNemar's Test [122] and Holm-Bonferroni Test [71] are performed.

Friedman Test

Assumptions made for the Friedman test are as follows: the null hypothesis (H_0) states that all the algorithms are ranked equally while the alternate hypothesis (H_a) states that the algorithms are ranked as mentioned in Table 2.6. Friedman statistic follows a χ^2_F-distribution with N_{DS} number of datasets, k_{algo} number of comparative algorithms and

$(k_{algo} - 1)$ degrees of freedom as given below:

$$\chi_F^2 = \frac{12N_{DS}}{k_{algo}(k_{algo}+1)} \left[\sum_{j=1}^{k_{algo}} R_{F,j}^2 - \frac{k_{algo}(k_{algo}+1)^2}{4} \right]. \tag{2.14}$$

This test considers $N_{DS} = 4$, $k_{algo} = 6$ (for aDECOR) or $k_{algo} = 7$ (for fDECOR). The average rank of the j^{th} algorithm ($R_{F,j}$) based on its performance is specified in Table 2.6.

For aDECOR, convergence metric attains $\chi_F^2 > \chi_{5,0.05}^2 = 11.07$ (critical value for five degrees of freedom and 95% confidence interval) and thus, the test rejects H_0. For hypervolume indicator, $\chi_F^2 < \chi_{5,0.05}^2 = 11.07$ and hence, it fails to reject H_0. For fDECOR also, convergence metric attains $\chi_F^2 > \chi_{6,0.05}^2 = 12.59$, and thus, the test rejects H_0. For hypervolume indicator, $\chi_F^2 < \chi_{6,0.05}^2 = 12.59$ and hence, it fails to reject H_0.

McNemar's Test

McNemar's test is another non-parametric test following χ_M^2-distribution which compares algorithms on a one-against-one basis. Assuming n_{AB} and n_{BA} denote the number of times algorithm A outperforms algorithm B and vice-versa, respectively, the McNemar's statistic is given as follows:

$$\chi_M^2 = \frac{(|n_{AB} - n_{BA}| - 1)^2}{(n_{AB} + n_{BA})}. \tag{2.15}$$

For each of the four problem types (DTLZ1 to DTLZ4) within each category ($M = 10$ or 20), each algorithm is executed 50 times and thus, the number of discordant pairs is 200 ($= 50 \times 4$). The null hypothesis ($H_{0,i}$) is assumed as the two competitor algorithms (i.e., aDECOR/fDECOR and the i^{th} algorithm listed in Table 2.6) have equal tendencies to approximate PF. When $\chi_M^2 > \chi_{1,0.05}^2 = 3.84$ (critical value for one degree of freedom and 95% confidence interval), $H_{0,i}$ is rejected.

Holm-Bonferroni Test

For multiple comparisons, this post-hoc test controls the family-wise error rate by applying Bonferroni corrections to the significance level of each individual hypotheses. The p-values of different results are listed under the McNemar's test in Table 2.6. Within each family, the p-values are ranked from lowest to highest as given by $R_{H,i}$ in Table 2.6. This test rejects those hypotheses for which p-values are smaller than the adjusted significance level (α'_{HB}) for 95% confidence interval, given as follows:

$$\alpha'_{HB} = \frac{0.05}{(k_{algo} + 1 - R_{H,i})}. \tag{2.16}$$

2.6 Discussion

Based on the results presented in Section 2.5, an analysis of the performance of DECOR is presented as follows:

Table 2.6: Parameters and results of Friedman Test (FT), McNemar's Test (MNT) and Holm-Bonferroni Test (HBT) to validate the performance of DECOR [142].

Algorithms	**10 objectives**			**20 objectives**		
	FT[a]	MNT[b]	HBT[c]	FT[a]	MNT[b]	HBT[c]
(a) aDECOR vs. others (for observations in Table 2.2)						
NSGA-II	5.50	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R	5.00	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R
MOEA/D	2.50	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R	2.75	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R
HypE	4.25	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R	4.25	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R
DEMO	4.50	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R	5.00	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R
α-DEMO-revised	3.25	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R	3.00	200, 0, 198.005, $< 10^{-5}$, R	1, 0.01, R
aDECOR	1.00	-	-	1.00	-	-
χ^2_F, H_0 (FT)	14.71, R	-	-	13.86, R	-	-
(b) aDECOR vs. others (for observations in Table 2.3)						
NSGA-II	4.50	150, 50, 49.005, $< 10^{-5}$, R	1, 0.01, R	3.75	100, 100, 0.005, 0.94363, A	2, 0.0125, A
MOEA/D	1.75	100, 100, 0.005, 0.94363, A	2, 0.0125, A	2.00	100, 100, 0.005, 0.94363, A	2, 0.0125, A
HypE	3.75	150, 50, 49.005, $< 10^{-50}$, R	1, 0.01, R	4.00	150, 50, 49.005, $< 10^{-50}$, R	1, 0.01, R
DEMO	3.50	100, 100, 0.005, 0.94363, A	2, 0.0125, A	3.75	100, 100, 0.005, 0.94363, A	2, 0.0125, A
α-DEMO-revised	3.50	160, 40, 70.805, $< 10^{-50}$, R	1, 0.01, R	3.00	140, 60, 31.205, $< 10^{-50}$, R	1, 0.01, R
aDECOR	2.75	-	-	3.00	-	-
χ^2_F, H_0 (FT)	-4.64, A	-	-	-8.43, A	-	-
(c) fDECOR vs. others (for observations in Table 2.4)						
NSGA-II	6.75	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	6.25	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
MOEA/D	4.00	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	3.75	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
HypE	5.25	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	5.25	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
DEMO	5.75	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	6.00	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
α-DEMO-revised	3.25	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	3.75	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
Approach of [141]	2.00	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R	2.00	200, 0, 198.005, $< 10^{-50}$, R	1, 0.0083, R
fDECOR	1.00	-	-	1.00	-	-
χ^2_F, H_0 (FT)	22.07, R	-	-	20.36, R	-	-
(d) fDECOR vs. others (for observations in Table 2.5)						
NSGA-II	4.50	100, 100, 0.005, 0.94363, A	4, 0.0167, A	4.50	150, 50, 49.005, $< 10^{-5}$, R	1, 0.0083, R
MOEA/D	2.50	100, 100, 0.005, 0.94363, A	4, 0.0167, A	2.50	100, 100, 0.005, 0.94363, A	3, 0.0125, A
HypE	4.75	150, 50, 49.005, $< 10^{-5}$, R	1, 0.0083, R	4.75	145, 55, 39.605, $< 10^{-5}$, R	1, 0.0083, R
DEMO	4.25	100, 100, 0.005, 0.94363, A	4, 0.0167, A	4.50	150, 50, 49.005, $< 10^{-5}$, R	1, 0.0083, R
α-DEMO-revised	3.75	95, 105, 0.405, 0.52452, A	3, 0.0125, A	4.00	140, 60, 31.205, $< 10^{-5}$, R	1, 0.0083, R
Approach of [141]	3.50	120, 80, 7.605, 0.00582, R	2, 0.01, R	3.75	125, 75, 12.005, 0.00053, R	2, 0.01, R
fDECOR	3.75	-	-	2.50	-	-
χ^2_F, H_0 (FT)	-3.86, A	-	-	-5.46, A	-	-

[a] For Friedman Test (FT): Average Ranks ($R_{F,j}$)
[b] For McNemar's Test (MNT): n_{AB}, n_{BA}, χ^2_M, p-value, Acceptance(A) or Rejection(R) of $H_{0,i}$ at 95% confidence interval
[c] For Holm-Bonferroni Test (HBT): Rank ($R_{H,i}$), Bonferroni corrected significance level (α'_{HB}) for 95% confidence interval, Acceptance(A) or Rejection(R) of $H_{0,i}$

1. *Benefit of Thresholding while Clustering*: The threshold (th) helps DECOR [142] to overcome the problems of singleton cluster (unlike [141]). It also leads to a reduced number of objective computations as seen from m determined by aDECOR (in Table 2.2), which are smaller than m determined by the approach of [141] (in Table 2.4). However, th varies with different problems and currently can only be set empirically.

2. *Convergence by DECOR*: Although superior convergence of DECOR is noted in Tables 2.2 and 2.4, its diversity is poor in some cases (Tables 2.3 and 2.5). Similar to Fig. 1.3b, this conflict in convergence metric and hypervolume indicator [134] can be due to the variation in solution distribution near the true PF which implies that performance of DECOR can further be improved.

3. *Diversity by DECOR*: Diversity of DECOR is studied from its hypervolume values. While a zero hypervolume indicates that the entire estimated PF is outside the hyper-rectangle [134], DECOR has succeeded in obtaining non-zero hypervolume in several cases. Moreover, DECOR outperforms other objective reduction based algorithms [10, 134] in 50% or more cases in Tables 2.3 and 2.5.

4. *Beneficial Attributes of DECOR*: DECOR integrates simultaneous objective reduction and optimization, allows the elimination of multiple objectives at a time, employs regulated elitism (Fig. 2.2a) to avoid dominance resistance of MaOO problems and uses a combination (D_{comb}) of crowding distance and distance from the ideal point for ranking. Using D_{ideal} during ranking of solutions not only helps in convergence along the center of the global PF but also along the center of those regions of PF which are induced by m objectives (local).

2.7 Conclusion

In this chapter, IDEMO, with revised elitist selection and ranking scheme, is used in order to improve the selection pressure, convergence and diversity of the solutions in the estimated PF. DECOR integrates IDEMO in a fast and online objective reduction framework with provision for elimination of multiple objectives in a turn. DECOR is applied on DTLZ problems for 10 and 20 objectives and results are noted in terms of convergence metric and hypervolume indicator. DECOR shows superior convergence to

PF as compared to several other algorithms. The diversity of the PF resulting from DECOR, is better than some of the MaOO approaches and is equivalent to a few other popular MaOO approaches. DECOR not only outperforms the recent objective reduction based MaOO approaches but also overcomes several of their drawbacks. In future, the integration of the revised elitist selection scheme and the objective reduction approach with other MOEAs could be explored.

An important observation is that there is a vast scope of improvement in terms of diversity as seen from the performance values of DECOR (Tables 2.3 and 2.5). This scope motivates research for further better many-objective evolutionary algorithms to tackle a broader spectrum of problem characteristics. Hence, in the next chapter, algorithms with decomposition-based strategies are considered for performance (convergence and diversity) improvement.

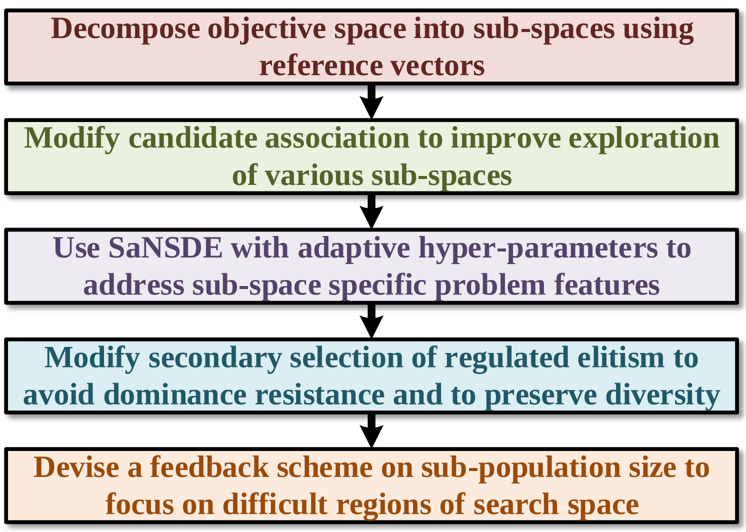
Chapter 3

ESOEA: Ensemble of Single Objective Evolutionary Algorithms for Many-Objective Optimization [138]

<u>**Outline**</u>

Objective: *To develop an adaptive optimization algorithm using reference-vector assisted decomposition of objective space and a feedback scheme on the allocation of candidates to the sub-spaces for addressing various kinds of many-objective optimization problems.*

Workflow:



3.1 Introduction

From the previous chapter, some Multi-Objective Evolutionary Algorithms (MOEAs) are noted to suffer from poor diversity. However, decomposition-based MOEAs are a promising alternative for Multi-Objective Optimization (MOO) or Many-Objective Optimization (MaOO) problems (Eq. (1.2)). These MOEAs use reference vectors to decompose the MOO/MaOO problems into multiple scalar problems which collaborate to get optimized.

Inspired by the success of decomposition-based MOEAs and the constant search for a versatile MaOO algorithm (adaptive to different problem characteristics), an optimization framework is developed by using an Ensemble of Single Objective Evolutionary Algorithms (ESOEA) [138]. It is characterized by reference-vector based decomposition and transformation of the MaOO problem into several single objective sub-problems to enhance the selection pressure. Additionally, with a feedback strategy, ESOEA explores difficult regions and thus, improves the search capability. For experimental validation, ESOEA is integrated with an adaptive Differential Evolution, and its performance is analyzed on several benchmark problems (from the DTLZ, WFG, IMB and CEC 2009 competition test suites) in terms of convergence metric, inverted generational distance, and hypervolume indicator. The estimated PFs are further visualized to establish the robustness of ESOEA.

Rest of this chapter is outlined as follows. The key concepts and the state-of-the-art of decomposition-based MOEAs are presented in Section 3.2. Thereafter, the algorithmic framework of ESOEA is described and discussed in Sections 3.3 and 3.4, respectively. Its performance is analyzed in Section 3.5 while highlighting its different modules. Finally, this chapter is concluded in Section 3.6, summarizing the overall observations.

3.2 Background of Reference Vector based Algorithms

Some basic concepts on reference vector based decomposition of objective space are presented in this section. Alongside, a brief review on the state-of-the-art of reference vector guided decomposition based algorithms is also presented.

3.2.1 Key Concepts

A decomposition-based MOEA partitions the objective space into multiple sub-spaces and thereby, decompose the MOO problem into multiple sub-problems (often single objective) associated with each sub-space. These sub-problems are then solved collaboratively. Hence, the decomposition of the objective space using reference vectors and the scalarization of the objective vectors are discussed next.

Reference Vector based Decomposition of the Objective Space

For partitioning the objective space into sub-spaces, an optimization algorithm is initialized with a set of reference vectors ($\mathcal{W}$) which is defined as follows:

$$\begin{gathered}\mathcal{W} = [\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_{n_{dir}}]^T, \\ \text{where } \mathbf{W}_i = [w_{i1}, w_{i2}, \cdots, w_{iM}] \text{ and } \sum_{j=1}^{M} w_{ij} = 1, \text{ for } i = 1 \text{ to } n_{dir}.\end{gathered} \tag{3.1}$$

The two-layered approach [40] is used to define $\mathcal{W}$ on a unit hyperplane in the first hyper-octant of the objective space. This approach is outlined using the following steps:

1. Das and Dennis' approach [40] is used to generate a set $\mathcal{H}_1$ of uniformly distributed $\binom{p_1+M-1}{M-1}$ vectors. If $M < 7$, $n_{dir} = \binom{p_1+M-1}{M-1}$ and $\mathcal{W} = \mathcal{H}_1$. An example of $\mathcal{W}$ with $M = 3$ and $p_1 = 4$ is illustrated in Fig. 3.1a and plotted in Fig. 3.1b.

   When $M = 7$, with $p_1 = 7$, $\mathcal{H}_1$ has $\binom{p_1+M-1}{M-1} = 1716$ reference vectors, which increases the computational burden of the MOEA. Also, with $p_1 < M$ only reference vectors on the boundary of the simplex are generated. Hence, when $M \geq 7$, a boundary layer is used to generate reference vectors (or the set $\mathcal{H}_1$) on the boundary of the simplex with $p_1 < M$ and another inside layer is used to generate reference vectors inside the simplex, as explained in the next step.

2. If $M \geq 7$, another set $\mathcal{H}_2$ of uniformly distributed $\binom{p_2+M-1}{M-1}$ vectors is generated. Each of the constituents of $\mathcal{H}_2$ is scaled and shifted to create a smaller simplex ($\mathcal{H}_2'$) inside the boundary simplex ($\mathcal{H}_1$). For each $h_{ij} \in \mathcal{H}_2$, $h'_{ij} = \frac{1-\tau_w}{M} + \tau_w \times h_{ij}$ is defined to create $\mathcal{H}_2'$ for $i = 1$ to $\binom{p_2+M-1}{M-1}$ and $j = 1$ to M. The parameter τ_w is called the shrinkage factor and $\tau_w = 0.5$ is considered, without loss of generality [109]. Then, $\mathcal{W} = \mathcal{H}_1 \cup \mathcal{H}_2'$ and $n_{dir} = \binom{p_1+M-1}{M-1} + \binom{p_2+M-1}{M-1}$. An example of $\mathcal{W}$ with $M = 3$, $p_1 = 2$ and $p_2 = 1$ is illustrated in Fig. 3.1c.

Once the set of reference vector $\mathcal{W}$ is defined, each reference vector $\mathbf{W}$ defines a unique sub-space in the objective space. To associate a point $\mathbf{F}(\mathbf{X})$ to a subspace, $d2$ is used which

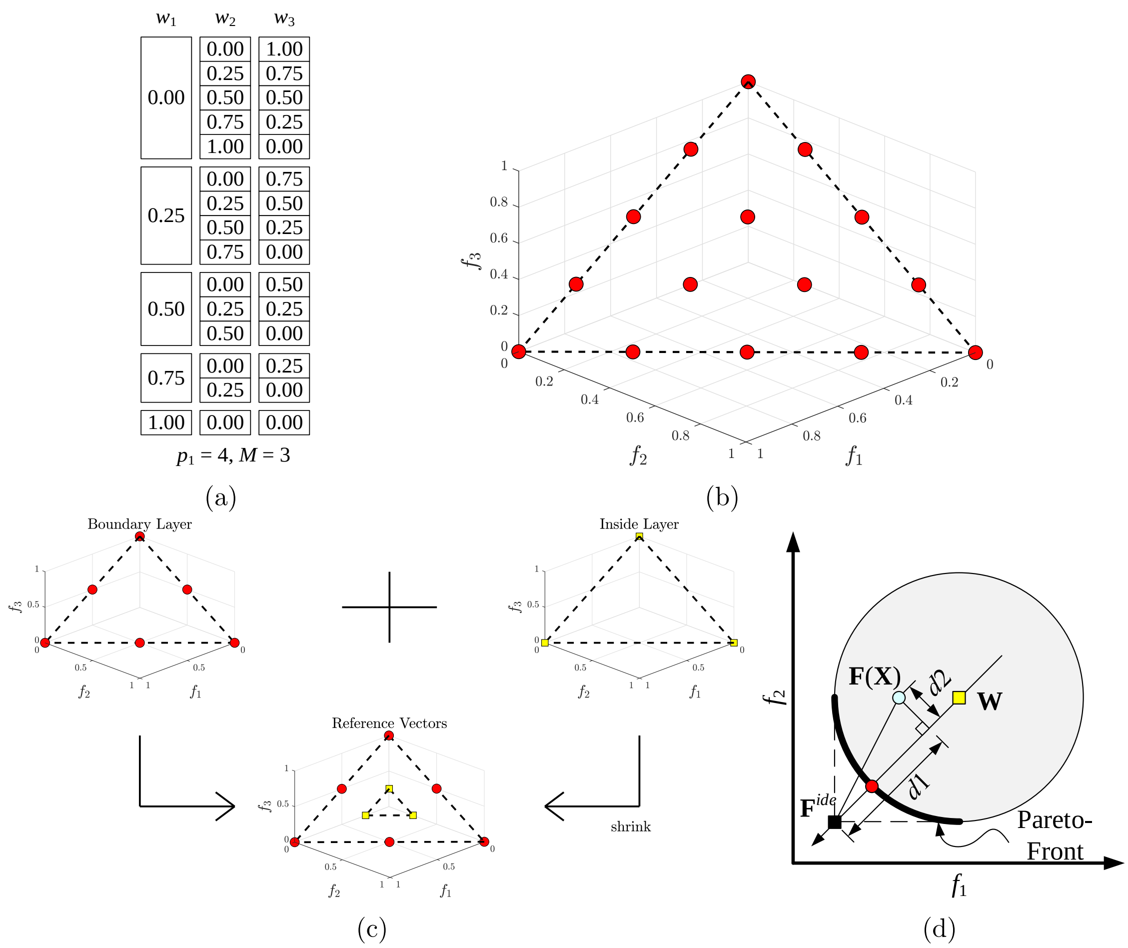


Figure 3.1: (a) One-layered approach of generating the reference vector set $\mathcal{W}$ with $p_1 = 4$ partitions in three-dimensional space using Das and Dennis' approach [40], (b) visualization of these $\binom{p_1+M-1}{M-1} = 15$ points (sampled from unit simplex), (c) two-layered approach of generating $\mathcal{W}$ with $p_1 = 2$ partitions in boundary layer and $p_2 = 1$ partitions in the inside layer, (d) PBI function combines $d1$ and $d2$ while associating $\mathbf{F}(\mathbf{X})$ with the reference vector $\mathbf{W}$ and aims to bring $\mathbf{F}(\mathbf{X})$ at the intersection of PF and $\mathbf{W}$, i.e., at $\bullet$ point.

is the perpendicular distance from $\mathbf{F}(\mathbf{X})$ to the reference vector passing through $\mathbf{W}_i$ and origin of the objective space. An illustration of this distance $d2$ is provided in Fig. 3.1d. Hence, a subspace is formed by all objective vectors $\mathbf{F}(\mathbf{X})$ that are associated to a reference vector $\mathbf{W}_i$ as follows:

$$\begin{aligned} &\text{Sub-space associated to } \mathbf{W}_i : \{\mathbf{F}(\mathbf{X}) \in \mathbb{R}^M | d2\,(\mathbf{X}|\mathbf{W}_i) \leq d2\,(\mathbf{X}|\mathbf{W}_j)\}, \\ &\quad \text{for } j = \{1, 2, \cdots, n_{dir}\},\ i \neq j \text{ and } \mathbf{X} \in \mathcal{D}. \end{aligned} \tag{3.2}$$

Scalarization Methods

The scalarization approaches obtain a scalar fitness value for an objective vector of the MOO problem [32, 85, 197]. The most commonly used scalarization functions [28, 109] are weighted sum, Tchebycheff and boundary intersection functions. Normal Boundary Intersection (NBI) method [40] considers a scalar transformation with an equality constraint. The unconstrained NBI variant uses a penalty parameter (θ_{pbi}) and is known as Penalty-based Boundary Intersection (PBI) method [40,163]. Due to its efficacy [40,109], several MOEAs [109, 138] adopt the PBI function. Using $\mathbf{F}^{ide}$ from Eq. (1.8), for a reference vector ($\mathbf{W}$), PBI generates a scalar optimization problem as follows:

$$\begin{aligned} &\text{Minimize: } f_{pbi}(\mathbf{X}|\mathbf{W},\mathbf{F}_{ide}) = d1 + \theta_{pbi} \times d2 \text{ where, } \mathbf{X} \in \mathcal{D},\ \theta_{pbi} \geq 0, \\ &d1 = \frac{\left\| \left(\mathbf{F}\left(\mathbf{X}\right) - \mathbf{F}^{ide}\right)^T \mathbf{W} \right\|}{\|\mathbf{W}\|} \text{ and } d2 = \left\| \mathbf{F}\left(\mathbf{X}\right) - \left(\mathbf{F}^{ide} + d1 \frac{\mathbf{W}}{\|\mathbf{W}\|}\right) \right\|. \end{aligned} \tag{3.3}$$

In PBI function (Eq. (3.3)), $d1$ denotes the convergence of the projection of $\mathbf{F}\left(\mathbf{X}\right)$ on $\mathbf{W}$ and $d2$ denotes the perpendicular distance from $\mathbf{F}\left(\mathbf{X}\right)$ to $\mathbf{W}$, as shown in Fig. 3.1d. Hence, $d2$ is the diversity parameter. Also, θ_{pbi} balances the degree of convergence and diversity such that $\mathbf{F}(\mathbf{X})$ evolves along the boundary of f_{pbi} for a given $\mathbf{W}$.

In [109], contours of PBI function along $\mathbf{W} = [0.5, 0.5]$ for $\theta_{pbi} = \{0, 1, 2\}$ are analyzed. It shows when $\theta_{pbi} = 0$, the PBI function (Eq. (3.3)) represents weighted sum and when $\theta_{pbi} = 1$, the PBI function (Eq. (3.3)) represents weighted Tchebycheff. The contour for weighted Tchebycheff with $\mathbf{W} = [0.5, 0.5]$ has same shape as the region dominating the solution $\mathbf{X}$ (green region in Fig. 1.1a). This contour becomes narrower around $\mathbf{W}$ as θ_{pbi} increases [109,163]. MOEAs with $\theta_{pbi} = 5$ shows promising performance [109,138,150,160].

3.2.2 Related Works

A reference vector assisted decomposition based MOEA is advantageous as it neither faces dominance resistance in high dimensional objective space like Pareto-dominance based algorithms nor it requires the extreme computational effort for hypervolume evaluation. However, such MOEAs suffer from the following shortcomings:

1. Replacing old solutions by new solutions is dictated by scalarization functions [123, 150] and may skip some search regions leading to a severe loss of population diversity.
2. Setting the reference vectors and the scalarization function is problem-specific [123, 150], which can be a daunting task for every new kind of a problem.
3. Performance of these MOEAs strongly depends on whether the distribution of reference vectors is consistent with the shape of Pareto-Front (PF) [84].

Hence, these MOEAs cannot regulate their exploration as per problem requirements. Some adaptation methods [151,190] report results only for 2 or 3-objective problems. Even the adaptive MaOO framework of [63] reports its performance up to 8-objective problems. Hence, their extensibility for problems with a large number of objectives is unknown.

Standard reproduction operators of Genetic Algorithm [44, 167] or Differential Evolution [15, 42] are essentially designed for candidate-wise perturbation, i.e., for single-objective EAs. In contrast, the solution of MOO problems is characterized by a set of candidate solutions. Yet most of the existing MOEAs directly adopt these vector-wise (instead of set-wise) reproduction operators, without proper justification [190]. Moreover, the existing MOEAs are specifically tailored for particular types (difficulties and shape) of PF [84, 173]. Hence, the literature of MOEAs still lacks a robust algorithm with adaptive search-ability.

Recent literature presents some reference-vector based adaptive MOEAs like Adaptive-NSGA-III (or A-NSGA-III) [89], MOEA/D-M2M [115, 117], MOEA based on Dominance and Decomposition (MOEA/DD) [109], Reference Vector guided Evolutionary Algorithm (RVEA) [28] and Adaptive Reference-vector based MOEA (AR-MOEA) [173]. The reference vectors ($\mathcal{W}$) represent different aspects in different algorithms. For example, in MOEA/D-M2M and RVEA, $\mathcal{W}$ specifies the sub-populations; in MOEA/DD and NSGA-III, $\mathcal{W}$ estimates the local density; while in AR-MOEA, $\mathcal{W}$ evaluates a scalar indicator to address the shape of the PF.

For addressing different shapes of PF, some reference vector adaptation strategies modify $\mathcal{W}$ based on the solution distribution in the current population such as in A-NSGA-III [89] and RVEA* [28] whereas other strategies are based on the solution distribution in an external archive such as in $pa\lambda$-MOEA/D [92] and MOEA/D-AWA [148]. The general steps of reference vector adaptation involve deletion of reference vectors from an empty niche or a sparsely populated region, followed by the addition of reference vectors randomly or to a densely populated region. However, these reference vector adaptation strategies perform better for MOO/MaOO problems with irregular PF[1] than those with regular PF due to perturbation of initial uniform distribution of the reference vectors [173].

Although the contemporary algorithms show excellent performance, yet these algorithms suffer from the following issues:

[1]Continuous, smooth and well-spread PF are called as regular PF and degenerate, disconnected, inverted PF or PF with sharp tails are called as irregular PF [173].

1. As MOEA/D-M2M is designed to deal with imbalance and variable linkage difficulties [115], its extension for problems with irregular PF is yet to be investigated.

2. Some algorithms (like MOEA/D-M2M [117], NSGA-III [45], and AR-MOEA [173]) inherit the entire rank-one solutions, which exhibits dominance resistance for problems with $M \geq 10$ [79]. Hence, such MOEAs are prone to get trapped in local optima as high-ranked diverse solutions have minimal survival chances [10].

3. For adaptive algorithms, like A-NSGA-III [45] and RVEA* [28], the candidate association with reference vectors becomes costlier as the location of the reference vectors are not constant, i.e., the reference vectors are adapted as per the shape of PF.

Thus, the need of the hour is an evolutionary algorithm which exhibits robust performance for various problem types and reduces the effect of the above drawbacks.

3.3 Algorithmic Framework of ESOEA

Motivated by the requirements specified in Section 3.2.2, an adaptive framework is developed for dealing with MaOO problems using an Ensemble of Single Objective Evolutionary Algorithms (ESOEA) [138]. Its significant contributions are as follows:

1. Being a decomposition-based approach, ESOEA considers multiple single objective sub-problems to maintain the selection pressure. The use of a Single Objective Evolutionary Algorithm (SOEA) justifies the use of candidate-wise reproduction operators. Moreoever, an adaptive SOEA helps in addressing the sub-space specific problem characteristics.

2. ESOEA uses a regulated elitism scheme (with a novel secondary selection) where only a fraction of rank-one solutions is inherited to avoid dominance resistance for MaOO problems with $M \geq 10$. As the regulated elitist selection is executed only after periodic intervals, it leads to a considerable saving in execution time.

3. ESOEA adaptively allocates candidates to SOEAs based on the contribution to form the current population. Thus, it boosts those SOEAs which perform poorly.

ESOEA [138] is the first MaOO algorithm that focusses on the above aspects.

Rest of this section describes several modules involved in building the entire framework of ESOEA for addressing MOO and MaOO problems using multiple instances of SOEAs, as illustrated in Fig. 3.2. Its constituent units are subsequently discussed.

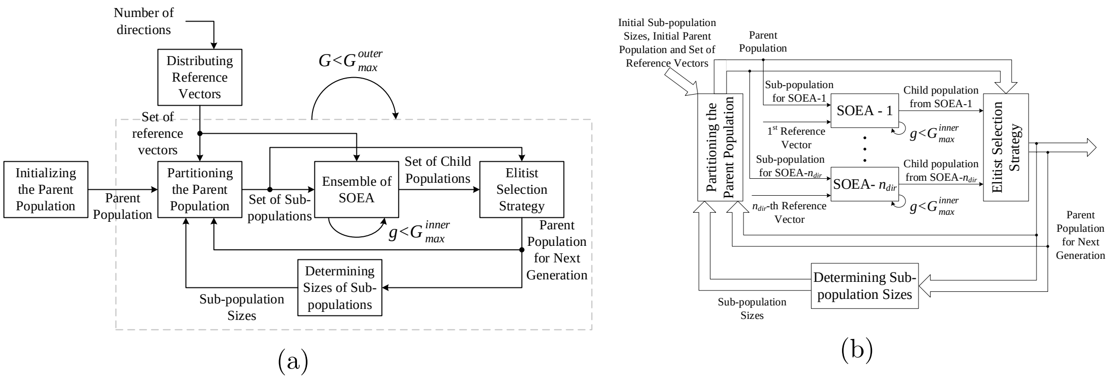


Figure 3.2: Architecture of ESOEA: (a) Overall framework [138], (b) Central loop [138].

3.3.1 Distributing Reference Vectors

The first step of ESOEA involves initializing a set of n_{dir} reference vectors ($\mathcal{W}$) using Das and Dennis' approach [40], as described in Section 3.2.1. This set of reference vectors is stored as a matrix ($\mathcal{W}$) of size $n_{dir} \times M$ as shown in Eq. (3.1).

3.3.2 Initializing the Parent Population

A population of n_{pop} candidate solutions is randomly initialized within the bounds of the search space ($\mathcal{D}$), as explained in Eq. (2.1), where $rand(0, 1)$ indicates a random real number between 0 and 1. This initial parent population is denoted as $\mathcal{A}_{G=0}$.

3.3.3 Partitioning the Parent Population (Decomposition)

Prior to partitioning the parent population ($\mathcal{A}_G$), objective scaling is done so that the scales of the objective functions have minimal influence on the other modules. For this purpose, a reference vector ($\mathbf{F}_G^r$) is obtained from the parent population as follows:

$$\mathbf{F}_G^r = \left[f_{1,G}^r, \cdots, f_{M,G}^r\right] \text{ with } f_{j,G}^r = \max_{\mathbf{X} \in \mathcal{A}_G} f_j(\mathbf{X}) - \min_{\mathbf{X} \in \mathcal{A}_G} f_j(\mathbf{X}) \text{ for } j = 1, \cdots, M. \quad (3.4)$$

This reference vector is updated in the G^{th} generation of the central loop with updates in the parent population $\mathcal{A}_G$. Using $\mathbf{F}_G^r$, an objective vector $\mathbf{F}(\mathbf{X})$ is scaled as follows:

$$\mathbf{F}^s(\mathbf{X}) = \left[\frac{f_1(\mathbf{X}) - \underset{\mathbf{X} \in \mathcal{A}_G}{min} f_M(\mathbf{X})}{f^r_{1,G}}, \cdots, \frac{f_M(\mathbf{X}) - \underset{\mathbf{X} \in \mathcal{A}_G}{min} f_M(\mathbf{X})}{f^r_{M,G}} \right], \text{ where } \mathbf{X} \in \mathcal{A}_G. \quad (3.5)$$

Population decomposition aims to select the candidates of the k^{th} sub-population ($\mathcal{A}^{sub}_{k,G}$) such that their corresponding scaled objective vectors ($\mathbf{F}^s(\mathbf{X})$ where $\mathbf{X} \in \mathcal{A}^{sub}_{k,G}$) are associated to the k^{th} reference vector ($\mathbf{W}_k$) among the n_{dir} reference vectors.

- *Usual Approach:* Usually, each candidate is associated to its nearest reference vector using the smallest $d2$ distance [45, 109, 160] (as shown in Fig. 3.1d and Eq. (3.2)) or using the smallest acute angle [28, 117] between $\mathbf{F}(\mathbf{X})$ and $\mathbf{W}_k$.

- *Association Approach of ESOEA:* This population decomposition is described in Algorithm 3.1. The first step (line 3) is to consider the array $\mathbf{P}^{arr}_G = \left[S^1_G, \cdots, S^{n_{dir}}_G\right]$ which stores the sub-population sizes. Using the $round(.)$ function on a real number to yield the nearest integer, the size of the k^{th} sub-population ($S^k_{G=0}$) is initialized as follows:

$$S^k_{G=0} = \begin{cases} round(n_{pop}/n_{dir}), & \text{for } k = 1, 2, \cdots, (n_{dir}-1) \text{ and} \\ n_{pop} - \sum_{i=1}^{(n_{dir}-1)} S^i_{G=0}, & \text{when } k = n_{dir}. \end{cases} \quad (3.6)$$

  This initialization considers that all n_{dir} sub-spaces are equally likely to be searched. When $G \neq 0$, the determination of S^k_G is discussed later in Section 3.3.6. In lines 4 to 9, $\mathcal{A}_G$ is sorted using the angle ϕ^E_{ik} between the scaled objective vector of $\mathbf{X}_i$ and the k^{th} reference vector ($\mathbf{W}_k$). This associating angle (ϕ^E_{ik}) is obtained as follows:

$$\phi^E_{ik} = \phi^E\left(\mathbf{F}^s(\mathbf{X}_i), \mathbf{W}_k\right) = \arccos\left(\frac{\mathbf{F}^s(\mathbf{X}_i) \cdot \mathbf{W}_k}{\|\mathbf{F}^s(\mathbf{X}_i)\| \times \|\mathbf{W}_k\|}\right), \text{ where } \mathbf{X}_i \in \mathcal{A}_G. \quad (3.7)$$

  Finally, the k^{th} sub-population $\mathcal{A}^{sub}_{k,G}$ is formed in lines 10 to 11 by sequentially assigning S^k_G candidates from sorted $\mathcal{A}_G$ to $\mathcal{A}^{sub}_{k,G}$.

After population decomposition, the sub-population $\mathcal{A}^{sub}_{k,G}$ is used as the initial population $\mathcal{A}^{par}_{k,g=0}$ of the k^{th} SOEA of the ensemble of single-objective optimizers.

Algorithm 3.1 Procedure for Partitioning the Population in ESOEA [138]

Input: $\mathbf{F}^s(\mathbf{X})$: Scaled objective vectors $\forall \mathbf{X} \in \mathcal{A}_G$; $\mathbf{P}_G^{arr}$: Array of sub-population sizes; n_{dir}: Number of reference vectors; $\mathcal{W}$: Set of reference vectors

Output: $\{\mathcal{A}_{1,G}^{sub}, \cdots, \mathcal{A}_{n_{dir},G}^{sub}\}$: n_{dir} sub-populations

procedure CANDASSOC($\mathcal{W}$, $\forall \mathbf{X} \in \mathcal{A}_G : \mathbf{F}^s(\mathbf{X})$, $\mathbf{P}_G^{arr}$, n_{dir})
    for $k = 1$ to n_{dir} (for each sub-population) **do**
        $S_G^k \leftarrow$ Size of k^{th} sub-population from $\mathbf{P}_G^{arr}$
        $\mathbf{\Phi}_k^E \leftarrow \emptyset$
        for $i = 1$ to n_{pop} (for each candidate) **do**
            Evaluate ϕ_{ik}^E using Eq. (3.7) where $\mathbf{W}_k \in \mathcal{W}$
            $\mathbf{\Phi}_k^E \leftarrow \mathbf{\Phi}_k^E \cup \phi_{ik}^E$
        end for
        $\mathbf{I}_{sort} \leftarrow$ Indices after sorting $\mathbf{\Phi}_k^E$ in ascending order
        $\mathbf{I}_{sub} \leftarrow$ First S_G^k indices from $\mathbf{I}_{sort}$
        $\mathcal{A}_{k,G}^{sub} \leftarrow \{\mathbf{X}_i \in \mathcal{A}_G | i \in \mathbf{I}_{sub}\}$
    end for
    return $\{\mathcal{A}_{1,G}^{sub}, \cdots, \mathcal{A}_{n_{dir},G}^{sub}\}$
end procedure

3.3.4 Ensemble of Single Objective Optimizers

In this module, n_{dir} instances of SOEA are used to evolve each sub-population for G_{max}^{inner} number of generations as shown in Fig. 3.2b. For differentiating, g denotes a generation of SOEA and G denotes a generation of ESOEA. The sub-population $\mathcal{A}_{k,g}^{par}$ is evolved to create a child sub-population $\mathcal{A}_{k,g}^{child}$. The sub-populations $\mathcal{A}_{k,g}^{par}$ and $\mathcal{A}_{k,g}^{child}$ are compared to yield $\mathcal{A}_{k,g+1}^{par}$ using PBI function (Eq. (3.3)) for scalarization of the MaOO problem.

After G_{max}^{inner} generations, the k^{th} SOEA generates the child sub-population $\mathcal{A}_{k,g=G_{max}^{inner}}^{child} = \mathcal{A}_{k,G}^{child}$. All candidates $\mathcal{A}_{k,G}^{child}$ undergo objective scaling by Eq. (3.5) using the same $\mathbf{F}_G^r$ which was determined over $\mathcal{A}_G$ in Eq. (3.4). Then, elitist selection is performed on the merged population $\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}_{k,G}^{child}\right)\right)$. As all objective vectors remain scaled using the same $\mathbf{F}_G^r$, the elitist selection is unbiased and yields the next population ($\mathcal{A}_{G+1}$).

3.3.5 Elitist Selection Strategy

Elitist selection is considered as elitism is necessary for MOEAs to guarantee convergence [155]. The usual approach of elitism (non-dominated sorting followed by sorting using crowding distance) and its problems are outlined in Section 2.3.2.

Using $ndset(.)$ from Eq. (1.4), ESOEA performs non-dominated sorting [47] on the merged population $\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}_{k,G}^{child}\right)\right)$ to yield $\mathcal{R}_{1,G}$ and $\mathcal{R}_{rest,G} = \bigcup_{p\geq 2} \mathcal{R}_{p,G}$ as follows:

$$\mathcal{R}_{p,G} = \begin{cases} ndset\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}_{k,G}^{child}\right)\right), \text{ for } p = 1 \text{ and} \\ ndset\left(\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}_{k,G}^{child}\right)\right) \setminus \left(\cup_{q=1}^{p-1} \mathcal{R}_{q,G}\right)\right), \text{ for } p \geq 2. \end{cases} \tag{3.8}$$

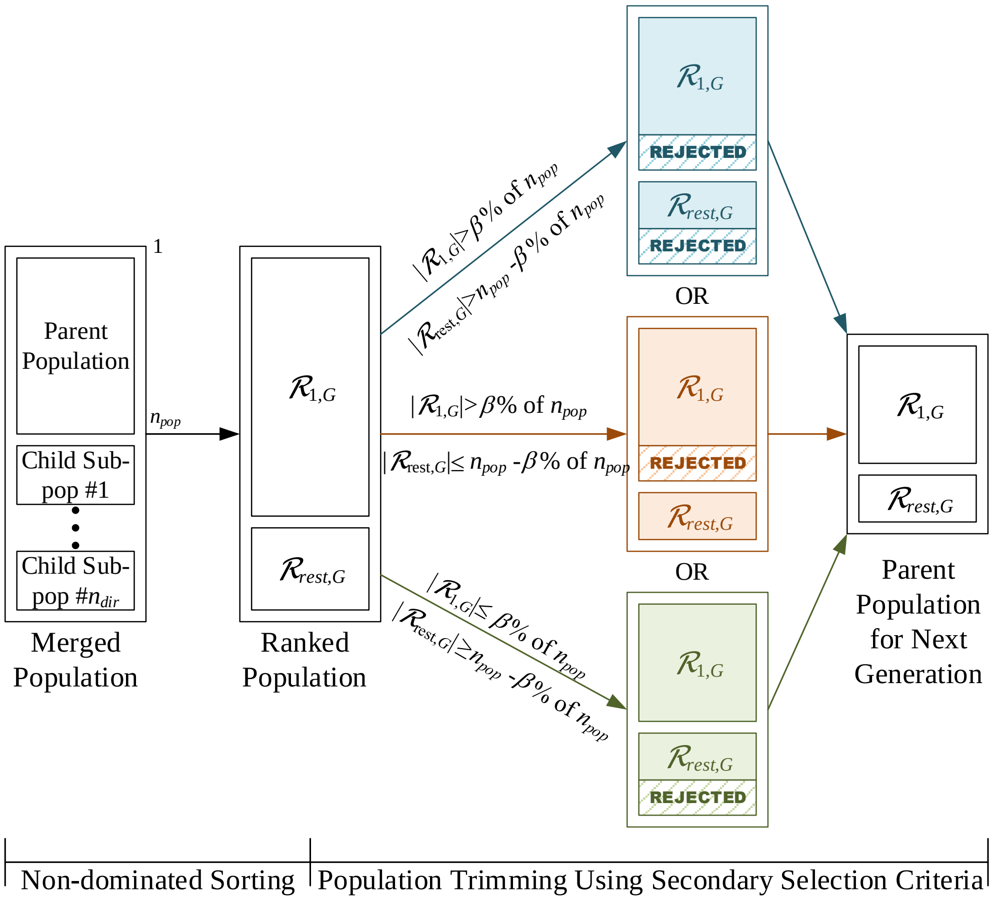


Figure 3.3: Regulated elitism strategy adopted in ESOEA [138].

Based on the parameter β (regulating the number of candidates of $|\mathcal{R}_{1,G}|$ and $|\mathcal{R}_{rest,G}|$ that are allowed to propagate), one of the following approaches (Fig. 3.3) is followed:

- *When* $|\mathcal{R}_{1,G}| > round\,(\beta\%\ of\ n_{pop})$ *and* $|\mathcal{R}_{rest,G}| > (n_{pop} - round\,(\beta\%\ of\ n_{pop}))$:

  $\mathcal{R}_{1,G}$ is sorted based on the secondary selection criteria and up to $round\,(\beta\%$ of $n_{pop})$ number of candidates are allowed to fill $\mathcal{A}_{G+1}$.

  The remaining of $\mathcal{A}_{G+1}$ is filled set-wise. Starting from $\mathcal{R}_{2,G}$, those ranks of solutions (which can be entirely accommodated) are added to $\mathcal{A}_{G+1}$ until the last required set $\mathcal{R}_{q,G}$ is reached from which only a fraction of solutions are added. The undesired candidates are eliminated from $\mathcal{R}_{q,G}$ using the secondary selection criteria.

- *When* $|\mathcal{R}_{1,G}| > round\,(\beta\%\ of\ n_{pop})$ *and* $|\mathcal{R}_{rest,G}| \leq (n_{pop} - round\,(\beta\%\ of\ n_{pop}))$:

The entire of $\mathcal{R}_{rest,G}$ is propagated to $\mathcal{A}_{G+1}$. From $\mathcal{R}_{1,G}$, $(n_{pop} - |\mathcal{R}_{rest,G}|)$ number of candidates are allowed to be propagated. The undesired candidates are elimated from $\mathcal{R}_{1,G}$ using the secondary selection criteria.

- *When* $|\mathcal{R}_{1,G}| \leq round\,(\beta\%\ of\, n_{pop})$ *and* $|\mathcal{R}_{rest,G}| \geq (n_{pop} - round\,(\beta\%\ of\, n_{pop}))$:

  The usual approach of elitist selection [47] is followed as described in Fig. 2.1b. However, instead of crowding distance, the undesired candidates are eliminated from $\mathcal{R}_{q,G}$ using the secondary selection criteria of ESOEA.

- *When* $|\mathcal{R}_{1,G}| < round\,(\beta\%\ of\, n_{pop})$ *and* $|\mathcal{R}_{rest,G}| < (n_{pop} - round\,(\beta\%\ of\, n_{pop}))$:

  This case is not possible as $|\mathcal{R}_{1,G}| + |\mathcal{R}_{rest,G}| \geq n_{pop}$.

Secondary Selection of ESOEA: Using Algorithm 3.2, a rank of solutions ($\mathcal{R}_{q,G}$) is sorted to yield $\mathcal{A}^{sort}$ and the required number of candidates are sequentially selected from $\mathcal{A}^{sort}$ to preserve diversity. At first, $d2$ is evaluated for all the candidates of $\mathcal{R}_{q,G}$ (lines 2 to 6). Then, the outer while loop (lines 8 to 18) selects candidate in the various unique directions as follows. An auxiliary array $\mathbf{P}^{labels}$ is used to store the indices of those reference vectors with which each candidate of $\mathcal{R}_{q,G}$ is associated. This association is established either through population partitioning or being generated as a child from a particular SOEA. For each unique reference vector ($\mathbf{W}_{r_{vec}}$), the associated sub-population within $\mathcal{R}_{q,G}$ (denoted as $\mathcal{A}^{r_{vec}}$) is selected and thereafter, the candidate ($\mathbf{X}_{select}$) with minimum $d2$ is chosen from $\mathcal{A}^{r_{vec}}$ in lines 11 to 13. This $\mathbf{X}_{select}$ is appended to $\mathcal{A}^{sort}$ and is removed from $\mathcal{R}_{q,G}$ (lines 14 to 16). The selection (lines 10 to 17) continues until a candidate is chosen from every unique direction, after which the outer while loop resumes with the reduced set of candidates in $\mathcal{R}_{q,G}$ and continues until all candidates of $\mathcal{R}_{q,G}$ have been assigned to $\mathcal{A}^{sort}$.

Thus, by elitist selection from the merged populations, i.e., from $\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}_{k,G}^{child}\right)\right)$, the parent population for the next generation ($\mathcal{A}_{G+1}$) is extracted.

3.3.6 Determining Sub-population Sizes for Next Generation

The adaptive property of ESOEA is imparted by regulating candidate allocation instead of perturbing the uniformly distributed reference vectors. The feedback (Fig. 3.2) provided to the population decomposition step is the updated sub-population sizes for next generation, i.e., $\mathbf{P}_{G+1}^{arr} = \left[S_{G+1}^{1}, S_{G+1}^{2}, \cdots, S_{G+1}^{n_{dir}}\right]$. This updation has the following characteristics:

Algorithm 3.2 Sorting based on Secondary Selection Criteria [138]

Input: $\mathcal{R}_{q,G}$: Candidates forming q^{th} non-dominated rank; $\mathbf{F}^s(\mathbf{X})$: Scaled objective vectors of candidates in the q^{th} non-dominated rank $\forall \mathbf{X} \in \mathcal{R}_{q,G}$; $\mathbf{P}^{labels}$: Indices of reference vectors with which candidates of $R_{q,G}$ are associated
Output: $\mathcal{A}^{sort}$: Candidates of $\mathcal{R}_{q,G}$ in sorted order

procedure SECONDSORT($\mathcal{R}_{q,G}$, $\forall \mathbf{X} \in \mathcal{R}_{q,G} : \mathbf{F}^s(\mathbf{X})$, $\mathbf{P}^{labels}$)
  $\mathbf{D}_{d2} \leftarrow \emptyset$
  for $\mathbf{X} \in \mathbf{R}_{q,G}$ (for each candidate) **do**
    $d2_q \leftarrow d2(\mathbf{X}|\mathbf{W}_k)$ using Eq. (3.3) with $\mathbf{F}^s(\mathbf{X})$ and $k = \mathbf{P}^{labels}(\mathbf{X})$
    $\mathbf{D}_{d2} \leftarrow \mathbf{D}_{d2} \cup d2_q$
  end for
  $n_{cand} = 1$, $\mathcal{A}^{sort} = \emptyset$
  while $n_{cand} \leq |\mathcal{R}_{q,G}|$ **do**
    $\mathbf{I}_{uniq}$ = Obtain all unique directions from $\mathbf{P}^{labels}$
    for $r_{vec} \in \mathbf{I}_{uniq}$ (for each direction) **do**
      $\mathcal{A}^{r_{vec}} = \{\mathbf{X} \in \mathcal{R}_{q,G} | \mathbf{P}^{labels}(\mathbf{X}) = r_{vec}\}$
      $\mathbf{D}_{d2}^{r_{vec}} = \{d2 \in \mathbf{D}_{d2} | \mathbf{P}^{labels}(\mathbf{X}) = r_{vec}\}$
      $\mathbf{X}_{select} = \arg\min_{\mathbf{X} \in \mathcal{A}^{r_{vec}}} \mathbf{D}_{d2}^{r_{vec}}$
      $\mathcal{A}^{sort} = \mathcal{A}^{sort} \cup \mathbf{X}_{select}$ (append sequentially)
      $n_{cand} = n_{cand} + 1$
      $\mathcal{R}_{q,G} = \mathcal{R}_{q,G} \setminus \mathbf{X}_{select}$
    end for
  end while
  return $\mathcal{A}^{sort}$
end procedure

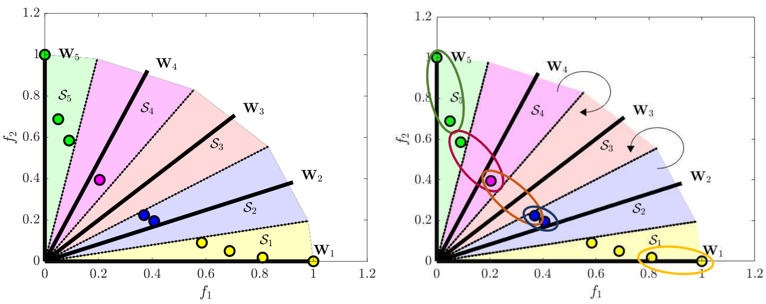
- It allocates more candidates to the poorly performing instances of SOEA (or equivalently along those sub-spaces where exploration is challenging).
- As a performance indicator, N_k^{share} represents the percentage of candidates contributed by the k^{th} instance of SOEA (SOEA-k) towards $\mathcal{A}_{G+1}$ as mentioned in Eq. (3.9), where n_k^{child} is the number of candidates contributed by $\mathcal{A}_{k,G}^{child}$ and n_{total}^{child} is the total number of candidates contributed by all the child sub-populations.
- As mentioned before, when $G = 0$, the sub-population sizes are initialized using Eq. (3.6). For subsequent generations, the k^{th} sub-population size (S_{G+1}^k) is negatively correlated to N_k^{share} as follows:

$$S_{G+1}^k = \begin{cases} round\left(\left(\frac{100 - N_k^{share}}{n_{dir}-1}\right) \times \left(\frac{n_{pop}}{100}\right)\right), & \text{if } k \neq n_{dir} \text{ and} \\ n_{pop} - \sum_{k=1}^{n_{dir}-1} S_{G+1}^k, & \text{if } k = n_{dir}, \end{cases}$$
$$\text{where, } N_k^{share} = (n_k^{child} / n_{total}^{child}) \times 100 \text{ with}$$
$$n_k^{child} = \left|\mathcal{A}_{k,G}^{child} \cap \mathcal{A}_{G+1}\right| \text{ and } n_{total}^{child} = \sum_{k=1}^{n_{dir}} n_k^{child}. \tag{3.9}$$

Table 3.1: Example to illustrate adaptive feedback of sub-population sizes [138].

k^{th} sub-population	Contribution of k^{th} sub-population, n_k^{child}	Performance indicator of k^{th} sub-population, N_k^{share}	Sub-population size before rounding, S_{G+1}^k	Sub-population size after rounding, S_{G+1}^k
1	10	20%	10.00	10
2	0	0%	12.50	13
3	25	50%	6.25	6
4	10	20%	10.00	10
5	5	10%	11.25	11
Total	**50**	**100%**	**50**	**50**

An example in Table 3.1 is discussed to explain the adaptive feedback. In Table 3.1, the sub-population along the 2[nd] reference vector does not contribute anything whereas the sub-population along the 3[rd] reference vector contributes half of the candidates to the parent population for the next generation. Thus, more resource should be spent to explore the region around the 2[nd] reference vector than that around the 3[rd] reference vector. This feedback is provided in terms of sub-population size for the next generation. Thus, by Eq. (3.9), the candidate allocation to the k^{th} sub-space is boosted (or damped) if it has a lesser (or higher) contribution in forming $\mathcal{A}_{G+1}$ as compared to other sub-spaces.



(a) Usual Strategy

(b) Strategy of ESOEA

k^{th} Sub-space ($\mathcal{S}_k$)	Existing S_G^k	ESOEA S_G^k
1	4	2
2	2	2
3	0	2
4	1	2
5	3	2
Total (n_{pop})	10	10

(c) Specifications

Figure 3.4: Difference in sub-population formation approaches: (a) usual strategy, (b) strategy of ESOEA, (c) corresponding details of population partitioning.

To explain why adaptively fixing S_G^k is a better strategy [26] than existing association strategy [28,45,109], the example in Fig. 3.4 is considered. The sub-populations formed by usual association scheme are shown in Fig. 3.4a, those formed by the association scheme of ESOEA are shown in Fig. 3.4b and the respective sub-population sizes are mentioned in Fig. 3.4c. The usual approach and the approach of ESOEA are compared as follows:

- In constrast to the usual association scheme (Fig. 3.4a), in the adaptive association scheme of ESOEA (Fig. 3.4b) S_G^k are determined before partitioning the population into sub-populations (Fig. 3.4c). Hence, unlike the usual approach, some solutions (such as a few from $\mathcal{S}_1$ in Fig. 3.4b) do not contribute to any sub-population.

- Each sub-population represents the mating pool corresponding to a sub-space. In the usual approach, all the candidates within a sub-space formed the mating pool. This approach is unfair as sub-spaces ($\mathcal{S}_3$ and $\mathcal{S}_4$) with a lower solution density need relatively more exploration. Some sub-spaces (such as $\mathcal{S}_1$ and $\mathcal{S}_5$) have a higher densities of solutions and thereby, result in bigger mating pools. To remove this bias, the approach of ESOEA adaptively determines S_G^k and tends to allocate an equal number of candidates in exploring all the sub-spaces (such as in Table 3.1).

- When a sub-space is empty, ESOEA borrows solutions from the neighboring sub-spaces similar to MOEA/DD [109]. For example, $\mathcal{A}_{3,G}^{sub}$ is formed using solutions from $\mathcal{S}_2$ and $\mathcal{S}_4$ in Fig. 3.4b. This is beneficial to improve exploration.

Thus, this adaptive association scheme of ESOEA assists in better exploration and in improving the overall diversity of the estimated PF.

3.4 Comparison of ESOEA with Related Algorithms

The similarities and differences of ESOEA with a few related works such as MOEA/D-M2M [117], MOEA/DD [109], RVEA [28] and AR-MOEA [173] are highlighted next.

3.4.1 Similarities with Related Algorithms

The following similarities are observed between ESOEA and other related algorithms:

a) Similar to MOEA/D-M2M and RVEA, ESOEA uses acute angle (Eq. (3.7)) between objective vectors and reference vectors for the association of candidates.

b) Similar to MOEA/DD, ESOEA uses PBI as the scalarization function.

c) Similar to MOEA/D-M2M, MOEA/DD, and AR-MOEA, ESOEA employs non-dominated sorting. Although RVEA is not based on Pareto-dominance, both RVEA and ESOEA performs an elitist selection of candidates.

d) Similar to MOEA/DD and MOEA/D-M2M, ESOEA exploits the neighborhood property of MaOO problems during mating of candidates.

3.4.2 Differences with Related Algorithms

The following differences are observed between ESOEA and other related algorithms:

a) *With MOEA/D-M2M:* MOEA/D-M2M [117] considers a fixed sub-population size. If a sub-population size exceeds this fixed size, the usual non-dominated sorting based selection [47, 117] is performed. Otherwise, candidates are randomly selected from the rest of the population to fill up the respective sub-population. Unlike this, ESOEA uses adaptive sub-population size and preserves the neighborhood property during the formation of parent sub-populations.

b) *With MOEA/DD:* In MOEA/DD [109], mating occurs either between neighboring sub-populations or within the global population, whereas in ESOEA, mating occurs only within each sub-population. Moreover, MOEA/DD (steady-state selection) preserves isolated solutions even with worst scalarized fitness in the last domination level. ESOEA performs the same by propagating the solutions closest to respective reference vectors, within a non-dominated rank of solutions.

c) *With RVEA:* In RVEA [28], mating occurs within the global population and candidate selection is guided by Angle-Penalized Distance (APD). Moreover, RVEA modifies initial reference vector distribution for scaled MaOO problems and for problems with irregular PF. Unlike this, ESOEA performs mating within each sub-population, performs selection guided by Pareto-dominance and PBI, and does not perturb the initial distribution of reference vectors.

d) *With AR-MOEA:* In AR-MOEA [173], non-dominated sorting is performed followed by secondary selection using IGD with Non-contributing Solutions (IGD-NS). Also, currently, AR-MOEA cannot handle problems with difficult regions [173]. Unlike this, ESOEA performs regulated elitism based on Pareto-dominance and PBI, and is capable of guiding the search process towards reference vectors where discovering solutions is difficult as for the imbalanced test problems [115].

The adaptive framework of ESOEA is implemented to assess its efficacy whose details are mentioned in the following section.

3.5 Performance Analysis

ESOEA [138] is implemented in Matlab R2017a using a 64-bit computer with 8 GB RAM and Intel Core i7 @2.20 GHz processor. The experimental specifications, in terms of benchmark test problems, performance indicators and parameter setting of ESOEA, are discussed in detail subsequently. The performance of ESOEA is compared with several other state-of-the-art approaches to establish its efficacy. Some experiments are also performed to demonstrate the importance of the different modules of ESOEA [138].

3.5.1 Benchmark Problems

For performance analysis of ESOEA, the following benchmark problems (described in Appendix A) are considered:

1. From the DTLZ test suite [49, 74], DTLZ1, DTLZ2, DTLZ3, DTLZ4 and DTLZ7 problems are considered where N is set as described in Section A.1.

2. From the WFG test suite [10, 45, 73], WFG1 and WFG2 problems are considered with $N = 24$ (except $N = 23$ for WFG2 when M is even) [187].

3. From the IMB test suite [115], IMB1 to IMB10 problems are considered with $N = 10$.

4. From the CEC 2009 competition test suite [191], UF1 to UF10 problems are considered with $N = 30$.

3.5.2 Performance Indicators

The performance of ESOEA is assessed in terms of convergence metric [10, 134, 141], Inverted Generational Distance (IGD) [32, 85, 197] and hypervolume indicator [9, 10, 142]. Moreover, some estimated PFs are visualized using Cartesian coordinate plots for problems with $M = 2$ or 3, and using polar coordinate plots [68] for problems with higher M.

A uniformly distributed set of points $\mathcal{H}_{IGD}$ ($= \mathcal{H}_{CM}$) is sampled over the true PF for IGD (or convergence metric) evaluation. For all problems (except UF5[2]), $|\mathcal{H}_{IGD}| = 5000$ is considered as per [173]. The size of the reference set $|\mathcal{H}_{HV}|$ and the reference point $\mathbf{R}_{HV}$ for hypervolume evaluation are specified later corresponding to each experiment.

3.5.3 Experimental Settings of ESOEA/DE

For implementing ESOEA [138], Self-adaptive Neighborhood Search based Differential Evolution (SaNSDE) [185] is used as the base optimizer (SOEA). The parameters of

[2]PF of UF5 consists of $(2k_U + 1)$ discrete Pareto-optimal solutions with $k_U = 10$ [191] (Appendix A).

SaNSDE [185] such as F^{DE} (sampled from normal or Cauchy distribution), mutation probability (to choose between DE/rand/1/bin and DE/current-to-best/2/bin) and CR (sampled from normal distribution) are learned independently for each sub-population and updated (using SaNSDE's scheme) after each generation g until G^{inner}_{max}. This hyper-parameter adaptation of SaNSDE helps in addressing the sub-space specific characteristics of the fitness landscape. For generation G of the central loop, the hyper-parameter adaptation is repeated until G^{outer}_{max} after which the estimated PF is obtained. This entire framework is referred to as ESOEA/DE, hereafter, and its steps are summarized in Algorithm 3.3 for a single generation G. Source code of ESOEA/DE is available at `http://worksupplements.droppages.com/esoea`.

Algorithm 3.3 Generation G of ESOEA/DE procedure [138]

Input: $\mathcal{W}$: Set of reference vectors; $\mathcal{A}_G$: Parent population; $\mathbf{P}^{arr}_G$: Set of sub-population sizes; n_{dir}: Number of reference vectors

Output: $\mathcal{A}_{G+1}$: Parent population for next generation; $\mathbf{P}^{arr}_{G+1}$: Set of sub-population sizes for next generation

1: **procedure** ESOEA/DE($\mathcal{W}$, $\mathcal{A}_G$, $\mathbf{P}^{arr}_G$, n_{dir})
2: $\quad$ $\mathbf{F}^r_G \leftarrow$ From Eq. (3.4) using $\mathbf{F}(\mathbf{X})$ over all $\mathbf{X} \in \mathcal{A}_G$
3: $\quad$ $\mathbf{F}^s(\mathbf{X}) \leftarrow$ Using $\mathbf{F}^r_G$ and Eq. (3.5) for all $\mathbf{X} \in \mathcal{A}_G$
4: $\quad$ $\{\mathcal{A}^{sub}_{1,G}, \cdots, \mathcal{A}^{sub}_{n_{dir},G}\} = \textsc{CandAssoc}(\mathcal{W}, \forall \mathbf{X} \in \mathcal{A}_G : \mathbf{F}^s(\mathbf{X}), \mathbf{P}^{arr}_G, n_{dir})$
5: $\quad$ **for** $k = 1$ to n_{dir} (for each sub-population) **do**
6: $\quad\quad$ Let $\mathcal{A}^{par}_{k,g=0} = \mathcal{A}^{sub}_{k,G}$
7: $\quad\quad$ **for** $g = 0$ to $(G^{inner}_{max} - 1)$ (for each SaNSDE-k) **do**
8: $\quad\quad\quad$ $\mathcal{A}^{child}_{k,g} \leftarrow$ By applying SaNSDE on $\mathcal{A}^{par}_{k,g}$
9: $\quad\quad\quad$ $\forall \mathbf{X} \in \mathcal{A}^{par}_{k,g+1}$: Get $\mathbf{F}^s(\mathbf{X})$ using $\mathbf{F}^r_G$ (from line 3.3) and Eq. (3.5)
10: $\quad\quad\quad$ $\forall \mathbf{X} \in \mathcal{A}^{par}_{k,g+1}$: Get $f_{pbi}(\mathbf{X}|\mathbf{W}_k)$ (Eq. (3.3)) using $\mathbf{F}^s(X)$ and $\mathbf{W}_k \in \mathcal{W}$
11: $\quad\quad\quad$ $\mathcal{A}^{par}_{k,g+1} \leftarrow$ Select by comparing $f_{pbi}(.)$ of candidates from $\mathcal{A}^{child}_{k,g}$ and $\mathcal{A}^{par}_{k,g}$
12: $\quad\quad$ **end for**
13: $\quad\quad$ Assign $\mathcal{A}^{child}_{k,G} = \mathcal{A}^{child}_{k,g=G^{inner}_{max}}$
14: $\quad$ **end for**
15: $\quad$ $(\mathcal{R}_{1,G}, \mathcal{R}_{2,G}, \cdots) = ndset\left(\mathcal{A}_G \cup \left(\cup_{k=1}^{n_{dir}} \mathcal{A}^{child}_{k,G}\right)\right)$ by Eq. (3.8)
16: $\quad$ $\mathcal{A}_{G+1} \leftarrow$ Form using one of the three selection approaches in Fig. 3.3 where secondary selection is done by Algorithm 3.2
17: $\quad$ Evaluate $\mathbf{P}^{arr}_{G+1} = \left[S^1_{G+1}, \cdots, S^{n_{dir}}_{G+1}\right]$ by Eq. (3.9) (adaptive feedback)
18: $\quad$ **return** $\mathcal{A}_{G+1}$ and $\mathbf{P}^{arr}_{G+1}$
19: **end procedure**

The specifications (p_1 and p_2) for defining $\mathcal{W}$ are mentioned in Table 3.2 alongwith the recommended values of G^{inner}_{max} and G^{outer}_{max}. As M increases, multi-modal problems (DTLZ1 and DTLZ3) and problems with sharp-tailed PF (WFG1) are observed to require higher G^{inner}_{max} (indicating irregular landscape) whereas unimodal problems (DTLZ2 and DTLZ4) needed higher G^{outer}_{max} (indicating smoother and flat regions in landscape). While all IMB1-

10 and UF1-10 are 2- or 3-objective problems, IMB problems require lesser generations to converge due to smaller N. Among the remaining parameters for ESOEA/DE, the sub-problem sizes ($S^k_{G=0}$) are initialized to 10 which is later adapted using Eq. (3.9). The penalty parameter ($\theta_{pbi} = 5$) for PBI is set as specified in [109,160,187] and the parameter ($\beta = 75$) for the elitist selection approach (Fig. 3.3) is set as specified in [10, 141, 142].

Table 3.2: Specifications of G^{inner}_{max} (tuned in the range 10 to 30) and G^{outer}_{max} (tuned in the range 25 to 350) and the number of divisions in the boundary layer (p_1) and the inside layer (p_2) for defining the reference vectors [138].

Problems	$M = 3$	$M = 5$	$M = 10$	$M = 20$	**Problems**	$M = 2$	$M = 3$	**Problems**	$M = 2$	$M = 3$
DTLZ1	20 and 35	10 and 150	30 and 50	30 and 70	IMB1	10 and 20	-	UF1	20 and 50	-
DTLZ2	10 and 25	10 and 50	10 and 200	10 and 250	IMB2	10 and 100	-	UF2	20 and 50	-
DTLZ3	10 and 100	20 and 50	20 and 75	20 and 100	IMB3	20 and 50	-	UF3	30 and 250	-
DTLZ4	10 and 50	10 and 100	10 and 200	10 and 350	IMB4	-	10 and 50	UF4	20 and 75	-
DTLZ7	20 and 25	20 and 50	20 and 100	20 and 100	IMB5	-	10 and 20	UF5	30 and 300	-
WFG1	30 and 50	10 and 200	20 and 100	20 and 125	IMB6	-	10 and 20	UF6	20 and 150	-
WFG2	10 and 50	10 and 75	30 and 50	10 and 150	IMB7	10 and 75	-	UF7	20 and 100	-
					IMB8	10 and 80	-	UF8	-	30 and 300
					IMB9	10 and 100	-	UF9	-	30 and 50
					IMB10	-	20 and 50	UF10	-	30 and 150
p_1, p_2	13, 0	6, 0	3, 2	2, 1	p_1, p_2	100, 0	13, 0	p_1, p_2	100, 0	13, 0

With these settings, the performance of ESOEA/DE [138] is analyzed subsequently. All the results are statistically validated using the Wilcoxon's rank-sum test [173] under the null hypothesis (H_0) that the performance of ESOEA/DE is equivalent to other competitor algorithms. The statistical significance is indicated using three signs: + denoting ESOEA/DE is superior, − denoting the competitor algorithm is superior, and ∼ indicating the algorithms are equivalent.

3.5.4 Effectiveness of ESOEA/DE to Address MOO Problems

The following experiments assess the performance of ESOEA/DE on MOO problems:

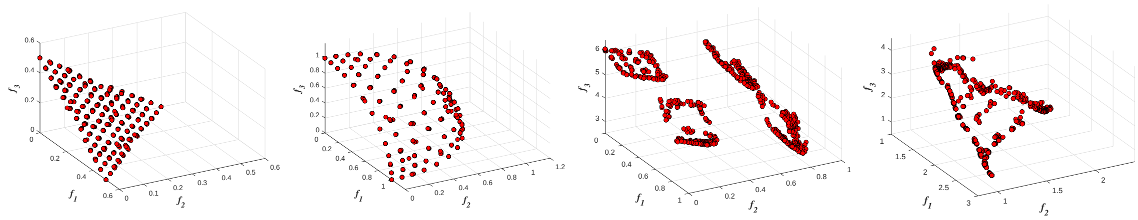
1) **On MOO problems having regular and irregular PF**: As per the specifications in [173], the mean and standard deviation of IGD for ESOEA/DE are compared in Table 3.3 with those for two of the most popular MOEAs (NSGA-II [47] and MOEA/D [150]) and the state-of-the-art MOEA (AR-MOEA [173]). The results show that ESOEA/DE performs in the best in five out of seven cases.

For some test cases in Table 3.3, the estimated PF from ESOEA/DE are visualized in Fig. 3.5 for the MOO problems. The estimated PFs of DTLZ2, DTLZ3 and DTLZ4 are identical. For DTLZ7, all the disconnected Pareto-optimal patches are obtained. For WFG1, the outline of the PF and a part of its sharp tail are obtained.

2) **On imbalanced test problems**: The performance of ESOEA/DE is studied

Table 3.3: Mean and standard deviation of IGD values over 30 independent runs for comparing MOEAs on 3-objective problems [138].

Problems	M	NSGA-II	MOEA/D	AR-MOEA	ESOEA/DE
DTLZ1	3	2.6772E-02 ± 1.36E-03 (+)	1.8973E-02 ± 3.89E-05 (+)	1.8972E-02 ± 3.52E-05 (+)	1.1561E-02 ± 1.75E-03
DTLZ2	3	6.7599E-02 ± 2.65E-03 (+)	5.1303E-02 ± 4.38E-04 (+)	5.0244E-02 ± 6.34E-05 (+)	4.6183E-02 ± 1.75E-03
DTLZ3	3	1.0247E-01 ± 1.73E-01 (+)	5.4281E-02 ± 2.52E-03 (∼)	5.2839E-02 ± 1.67E-03 (−)	5.4789E-02 ± 8.58E-04
DTLZ4	3	1.2481E-01 ± 2.23E-01 (+)	4.1204E-01 ± 3.65E-01 (+)	1.6466E-01 ± 2.11E-01 (+)	5.6493E-02 ± 8.46E-04
DTLZ7	3	7.4897E-02 ± 3.32E-03 (+)	1.2746E-01 ± 1.48E-03 (+)	6.2010E-02 ± 9.20E-04 (+)	3.4653E-02 ± 2.63E-03
WFG1	3	2.5333E-01 ± 3.02E-02 (+)	3.6315E-01 ± 3.72E-02 (+)	1.5906E-01 ± 1.17E-02 (∼)	1.6003E-01 ± 9.39E-03
WFG2	3	1.9063E-01 ± 1.27E-02 (+)	9.5329E-01 ± 7.30E-02 (+)	1.7238E-01 ± 4.52E-03 (+)	1.4627E-01 ± 9.83E-04
ESOEA vs. others (+/ − / ∼)		7/0/0	6/0/1	5/1/1	



(a) DTLZ1 ($M = 3$) (b) DTLZ4 ($M = 3$) (c) DTLZ7 ($M = 3$) (d) WFG1 ($M = 3$)

Figure 3.5: Estimated PFs from ESOEA/DE for 3-objective test problems [138].

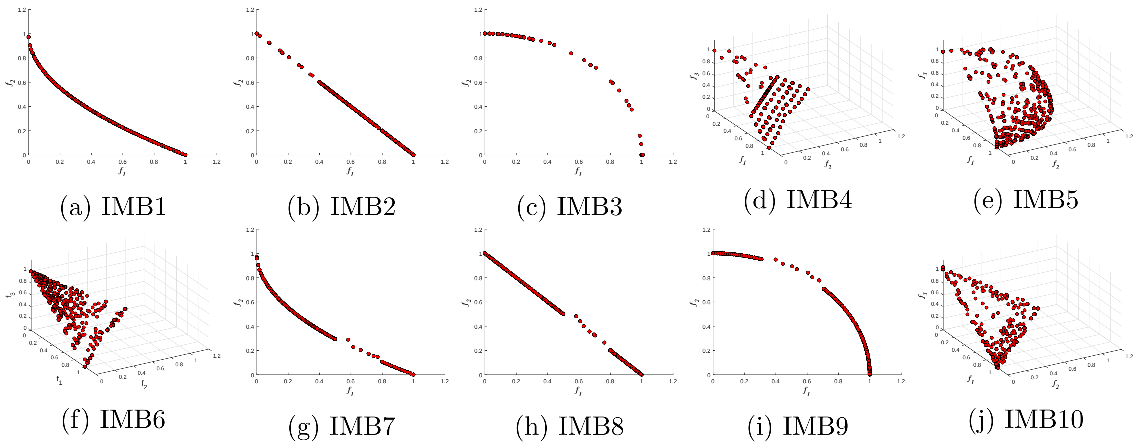
on the IMB test problems [115] using hypervolume indicator with $|\mathcal{H}_{HV}| = 10000$ and $\mathbf{R}_{HV} = \mathbf{F}^{nad} + \left[0.001, \overset{M}{\cdots}, 0.001\right]$ (where $\mathbf{F}^{nad}$ is given by Eq. (1.9)). The best, mean and worst hypervolume values of ESOEA/DE are compared in Table 3.4 with those of MOEA/D [150] and MOEA/D-M2M [115], according to the specifications in [115]. The estimated PFs of IMB problems from ESOEA/DE are shown in Fig. 3.6.

For the worst cases (Table 3.4), M2M approach performs better than ESOEA/DE as a combination of the PFs of the sub-space constrained MOO problems (from M2M) is equivalent to the true PF, whereas a combination of solutions of single-objective problems (from ESOEA) is only an approximation of the true PF [117]. Also, while ESOEA/DE is capable of exploring the difficult regions for IMB4 and IMB10 problems (Fig. 3.6d and 3.6j), it has not shown superior performance. Nonetheless, for most of the IMB problems, ESOEA/DE (due to its adaptive feedback strategy and adaptive parameters of SaNSDE) performs better than MOEA/D and MOEA/D-M2M.

3) **Against MOEAs using ensemble strategies**: While ESOEA is a decomposition-based optimization method using ensemble strategies, a few other ensemble-based MOEAs [182] are MOEA/D-DRA [101] and ENS-MOEA/D [195]. MOEA/D-DRA uses adaptive switching between simplex and center of mass crossover operators [101], and ENS-MOEA/D uses an ensemble of neighborhood sizes [195]. The mean and standard deviation of IGD values of ESOEA/DE are compared in Table 3.5 with those of MOEA/D-DRA

Table 3.4: Best, mean and worst hypervolume values over 30 independent runs for MOEAs on IMB problems [138].

IMB Test Problems	**MOEA/D**			**MOEA/D-M2M**			**ESOEA/DE**		
	best	**mean**	**worst**	**best**	**mean**	**worst**	**best**	**mean**	**worst**
IMB1	0.4684	0.4207	0.4154	0.6387	0.6375	0.6354	0.6712	0.6640	0.6569
IMB2	0.4436	0.3840	0.3390	0.4627	0.4608	0.4583	0.4902	0.4838	0.4700
IMB3	0.0385	0.0385	0.0385	0.1851	0.1836	0.1824	0.1923	0.1838	0.1728
IMB4	0.7361	0.7122	0.7030	0.7803	0.7795	0.7785	0.7716	0.7599	0.7488
IMB5	0.3916	0.3913	0.3910	0.4266	0.4229	0.4202	0.4306	0.4247	0.4205
IMB6	0.7783	0.7758	0.7751	0.7916	0.7909	0.7904	0.7998	0.7921	0.7843
IMB7	0.6327	0.6325	0.6322	0.6545	0.6540	0.6534	0.6682	0.6559	0.6415
IMB8	0.4504	0.4500	0.4496	0.4840	0.4830	0.4820	0.4885	0.4770	0.4676
IMB9	0.1716	0.1713	0.1712	0.1975	0.1960	0.1946	0.1989	0.1951	0.1911
IMB10	0.7815	0.7812	0.7807	0.7849	0.7842	0.7835	0.7698	0.7668	0.7613



(a) IMB1 (b) IMB2 (c) IMB3 (d) IMB4 (e) IMB5

(f) IMB6 (g) IMB7 (h) IMB8 (i) IMB9 (j) IMB10

Figure 3.6: Estimated PFs from ESOEA/DE for IMB test problems [138].

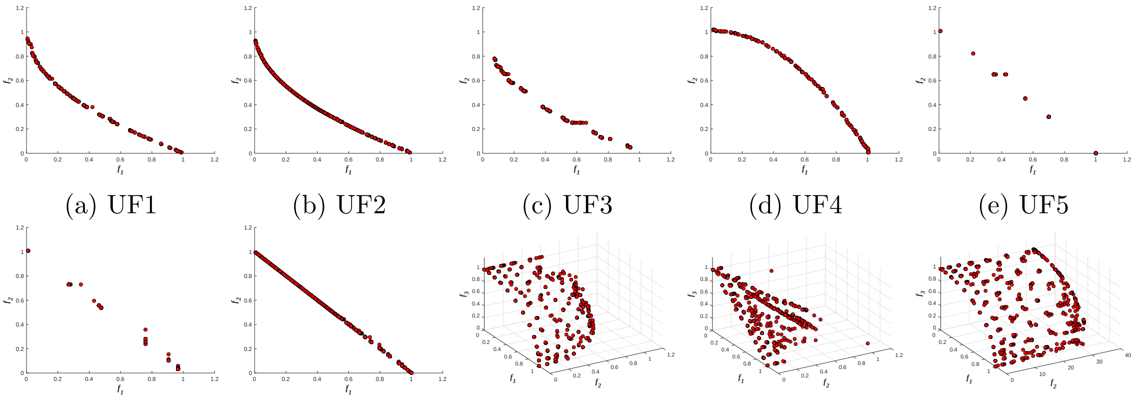
and ENS-MOEA/D on the CEC 2009 test suite [191], as per the specifications in [101]. Results show that ESOEA/DE is superior to MOEA/D-DRA in eight out of ten cases and is superior or equivalent to ENS-MOEA/D in six out of ten cases. The estimated PFs from ESOEA/DE (Fig. 3.7) bear good resemblance with the true PFs (Fig. A.4) which further supports its performance. This superior performance of ESOEA/DE is due to its adaptability in terms of both reproduction operators (SaNSDE [185]) as well as sub-population sizes. These experiments establish the versatility of ESOEA/DE for addressing MOO problems.

3.5.5 Comparison of ESOEA/DE with Various Categories of MOEAs

From each of the four categories of MOEAs (Section 1.3.1), four algorithms are chosen: NSGA-II [47] (Pareto-dominance based MOEA), MOEA/D [150] (decomposition-based MOEA), HypE [9] (indicator-based MOEA) and aDECOR [142] (objective reduction based MOEA). The respective mean and standard deviation of hypervolume indicator (Table 3.6

Table 3.5: Mean and standard deviation of IGD values over 30 independent runs for comparing MOEAs using ensemble strategies on CEC 2009 competition problems [138].

Problems	**MOEA/D-DRA**	**ENS-MOEA/D**	**ESOEA/DE**
UF1	4.2920E-03 ± 2.63E-04 (−)	1.6423E-03 ± 1.26E-04 (−)	7.3920E-03 ± 2.83E-03
UF2	5.6150E-03 ± 4.12E-04 (+)	4.0487E-03 ± 1.01E-03 (∼)	3.5153E-03 ± 3.56E-04
UF3	1.1165E-02 ± 1.31E-02 (−)	2.5916E-03 ± 4.56E-04 (−)	2.9207E-02 ± 1.15E-02
UF4	6.4145E-02 ± 4.24E-03 (+)	4.2070E-02 ± 1.33E-03 (+)	1.3010E-02 ± 4.89E-04
UF5	4.1851E-01 ± 1.36E-01 (+)	2.4811E-01 ± 4.26E-02 (+)	7.0236E-02 ± 1.37E-02
UF6	3.2736E-01 ± 1.86E-01 (+)	6.0847E-02 ± 1.98E-02 (+)	3.8250E-02 ± 3.67E-03
UF7	6.2620E-03 ± 3.31E-03 (+)	1.7286E-03 ± 8.52E-04 (−)	5.2724E-03 ± 4.50E-04
UF8	5.7443E-02 ± 3.37E-03 (+)	3.1006E-02 ± 3.01E-03 (+)	2.6375E-02 ± 3.21E-03
UF9	9.7693E-02 ± 5.43E-02 (+)	2.7874E-02 ± 9.57E-03 (−)	4.0991E-02 ± 6.99E-03
UF10	4.6265E-01 ± 3.87E-02 (+)	2.1173E-01 ± 1.99E-02 (∼)	1.9602E-01 ± 4.00E-02
ESOEA vs. others (+/ − / ∼)	8/2/0	4/4/2	

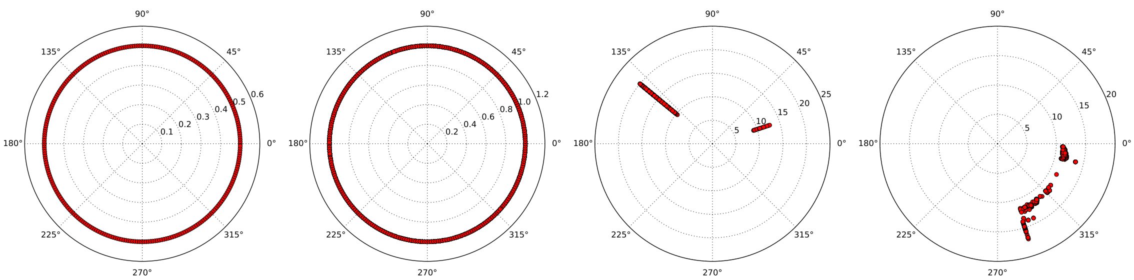


(a) UF1 (b) UF2 (c) UF3 (d) UF4 (e) UF5

(f) UF6 (g) UF7 (h) UF8 (i) UF9 (j) UF10

Figure 3.7: Estimated PFs from ESOEA/DE for CEC 2009 test instances [138].

with $|\mathcal{H}_{HV}| = 10000$ and $\mathbf{R}_{HV} = [3, \overset{M \text{ times}}{\cdots}, 3]$) and convergence metric (Table 3.7 with $|\mathcal{H}_{CM}| = 5000$) are compared with ESOEA/DE on 10- and 20-objective DTLZ problems, as per the specifications in [10, 142]. As ESOEA/DE is a DE-based method, its comparison is also done with DEMO [153]. These competitor algorithms are set as per the specifications in [10, 142]. For some test cases in Tables 3.6 and 3.7, the estimated PFs from ESOEA/DE are visualized in Fig. 3.8 using polar plots [68] (Appendix B).



(a) DTLZ1 ($M = 10$) (b) DTLZ4 ($M = 10$) (c) DTLZ7 ($M = 10$) (d) WFG1 ($M = 10$)

Figure 3.8: Estimated PFs from ESOEA/DE for 10-objective test problems [138].

From Table 3.6 (where ESOEA/DE has the best or second-best hypervolume in eight out of ten cases) and Fig. 3.8, the following observations are noted:

- The superior hypervolumes resulting from MOEA/D and ESOEA/DE are due to their embedded diversity enforcing through the decomposition of objective space.
- Even for 10-objective DTLZ7 (Fig. 3.8c), the estimated PF is outside the hyper-rectangle when $\mathbf{R}_{HV} = [3, \overset{10}{\cdots}, 3]$ which results in zero hypervolume in all test cases except from aDECOR. As aDECOR operates on reduced objective set, it has discovered a few points within the concerned hyper-rectangle.
- Overall, ESOEA/DE is superior to these six MOEAs based on hypervolume.

From Table 3.7 (where ESOEA/DE has the best or second-best convergence in nine out of ten cases) and Fig. 3.8, the following observations are noted:

- The better convergence of aDECOR and ESOEA/DE is due to the utilization of regulated elitism (Fig. 3.3) that avoids dominance resistance and trapping at local optima. Moreover, better convergence but worse diversity of aDECOR may be due to convergence of only a few solutions on the true PF as opposed to ESOEA/DE.
- For unimodal problems (DTLZ2 and DTLZ4), ESOEA/DE is only outperformed by aDECOR due to the better selection pressure for the reduced objective problem.

Table 3.6: Mean and standard deviation of hypervolume values over 30 independent runs for comparing MOEAs on DTLZ problems [138].

Problems	M	**NSGA-II**	**MOEA/D**	**HypE**	**DEMO**	**aDECOR**	**ESOEA/DE**
DTLZ1	10	0.0044 ± 0.0016 (+)	0.8132 ± 0.0984 (+)	0.0000 ± 0.0000 (+)	0.0000 ± 0.0000 (+)	0.9915 ± 0.0098 (+)	0.9970 ± 0.0015
	20	0.0000 ± 0.0000 (+)	0.7233 ± 0.1172 (+)	0.0000 ± 0.0000 (+)	0.0000 ± 0.0000 (+)	0.9994 ± 0.0206 (−)	0.9903 ± 0.0016
DTLZ2	10	0.8399 ± 0.0079 (+)	1.0000 ± 0.0000 (∼)	0.9514 ± 0.0034 (+)	0.8863 ± 0.0059 (+)	0.8765 ± 0.0018 (+)	1.0000 ± 0.0000
	20	0.8280 ± 0.0070 (+)	1.0000 ± 0.0000 (∼)	0.9372 ± 0.0019 (+)	0.8487 ± 0.0059 (+)	0.8016 ± 0.0050 (+)	1.0000 ± 0.0000
DTLZ3	10	0.0000 ± 0.0000 (+)	0.0235 ± 0.0188 (+)	0.0000 ± 0.0000 (+)	0.0000 ± 0.0000 (+)	0.9879 ± 0.0105 (+)	0.9997 ± 0.0002
	20	0.0000 ± 0.0000 (+)	0.0301 ± 0.0031 (+)	0.0000 ± 0.0000 (+)	0.0000 ± 0.0000 (+)	0.9964 ± 0.0098 (+)	0.9985 ± 0.0007
DTLZ4	10	0.9765 ± 0.0056 (+)	1.0000 ± 0.0000 (∼)	0.8741 ± 0.0169 (+)	0.9956 ± 0.0012 (+)	0.9488 ± 0.0072 (+)	1.0000 ± 0.0000
	20	0.9914 ± 0.0030 (+)	1.0000 ± 0.0000 (∼)	0.8963 ± 0.0103 (+)	0.9829 ± 0.0111 (+)	0.9420 ± 0.0155 (+)	1.0000 ± 0.0000
DTLZ7	10	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0102 ± 0.0002 (−)	0.0000 ± 0.0000
	20	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0000 ± 0.0000 (∼)	0.0097 ± 0.0003 (−)	0.0000 ± 0.0000
ESOEA vs. others (+/ − / ∼)		8/0/2	4/0/6	8/0/2	8/0/2	7/3/0	

Table 3.7: Mean and standard deviation of convergence metric over 30 independent runs for comparing MOEAs on DTLZ problems [138].

Problems	M	NSGA-II	MOEA/D	HypE	DEMO	aDECOR	ESOEA/DE
DTLZ1	10	225.4502 ± 5.9816 (+)	2.4800 ± 1.0351 (+)	146.3039 ± 2.2147 (+)	142.2519 ± 3.1073 (+)	0.3993 ± 0.0042 (+)	0.2455 ± 0.0523
	20	176.2357 ± 3.6600 (+)	3.2397 ± 1.1651 (+)	305.1945 ± 9.7488 (+)	143.5408 ± 2.7434 (+)	0.3307 ± 0.0310 (∼)	0.3491 ± 0.0405
DTLZ2	10	1.4716 ± 0.0317 (+)	0.7419 ± 0.0101 (+)	1.3979 ± 0.0156 (+)	1.3891 ± 0.0161 (+)	0.4088 ± 0.0111 (−)	0.6132 ± 0.0078
	20	1.9273 ± 0.0224 (+)	1.3116 ± 0.0050 (+)	1.9240 ± 0.0144 (+)	1.9009 ± 0.0092 (+)	0.4696 ± 0.0177 (−)	1.2169 ± 0.0367
DTLZ3	10	1048.0740 ± 39.3631 (+)	24.8627 ± 4.5587 (+)	409.5137 ± 3.9870 (+)	939.7426 ± 9.8824 (+)	0.5256 ± 0.0153 (∼)	0.8351 ± 0.0565
	20	978.3490 ± 44.9975 (+)	37.8409 ± 7.2125 (+)	911.8077 ± 5.5582 (+)	1024.4046 ± 12.5577 (+)	0.4925 ± 0.0293 (−)	1.3600 ± 0.0090
DTLZ4	10	1.1784 ± 0.0264 (+)	0.7461 ± 0.0102 (+)	0.8914 ± 0.0106 (+)	1.2663 ± 0.0347 (+)	0.4768 ± 0.0092 (−)	0.6118 ± 0.0088
	20	1.4337 ± 0.0309 (+)	1.0818 ± 0.0070 (+)	0.9572 ± 0.0077 (−)	1.6816 ± 0.0370 (+)	0.4768 ± 0.0307 (−)	0.9771 ± 0.0076
DTLZ7	10	42.6764 ± 0.7278 (+)	2.3922 ± 0.1161 (+)	40.0715 ± 0.2762 (+)	41.6292 ± 0.2632 (+)	4.4121 ± 0.0977 (+)	0.5856 ± 0.0090
	20	78.0439 ± 0.4853 (+)	6.8244 ± 0.5152 (+)	82.4481 ± 0.3383 (+)	84.4237 ± 0.5181 (+)	4.5302 ± 0.0687 (+)	1.7615 ± 0.1327
ESOEA vs. others (+/ − / ∼)		10/0/0	10/0/0	9/1/0	10/0/0	3/5/2	

Also, for DTLZ4 (biased solution density towards $f_M - f_1$ plane) [49], ESOEA/DE is not very far behind aDECOR in convergence. The adaptive feedback strategy of ESOEA/DE helps in the exploration of the search space for such problems.

- For multi-modal problems (DTLZ1 and DTLZ3), ESOEA/DE is better than most of the MOEAs which can be attributed to the fact that NSDE part of SaNSDE is useful for escaping from local minima [185].
- For DTLZ7 with disconnected PF (Fig. 3.8c shows two Pareto-optimal patches), ESOEA/DE outperforms other MOEAs. This improved performance is because the decomposition of the objective space with the adaptive feedback strategy of ESOEA/DE ensures the proper balance between exploration and exploitation.

Thus, ESOEA/DE has superior performance on a variety of problem characteristics as compared to popular MOO and MaOO algorithms from various categories of MOEAs.

3.5.6 Comparison of ESOEA/DE with MOEAs using Reference Vectors

The performance of ESOEA/DE are compared with other contemporary reference vector based approaches like NSGA-III [45], MOEA/DD [109], RVEA [28] and AR-MOEA [173]. As per the specifications in [173], the mean and standard deviation of hypervolumes of

Table 3.8: Mean and standard deviation of hypervolume values over 30 independent runs for comparing ESOEA/DE with other reference vector associated MaOO algorithms [138].

Problems	M	**NSGA-III**	**MOEA/DD**	**RVEA**	**AR-MOEA**	**ESOEA/DE**
DTLZ1	5	9.7456E-01 ± 4.86E-04 (−)	9.7487E-01 ± 1.94E-04 (−)	9.7478E-01 ± 3.14E-04 (−)	9.7492E-01 ± 1.53E-04 (−)	9.7113E-01 ± 1.05E-03
	10	9.8390E-01 ± 4.66E-02 (−)	9.9957E-01 ± 4.55E-05 (−)	9.9967E-01 ± 2.82E-05 (−)	9.9971E-01 ± 8.84E-06 (−)	9.1338E-01 ± 9.48E-02
DTLZ2	5	7.9035E-01 ± 8.70E-04 (+)	7.9294E-01 ± 5.49E-02 (∼)	7.9209E-01 ± 6.51E-04 (+)	7.9047E-01 ± 8.44E-04 (+)	8.1198E-01 ± 1.21E-03
	10	9.4923E-01 ± 3.10E-02 (+)	9.6735E-01 ± 2.41E-04 (+)	9.6751E-01 ± 2.27E-04 (+)	9.6432E-01 ± 8.26E-04 (+)	9.6852E-01 ± 4.00E-04
DTLZ3	5	5.9177E-01 ± 2.97E-01 (+)	7.7880E-01 ± 1.20E-02 (+)	7.3843E-01 ± 7.61E-02 (+)	7.7240E-01 ± 7.36E-03 (+)	8.5123E-01 ± 8.62E-03
	10	3.8532E-01 ± 3.38E-01 (+)	9.6669E-01 ± 2.00E-03 (−)	9.6065E-01 ± 6.15E-03 (−)	9.6723E-01 ± 2.93E-03 (−)	8.8142E-01 ± 8.66E-02
DTLZ4	5	7.8203E-01 ± 2.89E-02 (∼)	7.9366E-01 ± 5.01E-04 (+)	7.9307E-01 ± 4.99E-04 (+)	7.9077E-01 ± 6.88E-04 (+)	8.0225E-01 ± 9.41E-04
	10	9.6625E-01 ± 9.90E-04 (+)	9.6837E-01 ± 3.23E-03 (∼)	9.6964E-01 ± 2.83E-04 (∼)	9.6902E-01 ± 5.57E-04 (∼)	9.6966E-01 ± 8.06E-04
DTLZ7	5	2.4167E-01 ± 4.33E-03 (+)	9.0909E-02 ± 4.94E-07 (+)	2.0007E-01 ± 9.91E-03 (+)	2.3599E-01 ± 2.48E-03 (+)	5.0892E-01 ± 1.54E-02
	10	1.9584E-01 ± 1.26E-02 (+)	1.1971E-03 ± 3.11E-04 (+)	1.4380E-01 ± 1.51E-02 (+)	1.4646E-01 ± 7.03E-03 (+)	6.0554E-01 ± 3.05E-02
WFG1	5	7.8837E-01 ± 3.33E-02 (+)	7.7075E-01 ± 5.71E-02 (+)	8.6621E-01 ± 4.04E-02 (+)	9.0787E-01 ± 2.65E-02 (+)	9.5295E-01 ± 6.42E-03
	10	7.0682E-01 ± 4.79E-02 (+)	9.8947E-01 ± 2.24E-02 (∼)	9.8712E-01 ± 2.83E-02 (+)	9.4718E-01 ± 3.69E-02 (+)	9.9064E-01 ± 1.74E-03
WFG2	5	9.9246E-01 ± 1.19E-03 (−)	9.6933E-01 ± 4.70E-03 (−)	9.8809E-01 ± 1.99E-03 (−)	9.9469E-01 ± 5.81E-04 (−)	7.0163E-01 ± 3.58E-02
	10	9.9671E-01 ± 1.69E-03 (−)	9.6285E-01 ± 6.50E-03 (−)	9.8615E-01 ± 3.22E-03 (−)	9.9508E-01 ± 1.06E-03 (−)	7.3139E-01 ± 2.75E-02
ESOEA vs. others (+/ − / ∼)		9/4/1	6/5/3	8/5/1	8/5/1	

normalized PFs (between $\mathbf{F}^{nad}$ and $\mathbf{F}^{ide}$) [134] are noted in Table 3.8 with $|\mathcal{H}_{HV}| = 1,000,000$ and $\mathbf{R}_{HV} = \left[1.1, \overset{M \text{ times}}{\cdots}, 1.1\right]$. ESOEA/DE outperforms these MOEAs in nine out of 14 test cases, followed by AR-MOEA which shows superior performance in four out of 14 test cases. The following observations are made from Table 3.8:

- The superior performance of ESOEA/DE over NSGA-III is due to its regulated elitist selection (Fig. 3.3) and adaptive feedback scheme (Eq. (3.9)).
- Adaptive feedback of ESOEA/DE is also beneficial for DTLZ7 (having disconnected PF) and WFG1 (having sharp-tailed PF) over MOEA/DD which benefits problems like DTLZ1 and DTLZ3 (having multi-modal nature but regular PF).
- ESOEA/DE outperforms RVEA [28] as the former algorithm performs mating in a local neighborhood to improve exploitation to the search space.
- Unlike ESOEA/DE, the reference vector adaptation of AR-MOEA becomes detrimental for problems like DTLZ4 (having biased solution density).

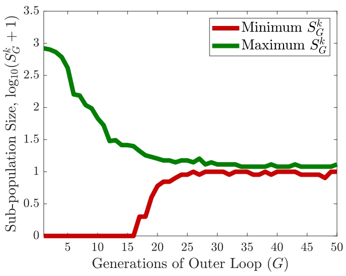
Table 3.9: Mean IGD values over 30 independent runs for comparing adaptive MaOO algorithms on multi-modal problems with regular Pareto-Fronts [138].

Problems	M	A-NSGA-III	RVEA*	AR-MOEA	ESOEA/DE
DTLZ1	3	2.3434E-02 (+)	2.8841E-02 (+)	1.8931E-02 (+)	1.1561E-02
	5	6.3446E-02 (+)	7.1247E-02 (+)	6.2861E-02 (+)	4.0492E-02
	10	1.7341E-01 (−)	2.5566E-01 (−)	1.4292E-01 (−)	4.0598E-01
DTLZ3	3	6.8590E-02 (+)	7.2553E-02 (+)	5.0276E-02 (−)	5.9847E-02
	5	2.0753E-01 (+)	2.8049E-01 (+)	1.9531E-01 (+)	1.1701E-01
	10	2.0584E+00 (+)	6.9093E-01 (−)	4.9583E-01 (−)	1.2617E+00
ESOEA vs. others (+/ − / ∼)		5/1/0	4/2/0	3/3/0	

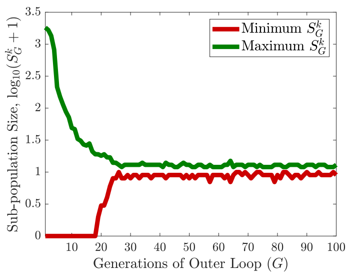
For comparing the adaptive tendency of ESOEA/DE with reference vector adaptation based MOEAs like A-NSGA-III [89], RVEA* [28] and AR-MOEA [173], the mean IGD values from DTLZ1 and DTLZ3 problems are noted in Table 3.9 using the specifications of [173]. Results show that ESOEA/DE is better than A-NSGA-III and RVEA* as the adaptive strategies of A-NSGA-III and RVEA* are designed specifically for problems with irregular PF. However, ESOEA/DE ties with AR-MOEA for problems with regular PF.

3.5.7 Analysis of the Adaptive Feedback Scheme of ESOEA

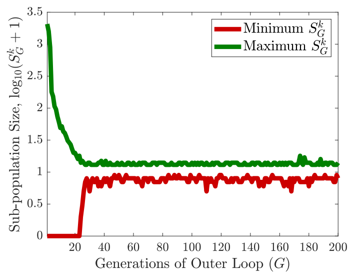
The adaptive feedback scheme of ESOEA intends to redistribute the solutions such that all sub-spaces have an equal number of associated solutions. As the global population size is $10 \times n_{dir}$, the minimum sub-population size (S_G^k using Eq. (3.9)) over all the sub-spaces can atmost be 10. For a problem with biased solution density (DTLZ4), the variations in S_G^k are noted across generations (G) of ESOEA/DE in Fig. 3.9 when $M = \{3, 5, 10\}$. For all these test cases, the minimum S_G^k and the maximum S_G^k vary marginally after reaching a near-equilibrium value of 10 (or $\log_{10}(10 + 1) \approx 1$). Thus, this experiment shows the role of the feedback strategy of ESOEA in a uniform exploration of the search space.



(a) DTLZ4 ($M = 3$)



(b) DTLZ4 ($M = 5$)



(c) DTLZ4 ($M = 10$)

Figure 3.9: Variations of minimum and maximum sub-population size across generations of ESOEA/DE for a problem with biased solution density (DTLZ4).

3.5.8 Effectiveness of Components of ESOEA/DE

For investigating the effectiveness of various components of ESOEA/DE, the following experiments are performed where the basic framework of ESOEA/DE is kept intact, except changing one of its components:

1. *Experiment-I*: Using DE/rand/1/bin [149, 165] instead of SaNSDE
2. *Experiment-II*: Using non-dominated sorting with crowding distance (Fig. 2.1b) instead of the selection scheme of ESOEA (Fig. 3.3)
3. *Experiment-III*: Using weighted sum ($\theta_{pbi} = 0$) instead of PBI function (Eq. (3.3))
4. *Experiment-IV*: Initializing $\mathcal{W}$ based on maximization of minimum pair-wise distance between reference vectors instead of Das and Dennis' two-layered approach [40]

These experiments are performed on both MOO and MaOO problems. For the comparison, the mean hypervolumes of normalized PFs (between $\mathbf{F}^{nad}$ and $\mathbf{F}^{ide}$) [134] are noted in Table 3.10 with $|\mathcal{H}_{HV}| = 1,000,000$ and $\mathbf{R}_{HV} = \left[1.1, \overset{M \text{ times}}{\cdots}, 1.1\right]$. From Table 3.10, it is observed that ESOEA/DE performs best or second-best in 21 out of 28 test cases as compared to the other experimental frameworks. Even for 3-objective DTLZ3 and 20-objective WFG1 problems, the hypervolume value of ESOEA/DE is not significantly different from the second-best value. The following insights are obtained from Table 3.10:

- ESOEA/DE performs better or equivalent to Experiment-I in 21 out of 28 test cases as using SaNSDE (ESOEA/DE) instead of DE/rand/1/bin (Experiment-I) aids in local adaptability to problem characteristics. Also, Experiment-I shows the second-highest number of best or second-best performances, supporting the efficacy of the overall framework of ESOEA.

- For some 3-objective problems, the framework of Experiment-II performs better than the ESOEA/DE framework. As M increases, the effectiveness of the non-dominated sorting with crowding distance (Experiment-II) fades away in comparison to the regulated elitism scheme of ESOEA/DE.

- ESOEA/DE performs significantly better or equivalent to Experiment-III in 23 out of 28 test cases as using PBI (ESOEA/DE) instead of weighted sum (Experiment-III) presents a better trade-off between convergence and diversity.

- The alternate weight distribution scheme (Experiment-IV) gives an inferior result over ESOEA/DE when M is small. However, it performs better or equivalent to ESOEA/DE in 5 test cases when $M = 10$ or 20 because for such problems there are very few intermediate reference vectors (by Das and Dennis' approach in ESOEA/DE) [13], hampering the diversity of the resulting PF.

These observations demonstrate that the combination of modules of ESOEA/DE is significantly better than the usual existing modules.

Table 3.10: Mean hypervolume values over 30 independent runs from ESOEA/DE to establish the effectiveness of its different modules [138].

	Problems	ESOEA/DE	Experiment-I	Experiment-II	Experiment-III	Experiment-IV
$M = 3$	DTLZ1	8.5453E-01	8.4154E-01 (+)	8.5658E-01 (−)	8.1774E-01 (+)	7.9633E-01 (+)
	DTLZ2	5.7467E-01	5.6276E-01 (+)	5.8463E-01 (−)	5.7097E-01 (+)	5.1058E-01 (+)
	DTLZ3	5.6074E-01	5.5885E-01 (+)	5.8437E-01 (−)	5.6102E-01 (∼)	5.0282E-01 (+)
	DTLZ4	5.6318E-01	5.6136E-01 (+)	5.8660E-01 (−)	5.5706E-01 (+)	5.0291E-01 (+)
	DTLZ7	5.7580E-01	4.4063E-01 (+)	5.3227E-01 (+)	5.5827E-01 (∼)	4.4002E-01 (+)
	WFG1	8.0581E-01	7.8667E-01 (+)	7.6349E-01 (+)	7.5286E-01 (+)	7.3019E-01 (+)
	WFG2	6.9209E-01	6.6852E-01 (+)	6.8132E-01 (∼)	6.5564E-01 (+)	6.1144E-01 (+)
$M = 5$	DTLZ1	9.7113E-01	9.7427E-01 (−)	9.7014E-01 (∼)	9.8447E-01 (−)	8.5232E-01 (+)
	DTLZ2	8.1198E-01	8.0876E-01 (+)	8.0901E-01 (+)	8.0099E-01 (+)	6.8135E-01 (+)
	DTLZ3	8.5123E-01	8.0531E-01 (+)	8.4474E-01 (∼)	8.4419E-01 (∼)	7.1841E-01 (+)
	DTLZ4	8.0225E-01	8.0710E-01 (−)	8.0046E-01 (+)	7.9584E-01 (+)	6.7313E-01 (+)
	DTLZ7	5.0892E-01	5.3167E-01 (−)	5.6352E-01 (−)	5.5237E-01 (−)	3.9291E-01 (+)
	WFG1	9.5295E-01	9.4304E-01 (+)	9.3726E-01 (+)	9.3512E-01 (+)	5.1293E-01 (+)
	WFG2	7.0163E-01	6.9669E-01 (∼)	6.4977E-01 (+)	6.6672E-01 (+)	8.0880E-01 (−)
$M = 10$	DTLZ1	9.1338E-01	9.9814E-01 (−)	8.9365E-01 (∼)	9.6804E-01 (−)	7.4837E-01 (+)
	DTLZ2	9.6852E-01	9.1965E-01 (+)	9.2339E-01 (+)	9.6504E-01 (+)	7.3232E-01 (+)
	DTLZ3	8.8142E-01	7.7287E-01 (+)	6.9587E-01 (+)	8.3545E-01 (+)	8.8016E-01 (∼)
	DTLZ4	9.6966E-01	8.7187E-01 (+)	8.7648E-01 (+)	9.6828E-01 (+)	7.2930E-01 (+)
	DTLZ7	6.0554E-01	6.9478E-01 (−)	7.7617E-01 (−)	5.6036E-01 (∼)	6.2205E-01 (∼)
	WFG1	9.9064E-01	9.8725E-01 (+)	9.0984E-01 (+)	9.2766E-01 (+)	8.4420E-01 (+)
	WFG2	7.3139E-01	6.7190E-01 (+)	6.4810E-01 (+)	6.2490E-01 (+)	5.9579E-01 (+)
$M = 20$	DTLZ1	8.1368E-01	9.6994E-01 (−)	8.5046E-01 (∼)	9.3120E-01 (−)	9.9752E-01 (−)
	DTLZ2	9.8335E-01	9.6377E-01 (+)	9.1597E-01 (+)	9.6395E-01 (+)	8.5250E-01 (+)
	DTLZ3	7.5869E-01	6.7534E-01 (+)	5.7603E-01 (+)	6.8463E-01 (+)	7.1100E-01 (∼)
	DTLZ4	9.9875E-01	9.9871E-01 (∼)	9.9456E-01 (+)	9.9834E-01 (+)	9.7615E-01 (+)
	DTLZ7	8.0038E-01	3.1698E-01 (+)	8.4948E-01 (−)	6.4741E-01 (+)	5.2040E-01 (+)
	WFG1	7.5766E-01	7.8495E-01 (−)	7.1659E-01 (+)	7.7598E-01 (∼)	6.6224E-01 (+)
	WFG2	6.2019E-01	5.8943E-01 (+)	6.0747E-01 (+)	6.3564E-01 (−)	6.0435E-01 (∼)
ESOEA vs. others (+/ − / ∼)			19/7/2	16/7/5	18/5/5	22/2/4

3.6 Conclusion

Motivated by the need for an evolutionary optimizer which is effective for a broad spectrum of problem characteristics, this chapter presents ESOEA/DE, comprising an ensemble of SaNSDE. Using the PBI function, ESOEA/DE transforms MaOO problems into multiple single-objective sub-problems, each constrained to a sub-space created by the decomposition of the objective space. These sub-problems are solved collaboratively. Additionally, the adaptive feedback, the modified candidate association and the regulated elitist scheme with a novel secondary selection help ESOEA to tackle several MOO/MaOO problems.

While the results exhibit the robustness of ESOEA/DE by demonstrating good convergence and superior diversity over several problems (with different modalities, biased solution density, disconnected PFs, sharp-tailed PFs, imbalance difficulties and variable linkage difficulties), the reasoning behind its performance is established only through qualitative analysis. Lack of any theoretical analyses hinders the further understanding of the weaknesses and the scope of improvement of such optimization algorithms. Hence, the next chapter theoretically analyzes some basic concepts for decomposition-based MOEAs.

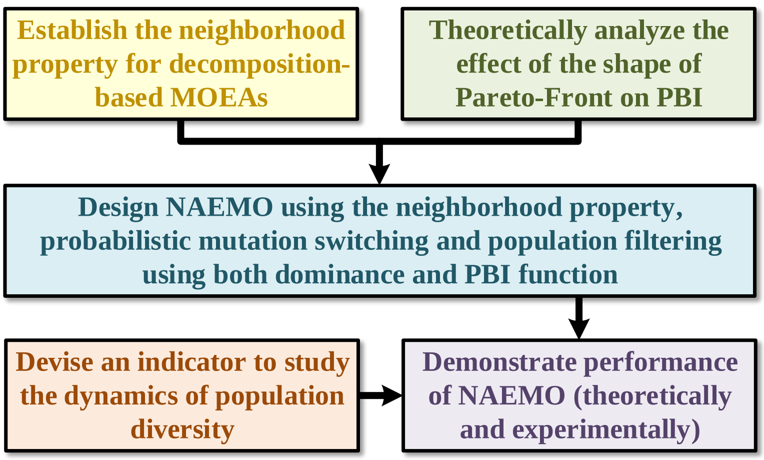
Chapter 4

NAEMO: Neighborhood-sensitive Archived Evolutionary Many-objective Optimization Algorithm [160]

<u>**Outline**</u>

Objective: *To establish the neighborhood property related to reference-vector assisted decomposition of objective space and to use it to develop an algorithm with theoretically and experimentally proven performance enhancement for many-objective optimization problems.*

Workflow:



4.1 Introduction

From the previous chapter, it is noted that one of the prominent techniques to address Multi-Objective Optimization (MOO) and Many-Objective Optimization (MaOO) problems (Eq. (1.2)) is the reference-vector assisted decomposition-based Multi-Objective Evolutionary Algorithms (MOEAs) [119]. However, literature severely lacks formal mathematical analysis to establish the reason behind superior performance of such methods.

Motivated by this research gap, the neighborhood property of MaOO problems is recognized and is used to develop the Neighborhood-sensitive Archived Evolutionary Many-objective Optimization (NAEMO) algorithm [160]. In NAEMO, mating occurs within a local neighborhood and every sub-space continues to retain at least one associated solution. This preservation of solutions leads to a monotonic improvement in diversity (theoretically and experimentally justified). Moreover, to combine the advantages of various mutation strategies, probabilistic mutation switching concept is introduced and to keep the archive size under control, periodic filtering modules are integrated with the NAEMO framework. In terms of inverted generational distance, hypervolume values and purity metric, NAEMO outperforms several state-of-the-art MOEAs on DTLZ1-4 test problems for up to 15 objectives. Further experiments show that NAEMO is also competitive to M2M-based algorithms on the IMB problems. Thus, NAEMO is a robust algorithm, which is additionally supported by theoretical foundations.

Rest of this chapter is structured as follows. The work presented in this chapter is motivated by the arguments discussed in Section 4.2. The neighborhood theorem is introduced in Section 4.3. The effect of the shape of the Pareto-Front (PF) on the PBI function (Eq. (3.3)) is theoretically analyzed in Section 4.4. Section 4.5 presents NAEMO along with preliminary theoretical analyses. Its performance is compared with other MOEAs in Section 4.6. The chapter is finally concluded in Section 4.7.

4.2 Research Gap Analysis

Among various approaches to deal with MaOO problems, reference-vector based algorithms such as MOEA/D [150], NSGA-III [45], θ-DEA [187], MOEA/DD [109] and their variants have been developed which perform well for problems with number of objectives (M) as high as 15. The concept of decomposition in MOEA/D has also been combined with other meta-heuristics such as Particle Swarm Optimization for yielding decomposition-based Multi-Objective Particle Swarm Optimization (dMOPSO) [121].

It is challenging for the general MOEAs to obtain the complete PF for IMB problems [115], which have *difficult* regions (as discussed in Section A.3). The M2M based algorithms [115, 117], which decomposes the MOO problems into simpler MOO problems, are effective for IMB problems (theoretically supported) [117].

Theoretical analyses and results are vital for understanding optimization problems and algorithms. The convergence of MOEAs are formally investigated in [66]. Some concepts of reference-vector assisted decomposition-based MOEAs are studied in [111]. However, much theoretical work on MaOO is still not present in the literature. The working and reasoning behind the performance of MOEAs are usually qualitative. Formal theoretical analysis will aid in finding the weaknesses of algorithms and in making improvements with a concrete theoretical basis. This chapter is dedicated to filling this research gap and has the following contributions:

1. The neighborhood property of decomposition-based MOEAs is identified and utilizing it Neighborhood-sensitive Archived Evolutionary Many-objective Optimization (NAEMO) algorithm [160] is developed. NAEMO introduces convergence-based filtering and diversity-based filtering schemes, supported by theoretical analyses.

2. NAEMO uses both PBI function and Pareto-dominance simultaneously. It also demonstrates a candidate vector generation scheme using the probabilistic mutation switching concept.

3. A theoretical analysis of the PBI function is presented to explain how the shape of the actual PF affects the final solution if the PBI function is used.

4. An indicator (*D_metric*) [161] is developed to study the diversity attainment behavior of MOEAs, which is used to compare NAEMO with other MOEAs.

4.3 Theoretical Outline of the Neighborhood Property

A notion of the spatial relationship between objective space and decision space, which is created by reference-vector assisted decomposition of the objective space, is conveyed by the following theorem based on which NAEMO is developed in [160].

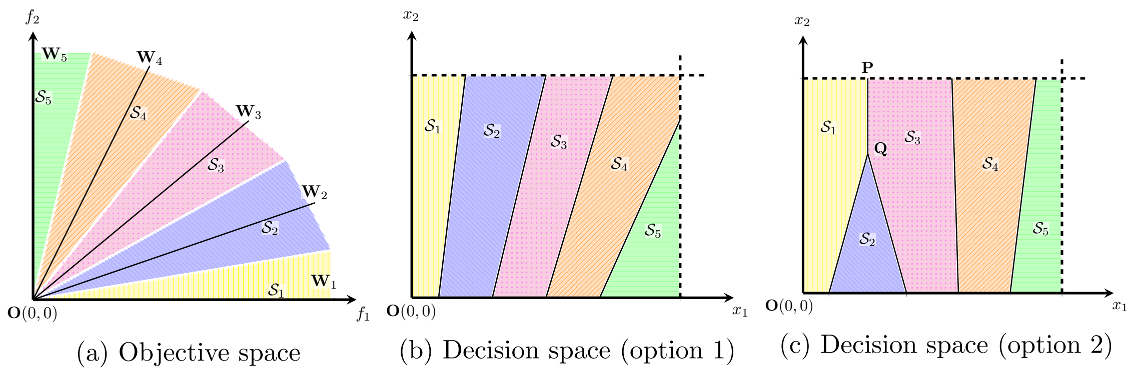


Figure 4.1: Neighborhood property [160]: (a) reference vectors decomposing the objective space in to 5 regions ($\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$, $\mathcal{S}_5$), (b) when adjacent regions of the objective space are also adjacent in the decision space, (c) when non-adjacent regions of the objective space ($\mathcal{S}_1$ and $\mathcal{S}_3$) are adjacent in the decision space sharing a common boundary $\overline{\mathbf{PQ}}$.

Theorem 4.1 (Neighborhood property)**.** *The regions corresponding to each reference vector in the objective space, which do not share a common boundary in the objective space, do not share a common boundary in the decision space either.*

Proof. The reference-vectors ($\mathbf{W}_1$, $\mathbf{W}_2$, $\mathbf{W}_3$, $\mathbf{W}_4$ and $\mathbf{W}_5$) decomposes the objective space into corresponding regions ($\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$ and $\mathcal{S}_5$) as shown in Fig. 4.1a.

It is first assumed that some regions (like $\mathcal{S}_1$ and $\mathcal{S}_3$) which do not have a common boundary in the objective space can have a common boundary in the decision space. Then, a possible decision space visualization might be as shown in Fig. 4.1c, where the line $\overline{\mathbf{PQ}}$ is common to both $\mathcal{S}_1$ and $\mathcal{S}_3$. Hence, when mapped to the objective space, this line $\overline{\mathbf{PQ}}$ would correspond to a curve common to both $\mathcal{S}_1$ and $\mathcal{S}_3$ in the objective space. However, such a curve can never be present as the region $\mathcal{S}_2$ always comes between $\mathcal{S}_1$ and $\mathcal{S}_3$ (Fig. 4.1a). Therefore, the initial assumption that $\mathcal{S}_1$ and $\mathcal{S}_3$ can have a common region in the decision space is wrong. This analysis proves the theorem by contraposition.

Thus, the non-adjacent regions (corresponding to the reference-vectors) in the objective space (Fig. 4.1a) will also have non-adjacent regions in the decision space (Fig. 4.1b). This property of MOO/MaOO problems is denoted as the *neighborhood property.* □

4.4 Analyzing the Penalty-based Boundary Intersection

The PBI function [40, 150] (Eq. (3.3)) is used over several MOEAs due to its efficacy. It provides a measure of fitness for the sub-problems (created by the decomposition of objective space). Due to its popularity in recent days, it is necessary to perform a theoretical

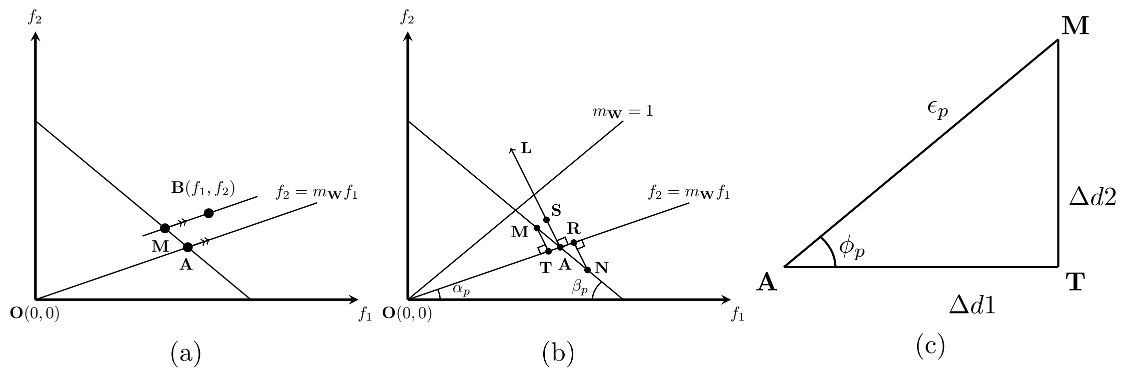


Figure 4.2: Given a reference vector $\overrightarrow{\mathbf{OA}}$, (a) for a non-optimal point $\mathbf{B}$, there is an optimal point $\mathbf{M}$, (b) for a optimal point $\mathbf{N}$, there is a better point $\mathbf{S}$ along $\overrightarrow{\mathbf{AL}}$ such that $\overline{\mathbf{AS}} = \overline{\mathbf{RN}} = d2_N$, (c) Triangle $\mathbf{AMT}$ to demonstrate $f_{pbi}(\mathbf{A}) < f_{pbi}(\mathbf{M}), \forall \mathbf{AM} = \epsilon_p$ [160].

analysis of the PBI function. This analysis is conducted for different shapes of the PF.

4.4.1 Linear Pareto-Front

For a problem with linear PF, the minimal objective values are attained using PBI approach with $\theta_{pbi} > 1$. This claim is analyzed, theoretically, using the following theorem.

Theorem 4.2. *The optimal point (objective values) for all sub-spaces for a multi-objective optimization problem with a linear Pareto-Front is the point of intersection of the reference vector with the true Pareto-Front, if the penalty factor is higher than 1, i.e.,* $\theta_{pbi} > 1$.

Proof. Let a problem with linear PF be considered. It is also assumed that the scales for all the objective functions are same and normalized such that

$$\text{Linear Pareto-Front: } f_1 + f_2 = 1. \tag{4.1}$$

Assuming $m_{\mathbf{W}}$ as the slope of the reference vector, the ideal optimal point $\mathbf{A}$ (at intersection of reference vector with PF) is shown in Fig. 4.2a and is given as follows:

$$\text{Ideal Optimal Point } \mathbf{A} : \left(\frac{1}{m_{\mathbf{W}} + 1}, \frac{m_{\mathbf{W}}}{m_{\mathbf{W}} + 1} \right). \tag{4.2}$$

The parameters ($d1_B$ and $d2_B$) of PBI function for a random point $\mathbf{B}$ is given as:

$$\begin{aligned} d1_B = d1(\mathbf{B}) = \overrightarrow{\mathbf{OB}} \cdot \widehat{\mathbf{OA}} \text{ and } d2_B = d2(\mathbf{B}) = \left\| \overrightarrow{\mathbf{OB}} - d1_B \widehat{\mathbf{OA}} \right\|, \\ \text{where } \widehat{\mathbf{OA}} = \frac{1}{\sqrt{1 + m_{\mathbf{W}}^2}} \hat{f}_1 + \frac{m_{\mathbf{W}}}{\sqrt{1 + m_{\mathbf{W}}^2}} \hat{f}_2. \end{aligned} \tag{4.3}$$

Thus, the fitness (through PBI-based scalarization using Eq. (3.3)) of $\mathbf{B}$ along the reference-vector $\overrightarrow{\mathbf{OA}} : f_2 = m_{\mathbf{W}} \times f_1$ is given as follows:

$$f_{pbi}(\mathbf{B}) = \frac{(f_1 + m_{\mathbf{W}} \times f_2) + \theta_{pbi} \left| f_2 - m_{\mathbf{W}} \times f_1 \right|}{\sqrt{1 + m_{\mathbf{W}}^2}}. \tag{4.4}$$

Hence, the penalty factor (θ_{pbi}) must satisfy the following equation:

$$\min\left(f_{pbi}(\mathbf{B})\right) \text{ is at } \mathbf{A} = \left(\frac{1}{m_{\mathbf{W}} + 1}, \frac{m_{\mathbf{W}}}{m_{\mathbf{W}} + 1}\right). \tag{4.5}$$

This analysis considers the following arguments:

1. A line is considered parallel to $f_2 = m_{\mathbf{W}} f_1$ and passing through $\mathbf{B}$ ($\overleftrightarrow{\mathbf{OA}} \| \overleftrightarrow{\mathbf{BM}}$ in Fig. 4.2a). This line intersects PF at $\mathbf{M}$. This point $\mathbf{M}$ has the same value of $d2$ as $\mathbf{B}$, i.e., $d2_B = d2_M$ as $\overleftrightarrow{\mathbf{OA}} \| \overleftrightarrow{\mathbf{BM}}$. But the $d1$ of $\mathbf{M}$ is better than $d1$ of $\mathbf{B}$, i.e., $d1_M < d1_B$. Hence, $f_{pbi}(\mathbf{M})$ $(= d1_M + \theta_{pbi} \times d2_M)$ is less than $f_{pbi}(\mathbf{B})$ $(= d1_B + \theta_{pbi} \times d2_B)$, which implies that for every point $\mathbf{B}$ not on the PF, there is always a better point $\mathbf{M}$ on the PF using the PBI function (Eq. (3.3)).

2. Next, a point $\mathbf{N}$ is considered on the other side of the reference vector ($\overrightarrow{\mathbf{OA}}$), opposite to $\mathbf{M}$. This point $\mathbf{N}$ is at ϵ_p distance from $\mathbf{A}$ and lies on the PF (Fig. 4.2b), i.e., $\overline{\mathbf{AM}} = \overline{\mathbf{AN}} = \epsilon_p$. The PBI parameters of $\mathbf{N}$ are $d1_N = \overline{\mathbf{OR}}$ and $d2_N = \overline{\mathbf{RN}}$. The vector ($\overrightarrow{\mathbf{AL}}$) is considered to be perpendicular to the reference vector $\overrightarrow{\mathbf{OA}} : f_2 = m_{\mathbf{W}} f_1$. Let the point on $\overrightarrow{\mathbf{AL}}$ at a distance $d2_N$ from $\mathbf{A}$ be $\mathbf{S}$. Such a point will have $d2_S = d2_N = \overline{\mathbf{RN}}$ and $d1_S = \overline{\mathbf{OA}}$. It can be seen that $d1_N = d1_S + \overline{\mathbf{AR}}$ with $\overline{\mathbf{AR}} > 0$. However, $\overline{\mathbf{AR}} \to 0$ as $\epsilon_p \to 0$. Thus, for every point $\mathbf{N}$ on the optimal surface, there exists a better point $\mathbf{S}$ on $\overrightarrow{\mathbf{AL}}$. Since $\mathbf{S}$ is not on the PF, there will again be a further better point on the PF using the PBI function similar to case 1.

3. Now, when traversing from $\mathbf{M}$ to $\mathbf{A}$ (Fig. 4.2b), $d1$ increases but $d2$ decreases. This change is evident from the associated parameters of PBI function as follows:

$$\begin{gathered} f_{pbi}(\mathbf{M}) = d1_M + \theta_{pbi} \times d2_M \text{ and } f_{pbi}(\mathbf{A}) = d1_A + \theta_{pbi} \times d2_A, \\ \text{where } d1_A = \overline{\mathbf{OA}} = \overline{\mathbf{OT}} + \overline{\mathbf{AT}} = d1_M + \overline{\mathbf{AT}}, \\ d2_A = 0 \text{ and } d2_M = \overline{\mathbf{MT}}. \end{gathered} \tag{4.6}$$

For $\mathbf{A}$ to be optimal after minimizing $f_{pbi}(.)$, θ_{pbi} must satisfy the following:

$$\begin{aligned}
& f_{pbi}(\mathbf{M}) > f_{pbi}(\mathbf{A}), \forall \epsilon_p \\
\implies & d1_M + \theta_{pbi} \times d2_M > d1_A + \theta_{pbi} \times d2_A \\
\implies & \theta_{pbi} > \frac{d1_A - d1_M}{d2_M - d2_A} = \frac{\Delta d1}{\Delta d2} = \frac{\overline{\mathbf{AT}}}{\overline{\mathbf{MT}}} \text{ (From Fig. 4.2c)} \\
\implies & \theta_{pbi} > \frac{1}{\tan \phi_p} = \frac{1}{\tan(\alpha_p + \beta_p)} \text{ (From Fig. 4.2b)} \\
\implies & \theta_{pbi} > \frac{1 - m_{\mathbf{W}}}{1 + m_{\mathbf{W}}} \\
\implies & \theta_{pbi} > 1, \text{ when } 0 \leq m_{\mathbf{W}} \leq 1.
\end{aligned} \tag{4.7}$$

Since everything is symmetric about the reference vector with $m_{\mathbf{W}} = 1$, considering $0 \leq m_{\mathbf{W}} \leq 1$ presents a generic analysis. Therefore, $\theta_{pbi} > 1$ ensures that $\mathbf{A}$ is optimal point for all values of $m_{\mathbf{W}}$ in case of a linear PF. □

4.4.2 General Pareto-Front

The shape of the final PF can also be concave or convex with different degrees of curvature (δ_c). Assuming a symmetric PF about the reference vector with $m_{\mathbf{W}} = 1$, the PF can be approximated as follows:

$$\text{General Pareto-Front: } f_1^{\delta_c} + f_2^{\delta_c} = 1. \tag{4.8}$$

A PF with $\delta_c > 1$ corresponds to a concave shape and $\delta_c < 1$ corresponds to a convex shape (Fig. 4.3). A very high or very low value of δ_c represents a high curvature.

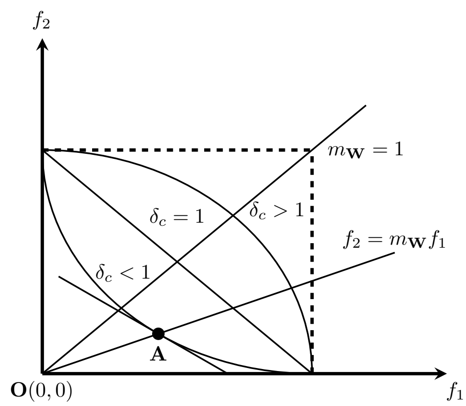


Figure 4.3: Illustration of convex, linear and concave PFs [160].

In terms of $m_{\mathbf{W}}$ and δ_c, the ideal optimal point $\mathbf{A}$ at the intersection of the PF (Eq.

(4.8)) and the reference vector $f_2 = m_{\mathbf{W}} f_1$ is expressed as follows:

$$\text{Ideal Optimal Point } \mathbf{A} : \left(\frac{1}{\left(1 + m_{\mathbf{W}}^{\delta_c}\right)^{\frac{1}{\delta_c}}}, \frac{m_{\mathbf{W}}}{\left(1 + m_{\mathbf{W}}^{\delta_c}\right)^{\frac{1}{\delta_c}}} \right). \tag{4.9}$$

Furthermore, a tangent at point $\mathbf{A}$ is given as follows:

$$\frac{df_2}{df_1} = -\frac{f_1^{\delta_c - 1}}{f_2^{\delta_c - 1}} = -\frac{1}{m_{\mathbf{W}}^{\delta_c - 1}}. \tag{4.10}$$

For $\mathbf{A}$ to be optimal, θ_{pbi} must satisfy the following in a small region around $\mathbf{A}$:

$$\theta_{pbi} > \frac{1}{\tan \phi_p} = \frac{1}{\tan\left(\alpha_p + \beta_p\right)} \implies \theta_{pbi} > \frac{m_{\mathbf{W}}^{\delta_c - 1} - 1}{m_{\mathbf{W}}^{\delta_c + 1} + 1}. \tag{4.11}$$

For all values of $m_{\mathbf{W}}$ and $\delta_c \geq 1$ (linear and concave PF), $\theta_{pbi} > 1$. This condition satisfies the condition for the optimality of $\mathbf{A}$.

However, for convex PF ($\delta_c < 1$), the condition (Eq. (4.11)) may not be satisfied for all values of $m_{\mathbf{W}}$. If the standard value of $\theta_{pbi} = 5$ is used and a convex PF (e.g., with $\delta_c = 0.5$) is assumed, the condition (Eq. (4.11)) is satisfied only till $m_{\mathbf{W}} = 0.03$, i.e., till the reference-vectors make an angle of 1.72° with the closest axes in the objective space. For reference vectors which make smaller angles with the axes, the optimal point cannot be obtained using PBI. Higher the value of θ_{pbi}, lesser the value of the limiting angle between the reference vector and axes. Thus, a higher value of θ_{pbi} increases the extent of the estimated PF for convex shape in a two-dimensional objective space. For larger number of objectives, future studies can be conducted to generalize the above theoretical analysis.

Thus, the PBI function may not always give the complete PF for all MOO problems. It is crucial to find out the effect of the value of θ_{pbi} on the other aspects of MOEAs (such as speed). Without such an analysis, it may not be advisable to increase the value of θ_{pbi} beyond the commonly used value of 5.

4.5 Algorithmic Framework of NAEMO

In this section, the different steps of NAEMO [160], and its underlying features are discussed in detail. The key concepts used in NAEMO are as follows:

1. *Using the neighborhood property:* NAEMO maintains an organized global archive ($\mathcal{A}_G$), which is divided into sub-archives to store the associated solutions for each of the n_{dir} reference-vectors. Candidate association occurs using Eq. (3.2) and the i^{th} sub-archive ($\mathcal{A}^{sub}_{i,G}$) belongs to $\mathcal{A}_G$ as follows:

$$\mathcal{A}_G = \left\{ \mathcal{A}^{sub}_{1,G}, \mathcal{A}^{sub}_{2,G}, \cdots, \mathcal{A}^{sub}_{n_{dir},G} \right\}. \tag{4.12}$$

2. *Periodic filtering of population:* In NAEMO, filtering operations are performed on $\mathcal{A}_G$ to enhance the convergence or diversity. When the size of $\mathcal{A}_G$ exceeds a specific value (l_{soft}), diversity-based filtering is performed to remove points with high PBI values from relatively crowded sub-archives. When a newly generated solution dominates some points in $\mathcal{A}_G$, the convergence-based filtering operation is performed to remove such dominated points while ensuring that the diversity is not hampered.

3. *Monotonic improvement in diversity:* NAEMO ensures that once a reference-vector obtains an associated point, it is never lost. This feature, along with the filtering operations, helps in diversity preservation as proven later in Section 4.5.8.

4. *Using the PBI function along with dominance:* As diversity-based filtering operation removes points having maximal PBI values, it further increases the selection pressure (as discussed in Section 4.4) in addition to the dominance-based selection.

5. *Improved mutation strategy:* NAEMO uses an improved mutation strategy (probabilistic mutation switching) with hyper-parameter adaptation.

4.5.1 Basic Steps of NAEMO

The NAEMO algorithm (Algorithm 4.1) starts with a randomly initialized global archive of size l_{soft}. A single generation (G) consists of iterations through all of the n_{dir} sub-spaces. The set of all reference-vectors is denoted by $\mathcal{W}$ and $|\mathcal{W}| = n_{dir}$.

A random candidate, associated to the j^{th} sub-archive ($\mathcal{A}^{sub}_{j,G}$), is selected as the parent candidate ($\mathbf{X}^{parent}_G$). However, if $\mathcal{A}^{sub}_{j,G}$ is empty, another sub-archive corresponding to a random reference-vector is selected from k_{nbr} non-empty reference-vectors closest to $\mathbf{W}_j \in \mathcal{W}$. The set ($\mathbf{N}_j$) of indices of such neighboring reference-vectors of $\mathbf{W}_j$ is stored in the j^{th} row of the matrix $\mathcal{N}$. The intuition behind this step is that mutation of points from

Algorithm 4.1 Framework of NAEMO [160]

Input: G_{max}: maximum number of generations; n_{dir}: number of reference lines; l_{hard} and l_{soft}: hard and soft limits on archive size; $flag1$, $flag2$, P_{mut} and η_m: parameters for probabilistic mutation switching (for Algorithm 4.2)
Output: $\mathcal{A}_{G_{max}}$: final archive at the end of G_{max} generations

Obtain $\mathcal{W}$ using the approach of Das and Dennis [40] (Section 3.2.1)
Randomly initialize archive $\mathcal{A}_{G=1}$ of size l_{soft} using Eq. (2.1)
Create $\mathcal{A}^{sub}_{i,G}$ by association, using Eq. (3.2) for $\mathbf{X} \in \mathcal{A}_G$ and for $\mathbf{W}_i \in \mathcal{W}$
for $G = 1$ to G_{max} **do**
    $\mathbf{S}_{\eta_c} = \emptyset$
    $\mathbf{S}_{F^{DE}} = \emptyset$
    $\mathbf{S}_{CR} = \emptyset$
    for $j = 1$ to n_{dir} **do**
        $I_{dir} = j$
        if $\mathcal{A}^{sub}_{I_{dir},G} = \emptyset$ **then**
            $I_{nbr} \leftarrow$ Sample a random index from $\mathbf{N}_j$
            $I_{dir} = I_{nbr}$
        end if
        $\mathbf{X}^{parent}_G \leftarrow$ A random candidate from $\mathcal{A}^{sub}_{I_{dir},G}$
        $\eta_c = Gaussian(\mu_{\eta_c}, 5)$
        $F^{DE} = Gaussian(\mu_{F^{DE}}, 0.1)$
        $CR = Gaussian(\mu_{CR}, 0.1)$
        $\mathbf{X}^{child}_G$ = Mutate $\mathbf{X}^{parent}_G$ using Algorithm 4.2
        if $\mathbf{X}^{parent}_G$ does not dominate $\mathbf{X}^{child}_G$ **then**
            Get index l of reference vector $\mathbf{W}_l$ where $\mathbf{X}^{child}_G$ associates (Eq. (3.2))
            $\mathcal{A}^{sub}_{l,G} \leftarrow \mathcal{A}^{sub}_{l,G} \cup \mathbf{X}^{child}_G$
            Convergence-based filtering (Algorithm 4.3) yields filtered $\mathcal{A}_G$
            if $|\mathcal{A}_G| > l_{soft}$ **then**
                Diversity-based filtering (Algorithm 4.4) yields filtered $\mathcal{A}_G$
            end if
            $\mathbf{S}_{\eta_c} = \mathbf{S}_{\eta_c} \cup \eta_c$
            $\mathbf{S}_{F^{DE}} = \mathbf{S}_{F^{DE}} \cup F^{DE}$
            $\mathbf{S}_{CR} = \mathbf{S}_{CR} \cup CR$
        end if
    end for
    $\mu_{\eta_c} = mean(\mathbf{S}_{\eta_c})$
    $\mu_{F^{DE}} = mean(\mathbf{S}_{F^{DE}})$
    $\mu_{CR} = mean(\mathbf{S}_{CR})$
end for
return $\mathcal{A}_{G_{max}}$

neighboring regions of an empty sub-space have a higher probability of generating a new point ($\mathbf{X}^{child}_G$) associated to that sub-space. After obtaining $\mathbf{X}^{parent}_G$, $\mathbf{X}^{child}_G$ is generated by mutating $\mathbf{X}^{parent}_G$ (by Algorithm 4.2). However, by the neighborhood property, the mutation operation is constrained only in the k_{nbr} closest non-empty neighborhood, i.e., the other parent vectors for mutation are selected from $\mathcal{A}^{sub}_{k,G}$ where $k \in \mathbf{N}_j$.

The new candidate solution ($\mathbf{X}^{child}_G$) is selected for addition to the archive only if

$\mathbf{X}_G^{parent}$ does not dominate $\mathbf{X}_G^{child}$. If selected, association operation is performed on the $\mathbf{X}_G^{child}$ to select the sub-archive to which $\mathbf{X}_G^{child}$ will be added. After this inclusion, if there exist points in $\mathcal{A}_G$ which are dominated by $\mathbf{X}_G^{child}$, the convergence-based filtering is executed. If the size of $\mathcal{A}_G$ exceeds the predefined value l_{soft}, diversity-based filtering is performed to reduce the number of points equal to a hard limit (l_{hard}).

4.5.2 Initialization

NAEMO (Algorithm 4.1) starts with initialization of $\mathcal{W}$. The reference-vectors can be placed with a higher density in the region of preference, as decided by the user. In the absence of any preference, $\mathcal{W}$ is initialized using Das and Dennis' approach [40] (as described in Section 3.2.1). NAEMO also initializes $\mathcal{A}_{G=1}$ of size l_{soft} and creates the structured archive, having sub-archives, as mentioned in Eq. (4.12).

4.5.3 Mutation Strategy

Probabilistic mutation switching of NAEMO involves switching between two or more mutation strategies according to a probability assigned for each of the mutation strategies. This switching between mutation techniques often helps in combining the benefits of individual mutation techniques. NAEMO uses Algorithm 4.2 to combine Simulated Binary Crossover (SBX) based mutation [44] and a Differential Evolution (DE) based mutation [41, 164]. These reproduction techniques are described as follows:

1. *SBX mutation:* For a parent solution ($\mathbf{X}_G^{parent}$) and a second parent solution ($\mathbf{X}_G^{par2}$), SBX crossover [44] combines the j^{th} parent decision variables ($x_{j,G}^{parent}$ and $x_{j,G}^{par2}$) to produce the j^{th} variable ($x_{j,G}^{child}$) of the child candidate ($\mathbf{X}_G^{child}$) as follows:

$$x_{j,G}^{child} = 0.5\left[(1+\beta_j)\, x_{j,G}^{parent} + (1-\beta_j)\, x_{j,G}^{par2}\right], \text{ where } j = 1, \cdots, N. \tag{4.13}$$

   Using η_c as the SBX crossover parameter, the parameter β_j is sampled from the following probability distribution:

$$P(\beta) = \begin{cases} 0.5\,(\eta_c+1)\,\beta^{\eta_c}, & \text{if } \beta \le 1 \\ 0.5\,(\eta_c+1)\,\frac{1}{\beta^{(\eta_c+2)}}, & \text{otherwise.} \end{cases} \tag{4.14}$$

2. *DE based mutation:* DE based mutation [41, 146, 162] involves generating the j^{th} decision variable ($x_{j,G}^{child} = x_{ij,G+1}^{trial}$) for the child candidate ($\mathbf{X}_G^{child}$) by using three different candidates from the archive ($\mathbf{X}_{r_1,G}$, $\mathbf{X}_{r_2,G}$ and $\mathbf{X}_{r_3,G}$) and employing the mutation (Eq. (2.2)) and binomial crossover (Eq. (2.3)) operations.

3. *Polynomial mutation:* The polynomial mutation [43] alters the j^{th} decision variable ($x_{j,G}$) of an N-dimensional candidate ($\mathbf{X}_G$) within the upper and lower bounds of the j^{th} decision variable (x_j^U and x_j^L, respectively) to generate the j^{th} decision variable ($x_{j,G}^{child}$) of the child candidate ($\mathbf{X}_G^{child}$) as follows:

$$x_{j,G}^{child} = x_{j,G} + \delta_j \times \left(x_j^U - x_j^L\right), \text{ where } j = 1, \cdots, N. \tag{4.15}$$

   Using η_m as the polynomial mutation parameter, δ_j is sampled from the following distribution:

$$P(\delta) = 0.5\left(\eta_m + 1\right)\left(1 - |\delta|\right)^{\eta_m}. \tag{4.16}$$

Probabilistic mutation switching (Algorithm 4.2) uses P_{mut} as the probability of performing an SBX-based mutation. Therefore, $(1 - P_{mut})$ becomes the probability of performing DE based mutation. The mutation strategy of NAEMO employs two flags, $flag1$ and $flag2$, which determines whether the SBX mutation and DE-based mutation are to be followed by polynomial mutation, respectively. This option is incorporated as the polynomial mutation helps in overcoming local optima [110].

Algorithm 4.2 Mutation Strategy of NAEMO [160]

Input: P_{mut}: mutation switching factor; $flag1$ and $flag2$: determine the use of polynomial mutation; $\mathcal{N}$: set of neighboring non-empty reference-vectors
Output: $\mathbf{X}_G^{child}$: Newly generated point

if $rand(0,1) > P_{mut}$ **then**
  $\mathbf{X}_G^{child} \leftarrow$ DE based mutation (Eq. (2.2)-(2.3))
  if $flag2$ is true **then**
    $\mathbf{X}_G^{child} \leftarrow$ Polynomial mutation (Eq. (4.15)-(4.16))
  end if
else
  $\mathbf{X}_G^{child} \leftarrow$ SBX mutation (Eq. (4.13)-(4.14))
  if $flag1$ is true **then**
    $\mathbf{X}_G^{child} \leftarrow$ Polynomial mutation (Eq. (4.15)-(4.16))
  end if
end if
return $\mathbf{X}_G^{child}$

4.5.4 Parameter Adaptation

The SBX mutation parameter η_c is sampled from the following Gaussian distribution:

$$\eta_c = Gaussian(\mu_{\eta_c}, 5). \tag{4.17}$$

The mean (μ_{η_c}) is initialized to a value of 30 and updated in each generation as the mean over $\mathbf{S}_{\eta_c}$ where $\mathbf{S}_{\eta_c}$ is a set to store all successful values of η_c over a generation G.

Similarly, the parameters F^{DE} and CR of DE-based mutation are also sampled from the following Gaussian distributions:

$$F^{DE} = Gaussian(\mu_{F^{DE}}, 0.1), \tag{4.18}$$

$$CR = Gaussian(\mu_{CR}, 0.1). \tag{4.19}$$

The mean values (μ_F^{DE} and μ_{CR}) are initialized to 0.5 and 0.2, respectively, and updated in each generation as the mean of $\mathbf{S}_{F^{DE}}$ and $\mathbf{S}_{CR}$, respectively.

The sampled values are then truncated to $[0, 1]$. The parameter P_{mut} is experimentally set at 0.75 as adaptation P_{mut} does not lead to any improvements. The parameters $flag1$ and $flag2$ are also not adaptive in this mutation strategy.

4.5.5 Convergence-based Filtering

This operation is performed every time $\mathbf{X}_G^{child}$ gets added. Convergence-based filtering (Algorithm 4.3) looks for all the points in $\mathcal{A}_G$, which are dominated by the $\mathbf{X}_G^{child}$ and removes them. However, it does not remove a dominated point if it is the only point associated with its corresponding reference-vector. This step ensures that no sub-space associated to a reference-vector is rendered empty, thus, preserves diversity. This step is useful for problems which have difficult regions, such as IMB problems [115, 160].

4.5.6 Diversity-based Filtering

Diversity-based filtering (Algorithm 4.4) is performed when the size of $\mathcal{A}_G$ exceeds the soft limit, l_{soft}. At first, the reference-vector with the highest number of associated points is obtained. Then, the point in the associated sub-archive is obtained, which has the highest PBI function value, and it is removed. This process continues until the total size of $\mathcal{A}_G$

Algorithm 4.3 Convergence-based filtering of NAEMO [160]

Input: $\mathcal{A}_G$: unfiltered archive consisting of all the i^{th} sub-archives $\mathcal{A}_{i,G}^{sub}$; n_{dir}: number of reference-vectors or sub-spaces in the objective space; $\mathbf{X}_G^{child}$: newly added point
Output: $\mathcal{A}_G$: filtered archive consisting of all the i^{th} filtered sub-archives $\mathcal{A}_{i,G}^{sub}$

for $i = 1$ to n_{dir} **do**
  $\mathcal{L} \leftarrow$ Set of points in $\mathcal{A}_{i,G}^{sub}$ dominated by $\mathbf{X}_G^{child}$
  for $j = 1$ to $|\mathcal{L}|$ **do**
    if $\left|\mathcal{A}_{i,G}^{sub}\right| > 1$ **then**
      $\mathcal{A}_{i,G}^{sub} \leftarrow \left\{\left(\mathcal{A}_{i,G}^{sub} - \mathbf{X}_j\right) | \mathbf{X}_j \in \mathcal{L}\right\}$
    end if
  end for
end for
return $\mathcal{A}_G$

Algorithm 4.4 Diversity-based filtering of NAEMO [160]

Input: $\mathcal{A}_G$: unfiltered archive consisting of all the i^{th} sub-archives $\mathcal{A}_{i,G}^{sub}$; n_{dir}: number of reference-vectors or sub-spaces in the objective space; $\mathbf{P}^{arr}$: array of length n_{dir} to store sub-archive sizes; l_{hard}: minimum size of archive
Output: : $\mathcal{A}_G$: filtered archive consisting of all the i^{th} filtered sub-archives $\mathcal{A}_{i,G}^{sub}$

$n_{arch} = 0$
$\mathbf{P}^{arr} \leftarrow \left[0, \overset{n_{dir}}{\cdots}, 0\right]$
for $i = 1$ to n_{dir} **do**
  $\mathcal{A}_{i,G}^{sub} \leftarrow$ Sort $\mathcal{A}_{i,G}^{sub}$ by PBI value (Eq. (3.3))
  $S_G^i \leftarrow \left|\mathcal{A}_{i,G}^{sub}\right|$ where S_G^i is the i^{th} element of $\mathbf{P}^{arr}$
  $n_{arch} = n_{arch} + \left|\mathcal{A}_{i,G}^{sub}\right|$
end for
while $n_{arch} > l_{hard}$ **do**
  $I_{ind} \leftarrow$ Index of maximum value from $\mathbf{P}^{arr}$
  Remove last element from $\mathcal{A}_{I_{ind},G}^{sub}$
  $S_G^{I_{ind}} = S_G^{I_{ind}} - 1$
  $n_{arch} = n_{arch} - 1$
end while
return $\mathcal{A}_G$

reduces to the hard limit, l_{hard}. Filtering the points, using the PBI value, creates selection pressure on the points towards the optimal point (Section 4.4).

4.5.7 An Indicator for the Dynamics of Population Diversity [161]

Using the sub-archive sizes ($S_G^k = \left|\mathcal{A}_{k,G}^{sub}\right|$), an indicator ($D\_metric$) is defined to measure the population diversity at a certain generation G. It is given as follows:

$$D\_metric^G = \frac{n_{dir}}{n_{arch}} \sqrt{\sum_{k=1}^{n_{dir}} (S_G^k - S_{ideal}^k)^2}, \text{ where } S_{ideal}^k = \frac{n_{arch}}{n_{dir}}. \tag{4.20}$$

While the ideal sub-archive size (S^k_{ideal}) is chosen to be equal for all reference-vectors [161], it can also be defined as per user's preference. For the best case, $D\_metric = 0$ when all the sub-archives have equal number of associated solutions, denoted by $S^k_{ideal} = n_{arch}/n_{dir}$ for $k = 1$ to n_{dir}.

4.5.8 Proof of Monotonic Improvement of Diversity

In this part, it is mathematically proven that NAEMO with its two filtering operations never leads to deterioration of the diversity.

Theorem 4.3. *Diversity-based filtering operation (Algorithm 4.4) always generates a monotonic improvement of diversity measured using D_metric (Eq.* (4.20)*).*

Proof. In NAEMO, diversity-based filtering reduces the population size (n_{arch}) from l_{soft} to l_{hard}. Thus, n_{arch} is constantly decreasing after each removal of point. This filtering operation finds the sub-archive with the highest S^k_G and removes from it the point with the largest PBI. The *D_metric* after removal of a point from two independent sub-archives, $\mathcal{A}^{l_1}_G$ and $\mathcal{A}^{l_2}_G$, are given as follows:

$$\begin{aligned}\text{Removal from } \mathcal{A}^{l_1}_G\text{: } D\_metric^G_1 &= \frac{n_{dir}}{n_{arch}-1}\left(\sum_{i=1, i\neq l1,l2}^{n_{dir}}\left(S^i_G - \frac{n_{arch}-1}{n_{dir}}\right)^2\right.\\ &\left.+\left(S^{l1}_G - 1 - \frac{n_{arch}-1}{n_{dir}}\right)^2 + \left(S^{l2}_G - \frac{n_{arch}-1}{n_{dir}}\right)^2\right)^{0.5}.\end{aligned} \tag{4.21}$$

$$\begin{aligned}\text{Removal from } \mathcal{A}^{l_2}_G\text{: } D\_metric^G_2 &= \frac{n_{dir}}{n_{arch}-1}\left(\sum_{i=1, i\neq l1,l2}^{n_{dir}}\left(S^i_G - \frac{n_{arch}-1}{n_{dir}}\right)^2\right.\\ &\left.+\left(S^{l2}_G - 1 - \frac{n_{arch}-1}{n_{dir}}\right)^2 + \left(S^{l1}_G - \frac{n_{arch}-1}{n_{dir}}\right)^2\right)^{0.5}.\end{aligned} \tag{4.22}$$

For an improvement in the diversity, the *D_metric* should be minimum, after removal of a point. For $D\_metric^G_1$ to be better than $D\_metric^G_2$, the following must occur:

$$\begin{aligned}& D\_metric^G_1 \leq D\_metric^G_2\\ \implies & \left(S^{l1}_G - 1 - \frac{n_{arch}-1}{n_{dir}}\right)^2 + \left(S^{l2}_G - \frac{n_{arch}-1}{n_{dir}}\right)^2 \leq\\ & \left(S^{l2}_G - 1 - \frac{n_{arch}-1}{n_{dir}}\right)^2 + \left(S^{l1}_G - \frac{n_{arch}-1}{n_{dir}}\right)^2\\ \implies & S^{l1}_G \geq S^{l2}_G.\end{aligned} \tag{4.23}$$

Thus, for removal of a point from $\mathcal{A}_G^{l1}$ to be a better decision than removal from $\mathcal{A}_G^{l2}$, the sub-archive size for $l1$ should be larger than that of $l2$. Since the diversity-based filtering always finds the index I_{ind} of the largest sub-archive from $\mathbf{P}^{arr}$ in line 9 of Algorithm 4.4, it always satisfies the inequality of Eq. (4.23) and thus, leads to the maximum possible decrease in the value of *D_metric*. □

Theorem 4.4. *Convergence-based filtering operation (Algorithm 4.3) preserves diversity in the long run.*

Proof. Convergence-based filtering operation removes all the points that are dominated by $\mathbf{X}_G^{child}$, except for those points removing which might render a sub-space empty. The removal of points by this operation might lead to an increase in *D_metric* as the inequality of Eq. (4.23) might not be met. However, NAEMO aims to obtain one point per reference-vector. Thus, finally $S_{ideal}^i = 1$ for $i = 1$ to n_{dir} as $n_{arch} = l_{hard} = n_{dir}$. Substituting $S_{ideal}^i = 1$ in Eq. (4.20), D_metric^G translates as follows:

$$D\_metric^G = \sqrt{\sum_{i=1}^{n_{dir}} \left(S_G^i - 1\right)^2}. \tag{4.24}$$

The value of this *D_metric* increases (i.e., diversity deteriorates) as more sub-spaces are rendered empty. However, since convergence-based filtering operation never renders a sub-space empty once an associated point is found, it can be claimed that the convergence-based filtering operator preserves diversity. □

4.5.9 Using the Neighborhood Property

NAEMO uses the neighborhood property (Theorem 4.1) in the following ways:

1. When an empty reference-vector is encountered during any iteration, the parent vector is chosen from a neighboring non-empty reference-vector.

2. The selection of parent vectors during reproduction is constrained within the k_{nbr} closest neighboring sub-spaces. This constraint increases the convergence of NAEMO immensely, as shown by the results in Section 4.6. Let the Region of Improvement (RoI) of a reference-vector be that region in decision space which correspond to better solutions than the current best point. This RoI is a subset of the region in decision space corresponding to the reference-vector. Neighboring to this RoI is the RoI of

other reference-vectors, by the neighborhood property (Theorem 4.1). Reproduction, using points in the neighbourhood, thus, has a much higher probability of producing a point in one of the RoIs than reproduction with other random points.

4.5.10 Computational Complexity of NAEMO

The complexity of one generation of NAEMO is computed by considering n_{dir} reference vectors, M number of objectives, l_{soft} as the soft limit and l_{hard} as the hard limit. For NAEMO, $l_{hard} = n_{dir}$ and $l_{soft} = Cn_{dir}$ where C is a real constant such that $C > 1$.

Apart from the values of M and n_{dir}, the time taken by NAEMO depends on how frequently the if conditions in lines 19 and 23 of Algorithm 4.1 are satisfied. The frequency of the if condition in line 23 being satisfied does not affect the complexity as it contributes a constant term as shown later. The if condition in line 19 depends on how frequently the $\mathbf{X}_G^{child}$ dominates $\mathbf{X}_G^{parent}$ and therefore, also does not affect the complexity.

From Algorithm 4.1-line 8, a for loop is observed with n_{dir} iterations. Thus, the complexity for operations inside this for loop considers the following:

- In Algorithm 4.1 - line 20, the association step requires $\mathcal{O}(Mn_{dir})$ operations.
- In Algorithm 4.1 - line 22, convergence-based filtering requires $\mathcal{O}(Ml_{soft}) = \mathcal{O}(Mn_{dir})$ operations.
- Diversity-based filtering, in Algorithm 4.1 - lines 23 and 24, is analyzed as follows. Within the if block,
  - Maximum number of times the if condition (in line 23) is satisfied within n_{dir} iterations is $\frac{n_{dir}}{l_{soft}-l_{hard}} = \frac{1}{C-1}$. Therefore, it is constant.
  - Diversity-based filtering (Algorithm 4.4) requires $\mathcal{O}(l_{soft}\log(l_{soft}))+\mathcal{O}(n_{dir}(l_{soft}-l_{hard})) = \mathcal{O}(n_{dir}\log(n_{dir})) + \mathcal{O}(n_{dir}^2) = \mathcal{O}(n_{dir}^2)$ operations.

Therefore, the total complexity over n_{dir} iterations is given by $n_{dir}(\mathcal{O}(Mn_{dir})+\mathcal{O}(n_{dir}^2)) = \mathcal{O}(Mn_{dir}^2 + n_{dir}^3)$.

The computational burden of several MOEAs is compared in Table 4.1. In the worst case, NAEMO has intermediate time requirements. It is neither the fastest nor the slowest among several other MOEAs. Hence, NAEMO is developed to yield competitive performance at similar time requirements. A comparison of execution time (in seconds) is provided later in Section 4.6.8.

Table 4.1: Worst case computational complexity for a single generation of several MaOEAs considering M as number of objectives and $l_{hard} = n_{dir}$ as the population size (which is nearly equal to the number of reference-vectors) [160].

Algorithm Name	Computational Complexity
NAEMO	$\mathcal{O}(Mn_{dir}^2 + n_{dir}^3)$
NSGA-III [45]	$\mathcal{O}(n_{dir}^2 \log^{M-2} n_{dir} + Mn_{dir}^2)$
θ-DEA [187]	$\mathcal{O}(Mn_{dir}^2)$
GrEA [184]	$\mathcal{O}(n_{dir}^3)$
HypE [9]	$\mathcal{O}(n_{dir}^M + Mn_{dir} \log n_{dir})$

4.6 Experimental Results and Interpretations

NAEMO is compared with other state-of-the-art MOEAs on DTLZ1-DTLZ4 problems from the DTLZ test suite [49] and IMB1-IMB9 problems (with $N = 10$) from IMB test suite [115]. These test-suites are described in Appendix A. For performance analysis, NAEMO is implemented in Python 3.4 and executed in a computer having 8 GB RAM with Intel Core i7 @ 2.5 GHz processor. The source code of NAEMO available at `http://worksupplements.droppages.com/naemo`.

4.6.1 Comparison Metrics

The estimated PFs are assessed in terms of convergence and diversity over the true PF. The most common performance measures are Inverted Generational Distance (IGD) and Hypervolume Indicator (HV) (described in Section 1.3.3).

Decomposition-based MOEAs have an unfair advantage over MOEAs not based on reference-vectors when compared using IGD because the former MOEAs explicitly target those points at the intersection of the reference-vectors and the true PF. These points also constitute the reference set $\mathcal{H}_{IGD}$ for IGD evaluation. Algorithms, not based on reference-vectors, do not target any specific points on the PF and thus, yield poorer IGD.

For comparing MOEAs (based on both decompostion and non-decomposition strategies) with NAEMO, the HV is considered. Using $\mathcal{A}_{\mathbf{F}}$ to denote the estimated PF, if the reference point in the objective space is $\mathbf{R}_{HV} = [r_{HV,1}, r_{HV,2}, \cdots, r_{HV,M}]$, then HV is evaluated as follows:

$$\begin{gathered} HV\left(\mathcal{A}_{\mathbf{F}}, \mathbf{R}_{HV}\right) = volume\left(\cup_{\mathbf{F}\in\mathcal{A}_{\mathbf{F}}} [f_1, r_{HV,1}] \times \cdots \times [f_M, r_{HV,M}]\right), \\ \text{where } \mathbf{F} = [f_1, f_2, \cdots, f_M]. \end{gathered} \tag{4.25}$$

For IMB test problems, $\mathbf{R}_{HV}$ is set at $\mathbf{F}^{nad} + \left[0.001, \overset{M}{\cdots}, 0.001\right]$ as per [116], where $\mathbf{F}^{nad}$ is given by Eq. (1.9). For DTLZ1, $\mathbf{R}_{HV}$ is set as $\left[1, \overset{M}{\cdots}, 1\right]$ and for DTLZ2 - DTLZ4, $\mathbf{R}_{HV}$ is set as $\left[2, \overset{M}{\cdots}, 2\right]$ as per [109]. The HV values for the DTLZ problems are further normalized to $[0, 1]$ by dividing with the total volume of the hyper-rectangle, $\prod_{i=1}^{M} r_{HV,i}$.

4.6.2 Parameter Settings of Algorithms

The parameters for other compared algorithms are set as suggested in [8, 45, 150, 184, 187]. The specifications of these parameters are summarized in Tables 4.2 and 4.3.

Table 4.2: Parameters for various MOEAs for qualitative comparison of NAEMO.

Category	Parameters	Values	Parameters	Values
NAEMO	μ_{η_c}	initialized as 30	k_{nbr}	$0.2 \times n_{dir}$
	$\mu_{F^{DE}}$	initialized as 0.5	l_{hard}	n_{dir}
	μ_{CR}	initialized as 0.2	l_{soft}	Table 4.3 for DTLZ 400 (for 2-objective IMB) 900 (for 3-objective IMB)
	$flag1$	true (only for DTLZ3)		
	$flag2$	true (only for DTLZ1)		
Reproduction Parameters	η_c	30 (for NSGA-III and θ-DEA) 20 (for other MOEAs)	crossover probability	1 (for all MOEAs)
	η_m	20 (for all MOEAs)	mutation probability	$1/N$ (for all MOEAs)
Decomposition Parameters	θ_{pbi}	5 (for NAEMO, θ-DEA, MOEA/D-PBI, MOEA/DD)	neighborhood size	20 (for MOEA/D and MOEA/DD)
Other Parameters	sampling size	10,000 (for HypE)	grid divisions	As per GrEA [184]

Table 4.3: Population size settings for experiments in [160].

No. of objectives (M)	**Divisions to decompose (p_1, p_2) [45]**	**No. of reference vectors (n_{dir})**	**Population size for NSGA-III [45]**	**l_{soft} for NAEMO**
3	12, 0	91	92	100
5	6, 0	210	212	220
8	3, 2	156	156	160
10	3, 2	275	276	280
15	2, 1	135	136	140

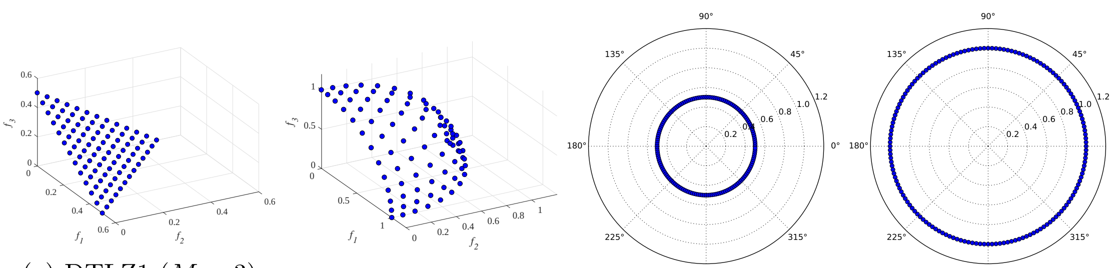
4.6.3 Comparison on DTLZ Problems

As per the specifications in [109, 187], the best, median and worst IGD (Tables 4.4 and 4.5) and HV values (Tables 4.6 and 4.7) of NAEMO are noted for DTLZ problems with $M \in \{3, 5, 8, 10, 15\}$. For establishing the efficacy of NAEMO, the IGD and HV values of other state-of-the-art MOEAs (NSGA-III, MOEA/D, θ-DEA*, MOEA/DD, GrEA and HypE) are mentioned alongside in Tables 4.4, 4.5, 4.6 and 4.7. The maximum number of generations (G_{max}), upto which the MOEAs are executed, are also mentioned in Tables 4.4 and 4.5. These G_{max} values are a standard setting as noted in [45, 109].

Table 4.4: Best, median, worst IGD values over 30 independent runs for comparing MOEAs on M-objective multimodal (DTLZ1 and DTLZ3) problems [160].

Problems	M	G_{max}	NAEMO	NSGA-III	MOEA/D	θ-DEA*	MOEA/DD	GrEA	HypE
DTLZ1			2.725E-5	4.880E-4	4.095E-4	3.006E-4	3.191E-4	2.759E-2	1.822E+1
	3	400	4.801E-5	1.308E-3	1.495E-3	9.511E-4	5.848E-4	3.339E-2	1.974E+1
			1.119E-3	4.880E-3	4.743E-3	2.718E-3	6.573E-4	1.351E-1	2.158E+1
			3.710E-5	5.116E-4	3.179E-4	3.612E-4	2.635E-4	7.369E-2	1.799E+1
	5	600	5.854E-5	9.799E-4	6.372E-4	4.259E-4	2.916E-4	3.363E-1	2.141E+1
			6.529E-5	1.979E-3	1.635E-3	5.797E-4	3.109E-4	4.937E-1	2.359E+1
			4.477E-4	2.044E-3	3.914E-3	1.869E-3	1.809E-3	1.023E-1	1.030E+1
	8	750	6.558E-4	3.979E-3	6.106E-3	2.061E-3	2.589E-3	1.195E-1	2.265E+1
			2.389E-1	8.721E-3	8.537E-3	2.337E-3	2.996E-3	3.849E-1	2.426E+1
			5.022E-4	2.215E-3	3.872E-3	1.999E-3	1.828E-3	1.176E-1	1.427E+1
	10	1000	8.536E-4	3.462E-3	5.073E-3	2.268E-3	2.225E-3	1.586E-1	1.693E+1
			1.762E-3	6.869E-3	6.130E-3	2.425E-3	2.467E-3	5.110E-1	2.034E+1
			1.782E-3	2.649E-3	1.236E-2	2.884E-3	2.867E-3	8.061E-1	1.797E+1
	15	1500	3.587E-3	5.063E-3	1.431E-2	3.504E-3	4.203E-3	2.057E+0	2.519E+1
			4.464E-3	1.123E-2	1.692E-2	3.922E-3	4.699E-3	6.307E+1	2.954E+1
DTLZ3			1.395E-4	9.751E-4	9.773E-4	8.575E-4	5.690E-4	6.770E-2	1.653E+2
	3	1000	1.682E-4	4.007E-3	3.426E-3	3.077E-3	1.892E-3	7.693E-2	1.700E+2
			2.871E-4	6.665E-3	9.113E-3	5.603E-3	6.231E-3	4.474E-1	1.757E+2
			4.173E-4	3.086E-3	1.129E-3	8.738E-4	6.181E-4	5.331E-1	1.826E+2
	5	1000	4.893E-4	5.960E-3	2.213E-3	1.971E-3	1.181E-3	8.295E-1	2.172E+2
			7.944E-4	1.196E-2	6.147E-3	4.340E-3	4.736E-3	1.124E+0	2.278E+2
			2.654E-3	1.244E-2	6.459E-3	6.493E-3	3.411E-3	7.518E-1	2.196E+2
	8	1000	3.476E-3	2.375E-2	1.948E-2	1.036E-2	8.079E-3	1.024E+0	2.700E+2
			5.102E-3	9.649E-2	1.123E+0	1.549E-2	1.826E-2	1.230E+0	2.949E+2
			1.760E-3	8.849E-3	2.791E-3	5.074E-3	1.689E-3	8.656E-1	1.720E+2
	10	1500	1.994E-3	1.188E-2	4.319E-3	6.121E-3	2.164E-3	1.145E+0	2.893E+2
			2.418E-3	2.082E-2	1.010E+0	7.243E-3	3.226E-3	1.265E+0	3.391E+2
			2.226E-3	1.401E-2	4.360E-3	7.892E-3	5.716E-3	9.391E+1	2.358E+2
	15	2000	3.017E-3	2.145E-2	1.664E-2	9.924E-3	7.461E-3	1.983E+2	2.635E+2
			3.640E-3	4.195E-2	1.260E+0	1.434E-2	1.138E-2	3.236E+2	3.451E+2

From Tables 4.4 and 4.5, it can be noted that in only six out of 60 cases, MOEA/DD performs slightly better than NAEMO. Similarly, from Tables 4.4 and 4.5, in only three out of 60 cases, θ-DEA* performs only slightly better than NAEMO. However, in all the remaining cases, NAEMO demonstrates improvement in IGD values, in some cases, even by order of magnitude. This large margin of improvement can be attributed to the efficient use of the neighborhood property (Theorem 4.1). NAEMO also uses the PBI function, which creates a selection pressure on the points towards the optimal point (Section 4.4). Tables 4.6 and 4.7, also show a similar trend in the performance of NAEMO.



(a) DTLZ1 ($M = 3$) (b) DTLZ4 ($M = 3$) (c) DTLZ1 ($M = 8$) (d) DTLZ2 ($M = 8$)

Figure 4.4: Estimated PFs from NAEMO for different types of DTLZ problems [160].

Table 4.5: Best, median, worst IGD values over 30 independent runs for comparing MOEAs on M-objective unimodal (DTLZ2 and DTLZ4) problems [160].

Problems	M	G_{max}	**NAEMO**	**NSGA-III**	**MOEA/D**	**θ-DEA***	**MOEA/DD**	**GrEA**	**HypE**
DTLZ2			2.350E-4	1.262E-3	5.432E-4	7.567E-4	6.666E-4	6.884E-2	6.732E-2
	3	250	3.542E-4	1.357E-3	6.406E-4	9.736E-4	8.073E-4	7.179E-2	6.910E-2
			4.463E-4	2.114E-3	8.006E-4	1.130E-3	1.243E-3	7.444E-2	7.104E-2
			4.589E-4	4.254E-3	1.219E-3	1.863E-3	1.128E-3	1.411E-1	2.761E-1
	5	350	5.895E-4	4.982E-3	1.437E-3	2.146E-3	1.291E-3	1.474E-1	2.868E-1
			7.831E-4	5.862E-3	1.727E-3	2.288E-3	1.424E-3	1.558E-1	2.922E-1
			1.977E-3	1.371E-2	3.097E-3	6.120E-3	2.880E-3	3.453E-1	5.475E-1
	8	500	2.410E-3	1.571E-2	3.763E-3	6.750E-3	3.291E-3	3.731E-1	6.033E-1
			3.053E-3	1.811E-2	5.198E-3	7.781E-3	4.106E-3	4.126E-1	6.467E-1
			1.753E-3	1.350E-2	2.474E-3	6.111E-3	3.223E-3	4.107E-1	6.778E-1
	10	750	2.105E-3	1.528E-2	2.778E-3	6.546E-3	3.752E-3	4.514E-1	6.901E-1
			2.429E-3	1.697E-2	3.235E-3	7.069E-3	4.145E-3	5.161E-1	6.917E-1
			2.209E-3	1.360E-2	5.254E-3	7.269E-3	4.557E-3	5.087E-1	6.237E-1
	15	1000	2.903E-3	1.726E-2	6.005E-3	8.264E-3	5.863E-3	5.289E-1	8.643E-1
			4.019E-3	2.114E-2	9.409E-3	9.137E-3	6.929E-3	5.381E-1	3.195E+0
DTLZ4			4.209E-5	2.915E-4	2.929E-1	1.408E-4	1.025E-4	6.869E-2	6.657E-2
	3	600	5.963E-5	5.970E-4	4.280E-1	1.918E-4	1.429E-4	7.234E-2	7.069E-2
			1.320E-4	4.286E-1	5.234E-1	5.321E-1	1.881E-4	9.400E-1	5.270E-1
			3.859E-5	9.849E-4	1.080E-1	2.780E-4	1.097E-4	1.422E-1	2.603E-1
	5	1000	5.285E-5	1.255E-3	5.787E-1	3.142E-4	1.296E-4	1.462E-1	2.676E-1
			7.452E-5	1.721E-3	7.348E-1	3.586E-4	1.532E-4	1.609E-1	5.301E-1
			6.595E-4	5.079E-3	5.298E-1	2.323E-3	5.271E-4	3.229E-1	4.792E-1
	8	1250	7.619E-4	7.054E-3	8.816E-1	3.172E-3	6.699E-4	3.314E-1	4.956E-1
			1.208E-3	6.051E-1	9.723E-1	3.635E-3	9.107E-4	3.402E-1	5.387E-1
			8.560E-4	5.694E-3	3.966E-1	2.715E-3	1.291E-3	4.191E-1	6.760E-1
	10	2000	1.025E-3	6.337E-3	9.203E-1	3.216E-3	1.615E-3	4.294E-1	6.828E-1
			1.189E-3	1.076E-1	1.077E+0	3.711E-3	1.931E-3	4.410E-1	6.877E-1
			9.607E-4	7.110E-3	5.890E-1	4.182E-3	1.474E-3	4.975E-1	5.986E-1
	15	3000	1.496E-3	3.431E-1	1.133E+0	5.633E-3	1.881E-3	5.032E-1	6.102E-1
			2.788E-3	1.073E+0	1.249E+0	6.562E-3	3.159E-3	5.136E-1	6.126E-1

The resulting archive ($\mathcal{A}_{\mathbf{F},G_{max}}$) in the objective space represents the estimated PF. It is visualized in Fig. 4.4 using Cartesian plots for 3-objective problems and polar plots [68] for higher-objective problems (Appendix B). From Fig. 4.4, these estimated PFs from NAEMO are noted to be similar to the true PF for DTLZ problems (Section A.1 and Fig. B.2). These observations establish the proficiency of NAEMO for DTLZ1-4 problems.

4.6.4 Comparison on IMB Problems

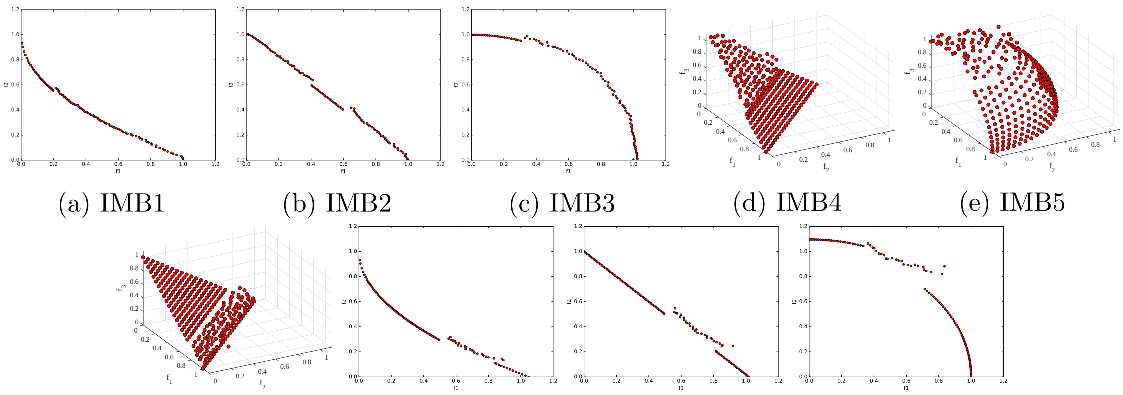
To establish the efficacy of NAEMO over the IMB problems, the best, mean and worst HV values of NAEMO are compared in Table 4.8 with those of the M2M-based MOEAs, according to the specifications in [115]. For M2M-based MOEAs, $G_{max} = 2000$, $n_{dir} = S_G^k = 10$ (when $M = 2$) or $n_{dir} = 30$ with $S_G^k = 10$ (when $M = 3$) are set as per [115] while for NAEMO, $n_{dir} = 100$ (when $M = 2$) or $n_{dir} = 276$ (when $M = 3$) are set. The respective HV values of ESOEA/DE [138] are also noted in Table 4.8 where the specifications from Table 3.2 are considered.

NAEMO outperforms the M2M-based MOEAs for IMB1, IMB2, IMB4 and IMB6

Table 4.6: Best, median, worst HV values over 30 independent runs for comparing MOEAs on M-objective multimodal (DTLZ1 and DTLZ3) problems [160].

Problems	M	NAEMO	NSGA-III	MOEA/D	MOEA/DD	GrEA	HypE
DTLZ1	3	0.973668	0.973519	0.973541	0.973597	0.967404	0.000000
		0.973668	0.973217	0.973380	0.973510	0.964059	0.000000
		0.973668	0.971931	0.972484	0.973278	0.828008	0.000000
	5	0.999897	0.998971	0.998978	0.998980	0.991451	0.000000
		0.999897	0.998963	0.998969	0.998975	0.844529	0.000000
		0.999897	0.998673	0.998954	0.998968	0.500179	0.000000
	8	0.999979	0.999975	0.999943	0.999949	0.999144	0.000000
		0.999979	0.993549	0.999866	0.999919	0.997992	0.000000
		0.994781	0.966432	0.999549	0.999887	0.902697	0.000000
	10	0.999999	0.999991	0.999983	0.999994	0.999451	0.000000
		0.999999	0.999985	0.999979	0.999990	0.998587	0.000000
		0.999978	0.999969	0.999956	0.999974	0.532348	0.000000
DTLZ3	3	0.926512	0.926480	0.926598	0.926617	0.924652	0.000000
		0.926411	0.925805	0.925855	0.926346	0.922650	0.000000
		0.925641	0.924234	0.923858	0.924901	0.621155	0.000000
	5	0.990532	0.990453	0.990543	0.990558	0.963021	0.000000
		0.990532	0.990344	0.990444	0.990515	0.808084	0.000000
		0.990428	0.989510	0.990258	0.990349	0.499908	0.000000
	8	0.999327	0.999300	0.999328	0.999343	0.953478	0.000000
		0.999325	0.924059	0.999303	0.999311	0.791184	0.000000
		0.999324	0.904182	0.508355	0.999248	0.498580	0.000000
	10	0.999923	0.999921	0.999922	0.999923	0.962168	0.000000
		0.999921	0.999918	0.999920	0.999922	0.735934	0.000000
		0.999921	0.999910	0.999915	0.999921	0.499676	0.000000

problems (with imbalanced difficulty). In cases of IMB7-IMB9 problems (with variable linkage difficulty), the performance of NAEMO is relatively poor. Even so, its HV values are quite comparable. While ESOEA/DE outperforms NAEMO in some test cases, it requires a larger global population size [138]. Hence, NAEMO uniformly explores the IMB problems to generate a well-diverse PF (Fig. 4.5). Thus, NAEMO is competitive to the state-of-the-art MOEAs for addressing the IMB problems.



(a) IMB1 (b) IMB2 (c) IMB3 (d) IMB4 (e) IMB5

(f) IMB6 (g) IMB7 (h) IMB8 (i) IMB9

Figure 4.5: Estimated PFs from NAEMO for IMB test problems [160].

Table 4.7: Best, median, worst HV values over 30 independent runs for comparing MOEAs on M-objective unimodal (DTLZ2 and DTLZ4) problems [160].

Problems	M	NAEMO	NSGA-III	MOEA/D	MOEA/DD	GrEA	HypE
DTLZ2	3	0.926683	0.926626	0.926666	0.926674	0.924246	0.925691
		0.926662	0.926536	0.926639	0.926653	0.923994	0.925650
		0.926651	0.926395	0.926613	0.926596	0.923675	0.925531
	5	0.990535	0.990459	0.990529	0.990535	0.990359	0.987889
		0.990535	0.990400	0.990518	0.990527	0.990214	0.987665
		0.990521	0.990328	0.990511	0.990512	0.990064	0.987545
	8	0.999352	0.999320	0.999341	0.999346	0.999991	0.997401
		0.999340	0.978936	0.999329	0.999337	0.999670	0.996551
		0.999329	0.919680	0.999307	0.999329	0.989264	0.995761
	10	0.999923	0.999918	0.999922	0.999952	0.997636	0.998995
		0.999923	0.999916	0.999921	0.999932	0.996428	0.998934
		0.999921	0.999915	0.999919	0.999921	0.994729	0.998913
DTLZ4	3	0.926733	0.926659	0.926729	0.926731	0.924613	0.926351
		0.926733	0.926705	0.926725	0.926729	0.924094	0.926223
		0.926652	0.799572	0.500000	0.926725	0.500000	0.800459
	5	0.990581	0.991102	0.990569	0.990575	0.990514	0.988150
		0.990569	0.990413	0.990568	0.990573	0.990409	0.988009
		0.990431	0.990156	0.973811	0.990570	0.990221	0.987743
	8	0.999382	0.999363	0.999363	0.999364	0.999102	0.997994
		0.999371	0.999361	0.998497	0.999363	0.999039	0.997730
		0.999327	0.994784	0.995753	0.998360	0.998955	0.997569
	10	0.999921	0.999915	0.999918	0.999921	0.999653	0.999019
		0.999921	0.999910	0.999907	0.999920	0.999608	0.998934
		0.999921	0.999827	0.999472	0.999917	0.999547	0.998921

4.6.5 Analyzing the Mutation Switching Scheme

To assess the efficacy of mutation switching scheme of NAEMO, instead of Algorithm 4.2, SBX crossover (Eq. (4.13)) followed by polynomial mutation (Eq. (4.15)) are considered in line 18 of Algorithm 4.1. The IGD value from this altered framework (NAEMO-SBX) is noted in Table 4.9 for 10-objective DTLZ1-4 problems. The values in Table 4.9 show that NAEMO is more robust than NAEMO-SBX. NAEMO's mutation switching (Algorithm 4.2) is a generic framework where other reproduction strategies could also be integrated to combine their advantages for covering a wider range of problem characteristics.

4.6.6 Diversity Plots

A comparison of the *D_metric* plots of NAEMO is presented for DTLZ1 and DTLZ3 problems (multi-modal) in Fig. 4.6. The multi-modal problems can cause changes in diversity while overcoming local optima [161]. The *D_metric* plots for NAEMO are monotonically decreasing as proven in Section 4.5.8. Not only NAEMO attains the ideal *D_metric* much faster than the other MOEAs, but also the diversity does not monotonically improve for

Table 4.8: Best, mean, worst HV values over 30 independent runs for comparing NAEMO with M2M-based MOEAs on IMB problems [160].

Problems	NAEMO	ESOEA/DE	NSGA-II -M2M	MOEA/D -M2M	SMS-EMOA -M2M	SPEA2 -M2M	GVEGA -M2M
IMB1	0.6477	0.6712	0.6375	0.6387	0.6402	0.6384	0.6475
	0.6441	0.6640	0.6360	0.6375	0.6386	0.6372	0.6408
	0.6399	0.6569	0.6353	0.6354	0.6363	0.6351	0.5969
IMB2	0.4734	0.4902	0.4605	0.4627	0.4639	0.4605	0.4750
	0.4710	0.4838	0.4577	0.4608	0.4592	0.4564	0.4509
	0.4657	0.4700	0.4537	0.4583	0.4411	0.4487	0.4224
IMB3	0.1801	0.1923	0.1828	0.1851	0.1845	0.1824	0.1964
	0.1745	0.1838	0.1815	0.1836	0.1834	0.1802	0.1950
	0.1639	0.1728	0.1801	0.1824	0.1819	0.1783	0.1914
IMB4	0.7886	0.7716	0.7445	0.7803	0.7792	0.7476	0.7798
	0.7812	0.7599	0.7424	0.7795	0.7786	0.7421	0.7790
	0.7531	0.7488	0.7398	0.7785	0.7783	0.7364	0.7784
IMB5	0.4117	0.4306	0.3874	0.4266	0.4169	0.3973	0.4215
	0.4026	0.4247	0.3842	0.4229	0.4140	0.3906	0.4209
	0.3974	0.4205	0.3802	0.4202	0.4119	0.3832	0.4205
IMB6	0.8046	0.7998	0.7700	0.7916	0.7859	0.7814	0.7837
	0.7996	0.7921	0.7686	0.7909	0.7856	0.7807	0.7833
	0.7961	0.7843	0.7675	0.7904	0.7853	0.7801	0.7828
IMB7	0.6501	0.6682	0.6499	0.6545	0.6559	0.6515	0.6540
	0.6471	0.6559	0.6482	0.6540	0.6550	0.6505	0.6537
	0.6443	0.6415	0.6464	0.6534	0.6542	0.6494	0.6531
IMB8	0.4777	0.4885	0.4798	0.4840	0.4852	0.4811	0.4863
	0.4703	0.4770	0.4774	0.4830	0.4835	0.4795	0.4857
	0.4519	0.4676	0.4756	0.4820	0.4820	0.4768	0.4848
IMB9	0.1836	0.1989	0.1925	0.1975	0.1974	0.1930	0.2011
	0.1777	0.1951	0.1912	0.1960	0.1961	0.1919	0.2005
	0.1719	0.1911	0.1896	0.1946	0.1947	0.1911	0.1998

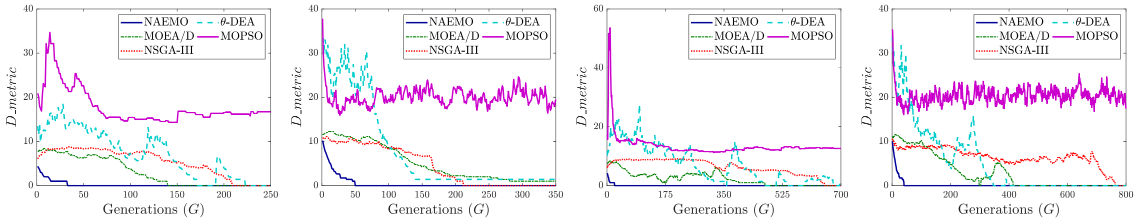
Table 4.9: Best, median, worst IGD values over 30 runs for demonstrating the effectiveness of NAEMO's mutation switching scheme on 10-objective DTLZ problems [160].

Problems	NAEMO-SBX	NAEMO	NSGA-III	MOEA/D	θ-DEA*	MOEA/DD	GrEA	HypE
DTLZ1	1.787E-3	5.022E-4	2.215E-3	3.872E-3	1.999E-3	1.828E-3	1.176E-1	1.427E+1
	2.668E-3	8.536E-4	3.462E-3	5.073E-3	2.268E-3	2.225E-3	1.586E-1	1.693E+1
	2.751E-3	1.762E-3	6.869E-3	6.130E-3	2.425E-3	2.467E-3	5.110E-1	2.034E+1
DTLZ2	1.827E-3	1.753E-3	1.350E-2	2.474E-3	6.111E-3	3.223E-3	4.107E-1	6.778E-1
	1.980E-3	2.105E-3	1.528E-2	2.778E-3	6.546E-3	3.752E-3	4.514E-1	6.901E-1
	2.369E-3	2.429E-3	1.697E-2	3.235E-3	7.069E-3	4.145E-3	5.161E-1	6.917E-1
DTLZ3	4.414E-3	1.760E-3	8.849E-3	2.791E-3	5.074E-3	1.689E-3	8.656E-1	1.720E+2
	1.842E-2	1.994E-3	1.188E-2	4.319E-3	6.121E-3	2.164E-3	1.145E+0	2.893E+2
	2.190E-2	2.418E-3	2.082E-2	1.010E+0	7.243E-3	3.226E-3	1.265E+0	3.391E+2
DTLZ4	9.227E-4	8.560E-4	5.694E-3	3.966E-1	2.715E-3	1.291E-3	4.191E-1	6.760E-1
	1.016E-3	1.025E-3	6.337E-3	9.203E-1	3.216E-3	1.615E-3	4.294E-1	6.828E-1
	1.042E-3	1.189E-3	1.076E-1	1.077E+0	3.711E-3	1.931E-3	4.410E-1	6.877E-1

any other MOEAs. This huge difference in the *D_metric* convergence is owing to the effective utilization of the neighborhood property (Theorem 4.1) in NAEMO.

4.6.7 Decomposition of Objective Space versus Objective Reduction

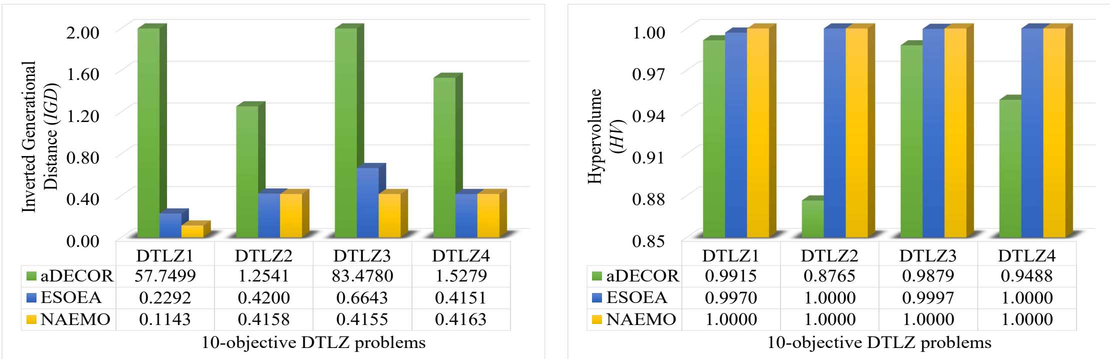
It is observed from Tables 3.6 and 3.7 that aDECOR (an objective reduction based MOEA) has superior convergence but poor diversity whereas ESOEA/DE (a reference vector assisted decomposition-based MOEA) improves the diversity of solutions over the estimated



(a) DTLZ1 ($M = 3$) (b) DTLZ1 ($M = 8$) (c) DTLZ3 ($M = 3$) (d) DTLZ3 ($M = 8$)

Figure 4.6: *D_metric* plots showing faster diversity attainment rate of NAEMO [160].

PF. Thus, the performance of NAEMO is compared with aDECOR and ESOEA/DE on DTLZ1-4 problems ($M = 10$) in Fig. 4.7 using the comparison framework of [83]. For HV evaluation, $\mathbf{R}_{HV} = \left[3, \overset{M}{\cdots}, 3\right]$ and $|\mathcal{H}_{HV}| = 10,000$ is considered as per [138,142]. For IGD evaluation, $|\mathcal{H}_{IGD}| = 5000$ ($\neq n_{dir}$) points are uniformly sampled from the true PF.



(a) Using IGD (b) Using HV

Figure 4.7: Mean HV and IGD values over 30 independent runs to compare decomposition-based MOEAs (ESOEA/DE and NAEMO) with objective reduction based MOEA (aDECOR) on 10-objective DTLZ1-4 problems where for better scaling, the maximum limit on y-axis of IGD is considered as 2.

It is seen from Fig. 4.7 that the reference-vector assisted decomposition-based MOEAs (ESOEA and NAEMO) have largely outperformed the objective reduction based MOEA (aDECOR) in all the cases. Although the difference in performance between NAEMO and ESOEA is small, NAEMO is superior for most of these test cases as DTLZ1-4 problems have regular PFs. Thus, this experiment establishes the superiority of NAEMO for problem characteristics similar to DTLZ1-4 problems.

4.6.8 Miscellaneous Experiments

These experiments compare NAEMO with variants of Multi-Objective Particle Swarm Optimization (MOPSO) [30,36,121], analyze its Purity metric [10,11], study its effect on scaled and disconnected PF and compare its computational time requirements.

1) **Comparison of NAEMO with MOPSO variants**: NAEMO is compared with MOPSO [36] and dMOPSO [121] using mean HV values in Table 4.10, as per the specifications in [187]. For HV evaluation, the estimated PF is normalized between $\mathbf{F}^{nad}$ and $\mathbf{F}^{ide}$, and $\mathbf{R}_{HV}$ is set at $1.1\mathbf{F}^{nad}$. NAEMO outperforms both MOPSO and dMOPSO in Table 4.10. A zero HV of MOPSO implies that its estimated PF is completely outside the hyper-rectangle used for calculating the HV.

Table 4.10: Mean HV over 30 runs for comparing NAEMO with MOPSO variants [160].

MaOEAs		Number of objectives (M) 3	5	8	10		Number of objectives (M) 3	5	8	10
NAEMO	DTLZ1	1.304662	1.609497	2.143047	2.593741	DTLZ2	0.744830	1.308778	1.980806	2.515441
MOPSO		0	0	0	0		0.638144	0.510065	0.060562	0.082047
dMOPSO		1.074976	1.482412	1.824428	2.317805		0.712523	1.239853	1.816420	2.428399
NAEMO	DTLZ3	0.744840	1.308723	1.980405	2.515377	DTLZ4	0.744848	1.308761	1.980838	2.515418
MOPSO		0	0	0	0		0	0	0	0
dMOPSO		0.665529	1.252229	1.428208	2.107556		0.677459	1.203429	1.829561	2.438748

2) **Performance of NAEMO based on Purity Metric**: The purity metric [10,11] compares two or more approximations of PF as described in Section 1.3.3. NAEMO is observed to be superior when compared to other MOEAs using purity metric in Table 4.11. Also, for DTLZ4, NAEMO, NSGA-III and θ-DEA have much higher purity values than HyPE and MOPSO. This result shows the necessity of decomposition-based MOEAs for problems with a biased solution density.

Table 4.11: Mean purity values over 30 independent runs for comparing MOEAs [160].

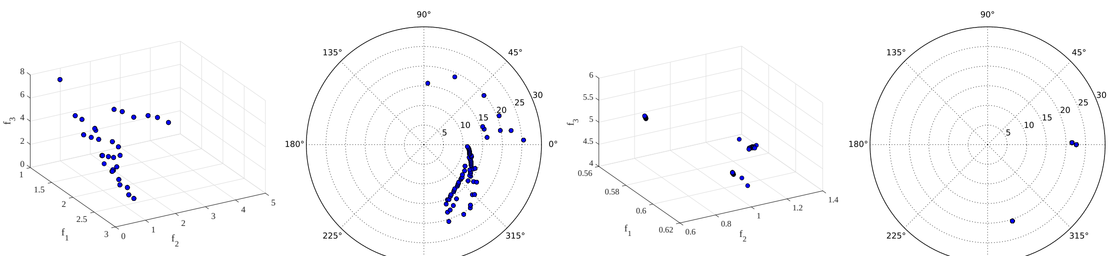
MaOEAs		Number of objectives (M) 3	5	8	10		Number of objectives (M) 3	5	8	10
NAEMO	DTLZ1	1.000000	1.000000	1.000000	1.000000	DTLZ2	1.000000	1.000000	1.000000	1.000000
HypE		0.000000	0.004762	0.012821	0.537879		0.318681	0.195238	0.339744	0.647273
MOPSO		0.017241	0.066667	0.326923	0.742424		0.406593	0.404762	0.551282	0.469091
NSGA-III		0.431034	0.633333	0.858974	0.946970		0.714286	0.638095	0.570513	0.512727
θ-DEA		0.965517	0.795238	0.980769	0.992424		0.637363	0.780952	0.750000	0.730909
NAEMO	DTLZ3	1.000000	1.000000	1.000000	1.000000	DTLZ4	1.000000	1.000000	1.000000	1.000000
HypE		0.092105	0.066667	0.500000	0.450909		0.844444	0.766667	0.224359	0.221818
MOPSO		0.026316	0.190476	0.282051	0.269091		0.077778	0.333333	0.538462	0.676364
NSGA-III		0.236842	0.509524	0.416667	0.505455		0.966667	0.990476	1.000000	1.000000
θ-DEA		0.421053	0.247619	0.634615	0.578182		0.977778	0.990476	0.993590	1.000000

3) **Weaknesses of NAEMO - Scaled and Disconnected Pareto-Fronts**: WFG1 and WFG2 problems are considered as examples of problems with scaled and disconnected PF (Section A.2). For WFG2 problem with even M, N is set to 23 and for all other cases, N is set to 24 as per [138, 187]. For these problems, the mean HV values of NAEMO is noted in Table 4.12, as done in [187]. For HV evaluation, the estimated PF is normalized between $\mathbf{F}^{nad}$ and $\mathbf{F}^{ide}$, and $\mathbf{R}_{HV}$ is set at $1.1\mathbf{F}^{nad}$. NAEMO has worst performance

in Table 4.12 as it does not have any explicit scaling mechanism integrated with the framework and it retains a solution in every sub-space if ever associated. For WFG1 (Fig. 4.8a, 4.8b) still a considerable part of the PF is estimated by NAEMO as opposed to the poor convergence for WFG2 (Fig. 4.8c, 4.8d).

Table 4.12: Mean HV values over 30 independent runs for comparing NAEMO with other MOEAs on WFG1 and WFG2 problems [160].

Problems	M	G_{max}	**NAEMO**	**NSGA-III**	**MOEA/D**	**GrEA**	**HypE**	**dMOPSO**
WFG1	3	400	0.387079	0.669729	0.657143	0.846287	0.976181	0.403170
	5	750	0.418739	0.859552	1.349888	1.268898	0.911020	0.461233
	8	1500	0.504653	1.424963	1.755326	1.769013	1.536599	0.484046
	10	2000	0.553354	2.249535	1.799394	2.365107	2.268813	0.536340
WFG2	3	400	0.289046	1.226956	1.111085	1.226099	1.244737	1.125810
	5	750	0.322285	1.598410	1.520168	1.570086	1.535704	1.478517
	8	1500	0.390006	2.136525	2.016854	2.102930	2.084336	1.971067
	10	2000	0.445272	2.588104	2.459026	2.570389	2.556327	2.406484



(a) WFG1 ($M = 3$) (b) WFG1 ($M = 10$) (c) WFG2 ($M = 3$) (d) WFG2 ($M = 10$)

Figure 4.8: Estimated PFs from NAEMO for WFG1 and WFG2 problems [160].

4) **Computational Time Requirement**: On the same platform, the average execution time of NAEMO [160] is compared with that of NSGA-III [45] for several cases of DTLZ problems (Fig. 4.9). NAEMO requires lesser execution time than NSGA-III as the main computation-intensive parts of NAEMO get initiated only when the if condition in line 19 of Algorithm 4.1 is satisfied. Although the execution time should increase with an increase in M, yet the MOEAs need more time for problems with $M = 5$ than for problems with $M = 8$. This requirement is because n_{dir} (and associatively, the number of candidates) is smaller when $M = 8$ than when $M = 5$ (Table 4.3) by Das and Dennis' approach of reference-vector initialization (Section 3.2.1).

All these experiments demonstrate the overall efficacy of NAEMO to tackle many-objective optimization problems (from DTLZ and IMB test suites) with several characteristics like unimodality, multi-modality, a biased density of solutions, meta-variable mapping, imbalance mapping difficulty and variable linkage difficulty.

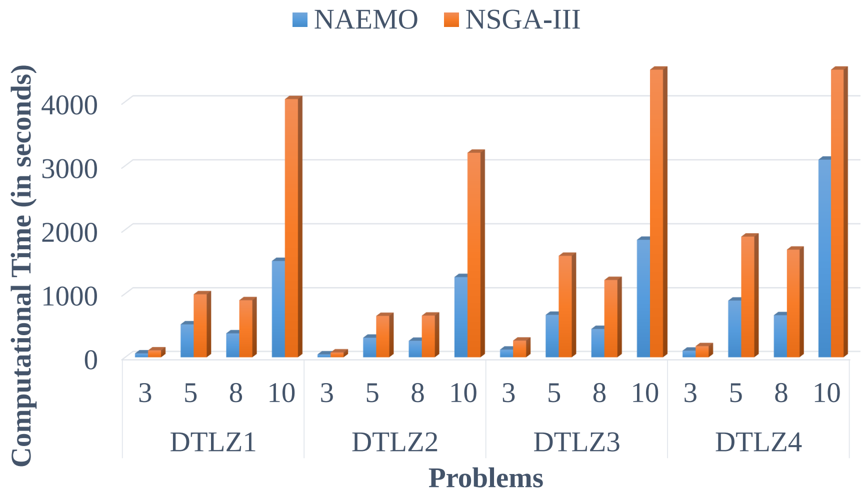


Figure 4.9: Computational time requirements of NAEMO and NSGA-III where for better scaling, the maximum limit along time-axis is 4500 seconds [160].

4.7 Conclusion

Motivated by the success of the decomposition-based MOEAs and the necessity of theoretical analyses to understand the working of such MOEAs, this chapter discusses the algorithmic framework of NAEMO where the neighborhood property of the MaOO problems is identified and used for selecting the mating candidate solutions for the generation of new candidate solutions. Moreover, NAEMO aims to preserve and monotonically improve the diversity through periodic filtering of the archive where if a candidate solution ever gets associated with a reference-vector, it is never lost along the evolutionary process. The robust performance of NAEMO to tackle MaOO problems with several characteristics like unimodality, multi-modality, biased solution density, meta-variable mapping, imbalance mapping difficulty and variable linkage difficulty, has been demonstrated through experiments on problems from DTLZ and IMB test suite. Results indicate that NAEMO outperforms several contemporary state-of-the-art MOEAs on these test problems.

While the usual algorithmic designs of MOEAs (Section 1.3.1), including those of DECOR [142], ESOEA [138] and NAEMO [160], deal with the solution distribution in the objective space, it is essential to analyze the solution distribution in the decision space as well. Such an analysis forms the basis of developing algorithms for Multi-Modal Multi-Objective Problems (MMMOPs) [171] and are extremely important from the practical perspective of decision-making. Hence, algorithms for MMMOPs are considered in the next chapter.

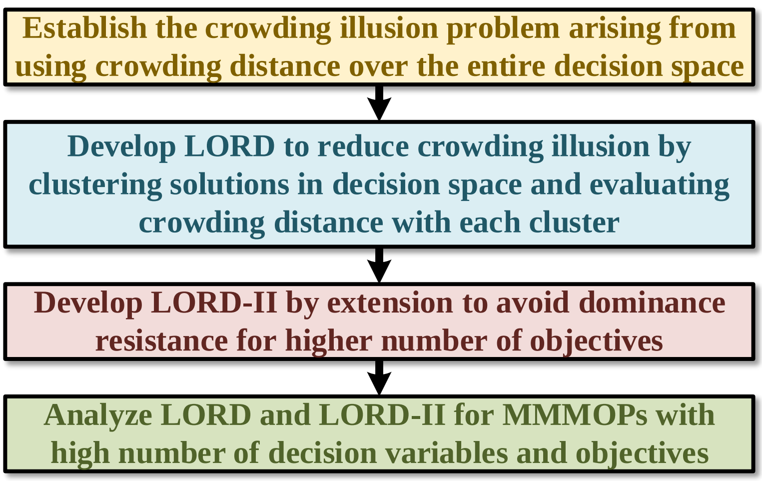
Chapter 5

Decomposition in Decision and Objective Space for Multi-Modal Multi-Objective Optimization [140]

Outline

Objective: *To develop a generic algorithm using reference-vector assisted decomposition of objective space and spectral clustering in the decision space for addressing many-objective optimization problems (including multi-modal problems).*

Workflow:



5.1 Introduction

Multi-Modal Multi-Objective Problem (MMMOP) [171] maps a set of k_{PS} (≥ 2) distinct decision vectors ($\mathcal{A}_M = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_{k_{PS}}\}$) to *almost same* objective vectors (formally given by Eq. (1.17), illustrated in Fig. 1.4). By the neighborhood property (Theorem 4.1), MMMOPs have k_{PS} partitions in the decision space. Thus, the Pareto-optimal Set (PS) of MMMOPs consists of multiple subsets, where each subset can independently generate the identical regions of the Pareto-Front (PF). However, it is observed from the previous chapters that the standard Multi-Objective Evolutionary Algorithms (MOEAs) [32, 127] (Section 1.3.1) focus mainly on the objective space and overlook the solution distribution in the decision space. Thus, MMMOPs are difficult for such MOEAs.

Research on MMMOPs is motivated to discover those k_{PS} alternative solutions for nearly the same objective values such that the non-numeric, domain-specific attributes of these solutions can be analyzed and compared during decision-making. Moreover, when the practical implementation of a solution is hindered, a nearly equivalent alternative can be beneficial. Such MMMOPs are seen in rocket engine design [103], feature selection problem [189] and path-planning problem [90].

In contrast to standard MOEAs, MOEAs for MMMOPs have improved diversity in the decision space but poor performance in the objective space [56, 113, 120, 188]. To explore this gap, this chapter explains the drawback of using crowding distance in the decision space when solving MMMOPs. Subsequently, graph Laplacian based Optimization using Reference-vector assisted Decomposition (LORD) is presented, which uses decomposition in both objective and decision space for dealing with MMMOPs. Its filtering step is further extended to present LORD-II algorithm, which demonstrates its dynamics on Multi-Modal Many-Objective Problems (MMMaOPs). The performance of these frameworks are compared on 34 test instances (obtained from the CEC 2019 test suite for MMMOPs [112]) with the state-of-the-art MOEAs for MMMOPs, Multi-Objective Optimization (MOO) or Many-Objective Optimization (MaOO) problems.

The rest of the chapter is organized as follows: Section 5.2 presents the related studies on the exploring the decision space, Section 5.3 explains the issue of directly using crowding distance, Section 5.4 outlines the frameworks of LORD and LORD-II, Section 5.5 investigates their performance and Section 5.6 concludes this chapter with a summary.

5.2 Related Studies on Manipulation of Solution Distribution in the Decision Space

Omni-optimizer [51] is the earliest work to consider the solution diversity in the decision space[1]. It uses of crowding distance in the decision space (CDX) after the non-dominated sorting [51] but hampers the solution diversity in the objective space. The work in [21] uses neighborhood count and Lebesgue contribution to promote solution diversity in the decision and objective spaces, respectively. The work in [198] considers CDX and a probabilistic model to estimate PS and PF but performs poorly when PS is a linear manifold.

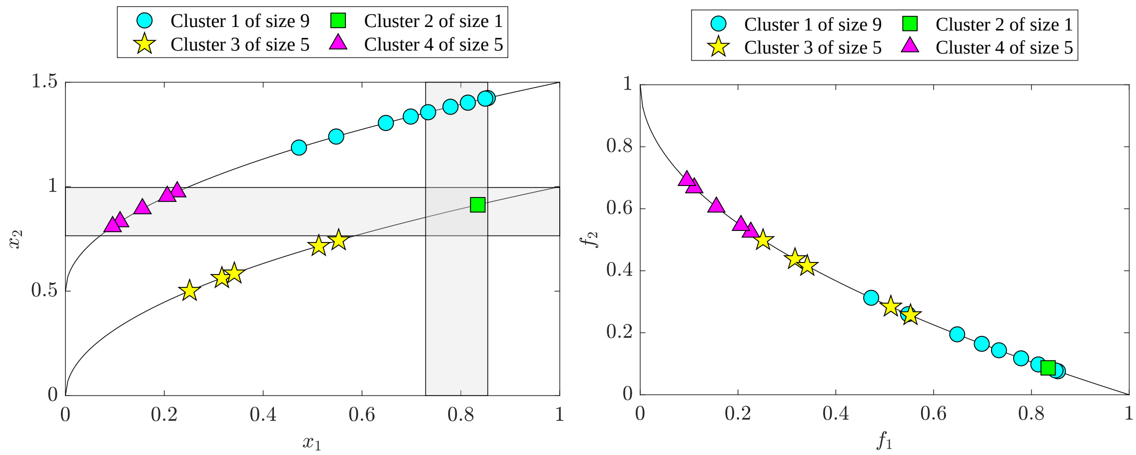
Extensive research on Multi-Modal Multi-Objective Evolutionary Algorithms (MMMOEAs) started with Decision-Niched NSGA-II (DN-NSGA-II) [113], which replaces the crowding distance in the objective space (CDF) with CDX in NSGA-II. Another MMMOEA combines NSGA-II with Weighted Sum Crowding Distance and Neighbor-hood Based Mutation (NSGA-II-WSCD-NBM) [91]. Unlike these preliminary MMMOEAs, MO_Ring_PSO_SCD [188] demonstrates that diversity preservation and niching methods (like ring topology) play vital roles for several MMMOPs. Although computationally expensive, Zoning Search (ZS) [56] further enhances its diversity in the decision space. MOEA/D with Addition and Deletion operators (MOEA/D-AD) [169] introduces the notion of *almost same* Pareto-optimal solutions. Multi-Modal Multi-Objective Evolutionary Algorithm with Two Archive and Recombination (TriMOEA_TA&R) [118] benefits those MMMOPs where a subspace can be extracted from the convergence-related decision variables [118]. Two recent studies: Differential Evolution for MMMOPs (DE-TriM) [137] and Multi-Modal NAEMO (MM-NAEMO) [120] use reference-vector assisted decomposition of objective space and adaptive reproduction strategies. However, these MMMOEAs have inferior performance in the objective space as compared to the standard MOEAs. Earlier in 2019, a Niching Indicator based Multi-Modal many-objective Optimizer (NIMMO) [170] demonstrated its performance on a few MMMaOPs. However, NIMMO [170] investigated its performance only on MMMaOPs [76] with 2-dimensional decision space.

Thus, several MMMOEAs [56, 113, 120, 188] exhibit poor performance in the objective space and have only been tested on non-scalable problems, which motivate the design of better MMMOEAs for problems with high numbers of variables (N) and objectives (M).

[1]In this thesis, decision space, variable space and solution space are considered as synonymous.

5.3 The Crowding Illusion Problem

Most MMMOEAs [51, 91, 113, 137, 188, 198] use CDX to assess the solution distribution. However, using CDX over the entire decision space can be illusional. To describe the problem, let the example in Fig. 5.1 be considered. It has an isolated ■ solution in the estimated PS. However, due to overlap along different dimensions of the decision space, ■ has nearby neighbors in both objective and decision space impacting the evaluation (perimeter of hyper-rectangle bounded by neighbors). Thus, by the crowding distance-based sorting approach of [137, 188], this ■ solution appears towards at the end of the sorted list as a more crowded solution. This ambiguity arising due to the use of CDX over the entire decision space is being termed as the crowding illusion problem, henceforth.



(a) Decision Space (b) Objective Space

Figure 5.1: Crowding illusion problem on the results of a benchmark test problem (MMF3 [112]) arises due to overlap along different dimensions of the decision space which gives the illusion that ■ is crowded. Usual sorting (without clustering of solutions in the decision space) from least crowded to most crowded generates ▲●●☆●▲▲☆☆☆☆▲●▲●●●●■● i.e., with ■ at 19$^{\text{th}}$ position whereas LORD's sorting (which relies on clustering of solutions in the decision space) generates ●■☆▲●☆▲●☆▲●☆▲●☆▲●●●● i.e., with ■ at 2$^{\text{nd}}$ position.

5.4 Algorithmic Frameworks of LORD and LORD-II [140]

Graph Laplacian based Optimization with Reference-vector guided Decomposition (LORD) [140] is developed for addressing a wide range of problems (MMMOPs or otherwise). It is further extended to LORD-II for MMMaOPs. In order to reduce the adverse effects of the crowding illusion problem, graph Laplacian based clustering (spectral clustering) is used in LORD to decompose the decision space while reference-vector based approach is used to decompose the objective space. Diversity preservation is collaboratively conducted in each decomposed sub-region.

The following aspects motivate the design of LORD and LORD-II:

1. As there is no standard formulation for the solution diversity in the decision space, it is either denoted by the solution distribution [137, 188] or by the number of optimal solutions [118]. Thus, LORD and LORD-II characterize the solution diversity in the decision space using both the number and distribution of solutions.

2. To reduce the effect of the crowding illusion problem (Section 5.3), LORD clusters a set of non-dominated solutions and, thereafter, computes crowding distance within each cluster. For the example in Fig. 5.1, the ■ solution appears as a much less crowded solution by using the sorting approach of LORD.

3. To yield the competitive performance in objective space similar to standard MOEAs, unlike other MMMOEAs, LORD demonstrates the synergism of diversity preservation, adaptation of hyper-parameters, reference-vector based decomposition of the objective space, and utilization of the neighborhood property [160] (Theorem 4.1) during mating pool formation and candidate selection.

Thus, LORD [140] utilizes decomposition in the decision space and is extended to LORD-II for investigating the scalability on MMMaOPs by varing M and N.

5.4.1 General Framework

LORD (Algorithm 5.1) considers the problem description ($prob(N, M)$), population size (n_{pop}), maximum function evaluations ($MaxFES$) and the set of reference-vectors ($\mathcal{W}$ by Eq. (3.1) to decompose the objective space, Section 3.2.1) as input. It estimates PS and PF as the output. Its major blocks are outlined next.

During initialization (line 2), the population ($\mathcal{A}_{G=1}$) is formed with n_{pop} candidates using Eq. (2.1). The mean values of reproduction parameters ($\mu_{F^{DE},G=1}$, $\mu_{CR,G=1}$ and $\mu_{\eta_c,G=1}$) are initialized. For the k^{th} reference vector ($\mathbf{W}_k$), the indices of other reference vectors are stored in the k^{th} row of the neighborhood lookup matrix, $\mathbf{N}_k \in \mathcal{N}$, sorted by distance from $\mathbf{W}_k$. Then, the for-loop (lines 3 to 13) executes different generations of LORD until $G_{max} = \lfloor MaxFES/n_{dir} \rfloor$. Each generation G iterates over all n_{dir} sub-spaces.

Algorithm 5.1 General Framework of LORD and LORD-II [140]

Input: $prob(N, M)$: An MMMOP having N-dimensional decision space (lower-bounded by $\mathbf{X}^L$ and upper-bounded by $\mathbf{X}^U$) and M-dimensional objective space; n_{pop}: Population size; $MaxFES$: Maximal of fitness evaluations; $\mathcal{W}$: Set of n_{dir} reference vectors (as in [40, 109])

Output: $\mathcal{A}_{G_{max}}$: Estimated PS; $\mathcal{A}_{\mathbf{F},G_{max}}$: Estimated PF

1: **procedure** LORD($prob$, n_{pop}, $MaxFES$, $\mathcal{W}$)
2:     Initialize $\mathcal{A}_G$, $\mathcal{N}$, $\mu_{F^{DE},G}$, $\mu_{CR,G}$, $\mu_{\eta_c,G}$, for $G = 1$
3:     **for** $G = 1$ to G_{max} **do**
4:         $\mathbf{S}_{F^{DE}} \leftarrow \emptyset$, $\mathbf{S}_{CR} \leftarrow \emptyset$, $\mathbf{S}_{\eta_c} \leftarrow \emptyset$
5:         **for** $k = 1$ to n_{dir} (for each direction) **do**
6:             $\left[\mathbf{X}^{child}, F^{DE}, CR, \eta_c\right] \leftarrow$ PERTURB using Algorithm 5.2
7:             $\mathcal{A}_G \leftarrow$ FILTER($\mathcal{A}_G$, $\mathbf{X}^{child}$) using Algorithm 5.4 (LORD) or 5.5 (LORD-II)
8:             **if** $\mathbf{X}^{child} \in \mathcal{A}_G$ **then**
9:                 $\mathbf{S}_{F^{DE}} \leftarrow \mathbf{S}_{F^{DE}} \cup F^{DE}$, $\mathbf{S}_{CR} \leftarrow \mathbf{S}_{CR} \cup CR$, $\mathbf{S}_{\eta_c} \leftarrow \mathbf{S}_{\eta_c} \cup \eta_c$
10:             **end if**
11:         **end for**
12:         $\mu_{F^{DE},G+1} \leftarrow mean(\mathbf{S}_{F^{DE}})$, $\mu_{CR,G+1} \leftarrow mean(\mathbf{S}_{CR})$, $\mu_{\eta_c,G+1} \leftarrow mean(\mathbf{S}_{\eta_c})$
13:     **end for**
14:     **return** $\mathcal{A}_{G_{max}}$ and $\mathcal{A}_{\mathbf{F},G_{max}} = \{\mathbf{F}(\mathbf{X}) | \mathbf{X} \in \mathcal{A}_{G_{max}}\}$
15: **end procedure**

Within one iteration, solution perturbation (line 6) and population filtering (line 7) are performed, as described in the next paragraphs. If the child candidate $\mathbf{X}^{child}$ survives the filtering step, the reproduction parameters involved in its creation are appended to respective success vectors ($\mathbf{S}_{F^{DE}}$, $\mathbf{S}_{CR}$ and $\mathbf{S}_{\eta_c}$) in lines 8 to 10. When the generation G ends, the mean of reproduction parameters are updated in line 12 using respective success vectors. The population ($\mathcal{A}_{G_{max}}$) at the end of G_{max} generations estimates PS and the set $\mathcal{A}_{\mathbf{F},G_{max}}$ of corresponding objective vectors represents the estimated PF.

The generation of $\mathbf{X}^{child}$ in line 6 of Algorithm 5.1 uses Algorithm 5.2. The first parent $\mathbf{X}_1$ is randomly chosen from the candidates associated with $\mathbf{W}_k$ (line 7) where candidate association is dictated by Eq. (3.2). The remaining parents (in line 10 or 17) and also $\mathbf{X}_1$ (if the k^{th} sub-space is empty in lines 3 to 5) are randomly chosen using the mating pool formation principle (described in next paragraph). The parameter P_{mut} chooses between DE/rand/1/bin [153, 168] and SBX crossover [44, 113] (in line 9 or 16). The reproduction parameters (η_c, F^{DE} and CR) are sampled from Gaussian distributions with mean values provided by $\mu_{\eta_c,G}$, $\mu_{F^{DE},G}$ and $\mu_{CR,G}$ and empirically chosen standard deviations, in line 11 or 18. Both SBX crossover and DE/rand/1/bin are followed by Polynomial mutation [110] in lines 14 and 21, respectively, as it helps to avoid local optima [110]. The sampled values of reproduction parameters and $\mathbf{X}^{child}$ are returned in line 24.

Algorithm 5.2 Reproduction of Child Candidate [140]

Input: $\mathcal{A}_G$: Population; $\mathbf{N}_k$: Mating pool; $\{\mu_{F^{DE},G},\ \mu_{CR,G},\ \mu_{\eta_c,G}\}$: Reproduction parameters; $\mathbf{W}_k$: k^{th} reference vector; P_{mut}: Probability of mutation switching
Output: $\mathbf{X}^{child}$: Child; $\{F^{DE},\ CR,\ \eta_c\}$: Reproduction parameters used

1: **procedure** PERTURB($\mathcal{A}_G$, $\mathbf{N}_k$, $\mu_{F^{DE},G}$, $\mu_{CR,G}$, $\mu_{\eta_c,G}$, $\mathbf{W}_k$, P_{mut})
2:     **if** no candidate is associated with $\mathbf{W}_k$ **then**
3:         $\mathbf{N}' \leftarrow$ First k_{nbr} non-empty vectors from $\mathbf{N}_k$
4:         $\mathbf{W}_r \leftarrow$ Reference vector for random index $r \in \mathbf{N}'$
5:         $\mathbf{X}_1 \leftarrow$ Random candidate associated with $\mathbf{W}_r$
6:     **else**
7:         $\mathbf{X}_1 \leftarrow$ Random candidate associated with $\mathbf{W}_k$
8:     **end if**
9:     **if** $rand(0,1) > P_{mut}$ **then**
10:         $\mathcal{A}^{mat}_{k,G} \leftarrow$ MATING_POOL($\mathbf{N}_k$, $\mathcal{A}_G$, 3) using Algorithm 5.3
11:         $F^{DE} \leftarrow N\left(\mu_{F^{DE},G}, 0.1\right)$, $CR \leftarrow N\left(\mu_{CR,G}, 0.1\right)$
12:         $[\mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4] \leftarrow$ Randomly from $\mathcal{A}^{mat}_{k,G}$
13:         $\mathbf{X}'_{child} \leftarrow$ DE/rand/1/bin [165] with $\mathbf{X}_1$ to $\mathbf{X}_4$, F^{DE}, CR using Eq. (2.2)-(2.3)
14:         $\mathbf{X}^{child} \leftarrow$ Polynomial mutation [110] on $\mathbf{X}'_{child}$ using Eq. (4.15)-(4.16)
15:         $\eta_c \leftarrow \emptyset$
16:     **else**
17:         $\mathcal{A}^{mat}_{k,G} \leftarrow$ MATING_POOL($\mathbf{N}_k$, $\mathcal{A}_G$, 1) using Algorithm 5.3
18:         $\eta_c \leftarrow N\left(\mu_{\eta_c,G}, 5\right)$
19:         $\mathbf{X}_2 \leftarrow$ Randomly from $\mathcal{A}^{mat}_{k,G}$
20:         $\mathbf{X}'_{child} \leftarrow$ SBX-crossover [44] with $\mathbf{X}_1$, $\mathbf{X}_2$, η_c using Eq. (4.13)-(4.14)
21:         $\mathbf{X}^{child} \leftarrow$ Polynomial mutation [110] on $\mathbf{X}'_{child}$ using Eq. (4.15)-(4.16)
22:         $F^{DE} \leftarrow \emptyset$, $CR \leftarrow \emptyset$
23:     **end if**
24:     **return** $\mathbf{X}^{child}$, F^{DE}, CR, η_c
25: **end procedure**

Algorithm 5.3 Mating Pool Formation [140]

Input: $\mathbf{N}_k$: Sorted array of nearest neighboring directions of $\mathbf{W}_k$; $\mathcal{A}_G$: Population in decision space; $n_{\mathcal{S}}$: Number of sub-spaces to be chosen
Output: $\mathcal{A}^{mat}_{k,G}$: Sub-population selected for mating

1: **procedure** MATING_POOL($\mathbf{N}_k$, $\mathcal{A}_G$, $n_{\mathcal{S}}$)
2:     $\mathbf{N}' \leftarrow$ First k_{nbr} non-empty vectors from $\mathbf{N}_k$
3:     $\{\mathbf{W}_{r_1}, \cdots, \mathbf{W}_{r_{n_{\mathcal{S}}}}\} \leftarrow$ Reference vectors for random indices $\{r_1, \cdots, r_{n_{\mathcal{S}}}\} \in \mathbf{N}'$
4:     $\mathcal{A}^{mat}_{k,G} \leftarrow$ Candidates of $\mathcal{A}_G$ associated with $\{\mathbf{W}_{r_1}, \cdots, \mathbf{W}_{r_{n_{\mathcal{S}}}}\}$
5:     **return** $\mathcal{A}^{mat}_{k,G}$
6: **end procedure**

The mating pool ($\mathcal{A}^{mat}_{k,G}$) formation in line 10 or 17 of Algorithm 5.2 uses Algorithm 5.3. It considers k_{nbr} nearest non-empty reference vectors of $\mathbf{W}_k$ (line 2), from which $n_{\mathcal{S}}$ random reference vectors $\{\mathbf{W}_{r_1}, \cdots, \mathbf{W}_{r_{n_{\mathcal{S}}}}\}$ are selected in line 3. The parameter $n_{\mathcal{S}}$ is the minimum number of candidates required as per a reproduction strategy. All candidates associated with $\{\mathbf{W}_{r_1}, \cdots, \mathbf{W}_{r_{n_{\mathcal{S}}}}\}$ form $\mathcal{A}^{mat}_{k,G}$ in line 4 and returned from line 5.

To maintain a constant n_{pop}, one of the candidates from $\mathcal{A}_G \cup \mathbf{X}^{child}$ is removed in line 7 of Algorithm 5.1 by calling the filtering operation, which is described after explaining the approach to decompose the population in the decision space.

5.4.2 Decomposition of the Decision Space

The filtering operation in line 7 of Algorithm 5.1 involves graph Laplacian based partitioning (spectral clustering) [178] of a set of solutions ($\mathcal{A}^{nd}$) in the decision space. This clustering operation has the following steps:

1) **Create nearest neighbor graph ($\mathcal{G}$):** All candidates of $\mathcal{A}^{nd}$ are used as the nodes of graph $\mathcal{G}$. Euclidean distances between all pairs of candidates in $\mathcal{A}^{nd}$ are evaluated. Edges are placed between pairs of candidates (nodes) where distance is less than a threshold of ε_L. Specifically, $\mathcal{G}$ (binary symmetric matrix) is the adjacency matrix representation.

2) **Obtain symmetric normalized graph Laplacian ($\mathcal{L}_{sym}$):** A diagonal matrix $\mathcal{G}_d$ is created using the degree of each node (row sum) of $\mathcal{G}$. Using the identity matrix I of the same order as $\mathcal{G}$ and $\mathcal{G}_d$, $\mathcal{L}_{sym}$ [178] is obtained as follows:

$$\mathcal{L}_{sym} = I - \mathcal{G}_d^{-1/2} \mathcal{G} \mathcal{G}_d^{-1/2}. \tag{5.1}$$

3) **Obtain number of connected components ($k_{\mathcal{CC}}$):** The algebraic multiplicity of 0 eigenvalue of $\mathcal{L}_{sym}$ [178] gives the number of connected components ($k_{\mathcal{CC}}$) of $\mathcal{G}$.

4) **Assign candidates (nodes) to $k_{\mathcal{CC}}$ clusters:** By Cheeger's inequality [19,23], the sparsest cut of $\mathcal{G}$ is approximated by the second smallest eigenvalue of $\mathcal{L}_{sym}$ [59]. Thus, all the eigenvectors from the second smallest to the $k_{\mathcal{CC}}^{\text{th}}$ eigenvalues are clustered ($\mathcal{C}_1, \cdots, \mathcal{C}_{k_{\mathcal{CC}}}$) using k-means [37] for assigning the candidates of $\mathcal{A}^{nd}$ to the clusters in the decision space. Examples of clustering of a non-dominated set of solutions are illustrated in Fig. 5.1a and Fig. 5.2a for benchmark test problems [112]: MMF3 and MMF2, respectively.

For reducing crowding illusion (Section 5.3), spectral clustering of $\mathcal{A}^{nd}$ is chosen over k-means clustering due to the following reasons: (1) k-means is effective only for globular structures whereas spectral clustering is effective for non-globular structure as well, (2) $k_{\mathcal{CC}}$ for k-means is not known apriori whereas $k_{\mathcal{CC}}$ for spectral clustering can be obtained mathematically and (3) k-means (involved in step 4 of decomposition of $\mathcal{A}^{nd}$) is independent of the number of decision variables (N).

5.4.3 Filtering Scheme of LORD

For maintaining the convergence and the solution diversity in both the objective and the decision spaces, the filtering operation of LORD (Algorithm 5.4) has the following steps:

1. *Obtain last non-dominated rank (maintaining convergence in objective space):* Using non-dominated sorting on $(\mathcal{A}_G \cup \mathbf{X}^{child})$, the solutions $(\mathcal{A}^{nd})$ in the last non-dominated rank [109, 125] are obtained in lines 2 to 4. If $\left|\mathcal{A}^{nd}\right| = 1$, lines 5 to 17 yield the only $\mathbf{X}_{del} \in \mathcal{A}^{nd}$ for elimination. Otherwise, some $\mathbf{X}_{del} \in \mathcal{A}^{nd}$ (from the least converged set of mutually non-dominated points) is eliminated by the next steps.

2. *Spectral clustering of candidates from $\mathcal{A}^{nd}$ (maintaining diversity in decision space):* The candidates in $\mathcal{A}^{nd}$ is partitioned in line 5 as mentioned in Section 5.4.2. Evaluating the crowding in the respective spaces, Special Crowding Distance (SCD) [137,188] combines CDF $(\sum_{j=1}^{M} D_{crowd}(\mathbf{X}|f_j))$ and CDX $(\sum_{k=1}^{N} D_{crowd}(\mathbf{X}|x_k))$ by obtaining $D_{crowd}(.)$ from Eq. (2.5) as follows:

$$SCD(\mathbf{X}) = \begin{cases} max\left(CDX(\mathbf{X}), CDF(\mathbf{X})\right), & \text{if } CDX(\mathbf{X}) > \text{mean}\left(CDX(\mathbf{X})\right) \\ & \text{or } CDF(\mathbf{X}) > \text{mean}\left(CDF(\mathbf{X})\right) \\ min\left(CDX(\mathbf{X}), CDF(\mathbf{X})\right), & \text{otherwise.} \end{cases} \tag{5.2}$$

   SCD is evaluated per cluster in line 6. A sorted set $(\mathcal{A}_s^{nd})$ of candidates is formed by rearranging $\mathcal{A}^{nd}$ in line 7 where at first the candidates with the highest SCD are selected from each cluster, then candidates with the second-highest SCD are selected from each cluster and so on. An example in Fig. 5.2b demonstrates $\mathcal{A}_s^{nd}$.

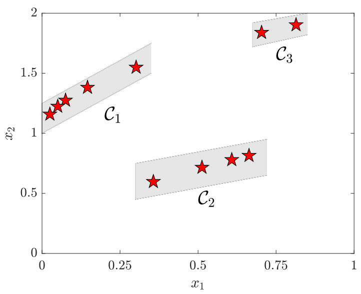
3. *Association based elimination of candidate from $\mathcal{A}_s^{nd}$ (maintaining diversity in objective space):* Starting from the last candidate (worst) in $\mathcal{A}_s^{nd}$, the reference vector $\mathbf{W}_k$ is obtained in line 9 with which $\mathbf{X}_j \in \mathcal{A}_s^{nd}$ associates. If multiple candidates of $(\mathcal{A}_G \cup \mathbf{X}^{child})$ are associated with $\mathbf{W}_k$ (implying a dense sub-space), $\mathbf{X}_{del} = \mathbf{X}_j$ is chosen for deletion (lines 10 to 13) as exemplified in Fig. 5.2 (see caption for details). If all the sub-spaces with which candidates of $\mathcal{A}_s^{nd}$ are associated have only one candidate, the last candidate from $\mathcal{A}_s^{nd}$ is chosen for deletion (lines 15 to 17). $\mathbf{X}_{del}$ is deleted from $(\mathcal{A}_G \cup \mathbf{X}^{child})$ in line 18 to yield the filtered $\mathcal{A}_G$ for the next

iteration. This filtered $\mathcal{A}_G$ is returned from line 19 of Algorithm 5.4 to line 7 of Algorithm 5.1.

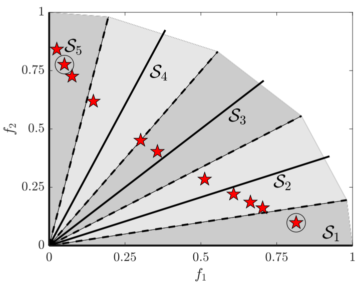
Algorithm 5.4 Filter for constant Population Size (LORD) [140]

Input: $\mathcal{A}_G$: Current population; $\mathbf{X}^{child}$: Child candidate
Output: $\mathcal{A}_G$: Filtered population of size n_{pop};

```
1:  procedure FILTER(A_G, X^child)
2:      A_F^all = {F(X) | X ∈ (A_G ∪ X^child)}
3:      A_F^nd ← Last non-dominated rank of A_F^all
4:      A^nd = {X | F(X) ∈ A_F^nd}
5:      {C_1, ··· , C_kCC} ← Spectral clustering of A^nd
6:      Evaluate SCD cluster-wise
7:      A_s^nd ← Rearrange A^nd by select one-by-one from C_1 to C_kCC w.r.t. SCD
8:      for j = |A_s^nd| to 1 (starting from most-crowded) do
9:          W_k ← Direction where X_j ∈ A_s^nd is associated
10:         if number of candidates associated with W_k > 1 then
11:             X_del ← Assign X_j for deletion
12:             Break loop
13:         end if
14:     end for
15:     if no X_del is chosen then
16:         X_del ← Last candidate of A_s^nd
17:     end if
18:     A_G ← (A_G ∪ X^child) − X_del
19:     return A_G
20: end procedure
```



$\mathbf{X} \in \mathcal{C}_1 \vert \mathbf{X}$ has highest SCD in $\mathcal{C}_1$
$\mathbf{X} \in \mathcal{C}_2 \vert \mathbf{X}$ has highest SCD in $\mathcal{C}_2$
$\mathbf{X} \in \mathcal{C}_3 \vert \mathbf{X}$ has highest SCD in $\mathcal{C}_3$
$\mathbf{X} \in \mathcal{C}_1 \vert \mathbf{X}$ has second-highest SCD in $\mathcal{C}_1$
$\mathbf{X} \in \mathcal{C}_2 \vert \mathbf{X}$ has second-highest SCD in $\mathcal{C}_2$
$\mathbf{X} \in \mathcal{C}_3 \vert \mathbf{X}$ has second-highest SCD in $\mathcal{C}_3$
$\mathbf{X} \in \mathcal{C}_1 \vert \mathbf{X}$ has third-highest SCD in $\mathcal{C}_1$
$\mathbf{X} \in \mathcal{C}_2 \vert \mathbf{X}$ has third-highest SCD in $\mathcal{C}_2$
$\mathbf{X} \in \mathcal{C}_1 \vert \mathbf{X}$ has fourth-highest SCD in $\mathcal{C}_1$
$\mathbf{X} \in \mathcal{C}_2 \vert \mathbf{X}$ has fourth-highest SCD in $\mathcal{C}_2$
$\mathbf{X} \in \mathcal{C}_1 \vert \mathbf{X}$ has fifth-highest SCD in $\mathcal{C}_1$



(a) Decision Space (MMF2) (b) Sorted Set ($\mathcal{A}_s^{nd}$) (c) Objective Space (MMF2)

Figure 5.2: Filtering steps of LORD on a non-dominated set of solutions ($\mathcal{A}^{nd}$) which rearranges candidates according to maximal SCD per cluster ($\mathcal{C}_1$ to $\mathcal{C}_{k_{CC}} = \mathcal{C}_3$) to form the sorted set ($\mathcal{A}_s^{nd}$). LORD removes one candidate from the end of $\mathcal{A}_s^{nd}$ if it is not the only candidate within a subspace (e.g., the encircled candidate from $\mathcal{S}_1$ will not be removed, whereas the encircled candidate from $\mathcal{S}_5$ can be removed) [140].

Explicit maintenance of the three essential properties is the most important characteristics of LORD as a novel MMMOEA. While SCD explicitly accounts for solution distribution in decision space, the candidates towards the end of $\mathcal{A}_s^{nd}$ come from the bigger clusters (e.g., Fig. 5.1 and Fig. 5.2) and are more likely to be deleted. Hence, LORD implicitly takes care of the neighborhood count also as a diversity criterion.

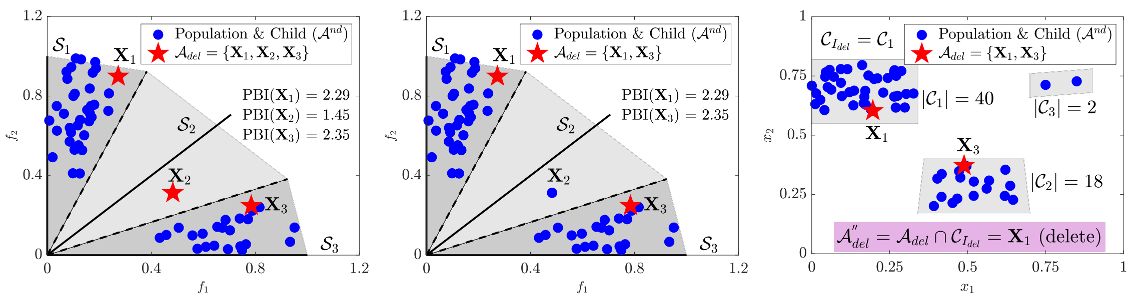
5.4.4 Filtering Scheme of LORD-II

For avoiding dominance resistance with high number of objectives [32, 127], the filtering operation is modified to yield LORD-II. It is based on Penalty-based Boundary Intersection (PBI, Eq. (3.3)) and uses Algorithm 5.5 which involves the following steps:

1. *PBI-based selection for deletion (maintaining convergence in objective space):* The objective vectors corresponding to all candidates of $\mathcal{A}_G$ and $\mathbf{X}^{child}$ are stored in $\mathcal{A}^{nd}_{\mathbf{F}}$ in line 3. From each sub-space, the candidate with the maximum PBI [40, 109, 138] is stored in $\mathcal{A}_{del}$ (lines 4 to 10) as potential candidates for deletion.

2. *Disregarding based on association (maintaining diversity in objective space):* Deletion of a candidate from any sub-space with only one candidate would hamper the diversity in the objective space. Hence, it is not considered in $\mathcal{A}_{del}$ (lines 7 to 10).

3. *Spectral clustering of candidates (maintaining diversity in decision space):* The candidates in $\mathcal{A}^{nd}$ are partitioned in line 11 as mentioned in Section 5.4.2. The cardinality is noted (lines 12 to 19) for those clusters, which share common element with $\mathcal{A}_{del}$ (line 14). From the largest cluster ($\mathcal{C}_{I_{del}}$), the candidates common to those in $\mathcal{A}_{del}$ are chosen to yield $\mathcal{A}''_{del}$ in line 20. The candidate with the largest PBI in $\mathcal{A}''_{del}$ is deleted (lines 21 to 22) to yield the filtered $\mathcal{A}_G$ for the next iteration.



(a) PBI-based creation of $\mathcal{A}_{del}$ (b) Tuning of $\mathcal{A}_{del}$ (c) Clustering in decision space

Figure 5.3: Filtering of LORD-II on a set of solutions ($\mathcal{A}^{nd}$): (a) candidates ($\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$) with maximal PBI from each sub-space form $\mathcal{A}_{del}$, (b) sub-spaces with only one candidate ($\mathcal{S}_2$ with $\mathbf{X}_2$) are disregarded in $\mathcal{A}_{del}$, (c) candidate $\mathbf{X}_1$, common to both largest cluster ($\mathcal{C}_{I_{del}} = \mathcal{C}_1$ of size 40) and $\mathcal{A}_{del}$, is deleted [140].

Algorithm 5.5 Filter for constant Population Size (LORD-II) [140]

Input: $\mathcal{A}_G$: Current population; $\mathbf{X}^{child}$: Child candidate
Output: $\mathcal{A}_G$: Filtered population of size n_{pop};

procedure FILTER($\mathcal{A}_G$, $\mathbf{X}^{child}$)
  $\mathcal{A}^{nd} \leftarrow \mathcal{A}_G \cup \mathbf{X}^{child}$
  $\mathcal{A}^{nd}_{\mathbf{F}} = \{\mathbf{F}(\mathbf{X}) | \mathbf{X} \in \mathcal{A}^{nd}\}$
  $\mathcal{A}_{del} \leftarrow \emptyset$
  for $k = 1$ to n_{dir} (for each direction) **do**
    $\mathcal{A}^{sub}_{\mathbf{F},k} \leftarrow$ Candidates of $\mathcal{A}^{nd}_{\mathbf{F},G}$ associated with $\mathbf{W}_k$
    if $\left|\mathcal{A}^{sub}_{\mathbf{F},k}\right| > 1$ **then**
      $\mathcal{A}_{del} \leftarrow \mathcal{A}_{del} \cup (\mathbf{X}$ with max PBI in $\mathcal{A}^{sub}_{\mathbf{F},k})$
    end if
  end for
  $[\mathcal{C}_1, \cdots, \mathcal{C}_{k_{CC}}] \leftarrow$ Spectral clustering of $\mathcal{A}^{nd}$
  $I_{del} = 0$, $M_{del} = 0$
  for $j = 1$ to k_{CC} (for all clusters) **do**
    if $\mathcal{A}_{del} \cap \mathcal{C}_j \neq \emptyset$ **then**
      if $M_{del} < |\mathcal{C}_j|$ **then**
        $I_{del} = j$, $M_{del} = |\mathcal{C}_j|$
      end if
    end if
  end for
  $\mathcal{A}''_{del} \leftarrow \mathcal{C}_{I_{del}} \cap \mathcal{A}_{del}$
  $\mathbf{X}_{del} \leftarrow$ Select candidate with max PBI from $\mathcal{A}''_{del}$
  $\mathcal{A}_G \leftarrow \mathcal{A}^{nd} - \mathbf{X}_{del}$
  return $\mathcal{A}_G$
end procedure

While the cluster size explicitly accounts for the number of optimal solutions, the spectral clustering implicitly accounts for the solution distribution in the decision space.

The combination of these operations allows the LORD and LORD-II to effectively address MMMOPs and MMMaOPs, respectively.

5.5 Experimental Results

For performance analysis, LORD and LORD-II are implemented in MATLAB R2018a using a 64-bit computer (8 GB RAM, Intel Core i7 @ 2.20 GHz). The experimental specifications of the benchmark MMMOPs, performance measures and parameter settings of various competitor MOEAs are provided in the following sub-sections.

5.5.1 Benchmark Problems

The benchmark problems from CEC 2019 test suite for MMMOPs [112] (defined in Section A.5) are considered with $MaxFES = 5000 \times N$ and $n_{pop} = 100 \times N$, as per [112]. Specifications of these MMMOPS are mentioned in Table 5.1. It should be noted that MMF12 has discontinuous PF, hence the number of subsets in the global PS (#PSs) is one per Pareto-optimal patch. While MMF10-13, MMF15 and MMF15_a have one global PS but these are multi-modal problems as these have local PSs close to their global PS.

Table 5.1: Specifications for M-objective MMMOPs in terms of N-dimensional decision space, upper and lower bounded between $\mathbf{X}^U$ and $\mathbf{X}^L$ having reference point at $\mathbf{R}_{HV}$ for HV calculation with N_{IGD} number of points in the reference set for IGD evaluation and number of subsets in the global PS (#PSs) [137, 140].

Problems	N	M	$\mathbf{X}^L$	$\mathbf{X}^U$	$\mathbf{R}_{HV}$	N_{IGD}	#PSs
MMF1	2	2	$[1,-1]$	$[3,1]$	$[1.1,1.1]$	400	2
MMF1_z	2	2	$[1,-1]$	$[3,1]$	$[1.1,1.1]$	400	2
MMF1_e	2	2	$[1,-20]$	$[3,20]$	$[1.1,1.1]$	400	2
MMF2	2	2	$[0,0]$	$[1,1]$	$[1.1,1.1]$	400	2
MMF3	2	2	$[0,0]$	$[1,1.5]$	$[1.1,1.1]$	400	2
MMF4	2	2	$[-1,0]$	$[1,2]$	$[1.1,1.1]$	400	4
MMF5	2	2	$[1,-1]$	$[3,3]$	$[1.1,1.1]$	400	4
MMF6	2	2	$[1,-1]$	$[3,2]$	$[1.1,1.1]$	400	4
MMF7	2	2	$[1,-1]$	$[3,1]$	$[1.1,1.1]$	400	2
MMF8	2	2	$[-\pi,0]$	$[\pi,9]$	$[1.1,1.1]$	400	4
MMF9	2	2	$[0.1,0.1]$	$[1.1,1.1]$	$[1.21,11]$	400	2
MMF10	2	2	$[0.1,0.1]$	$[1.1,1.1]$	$[1.21,13.2]$	400	1
MMF11	2	2	$[0.1,0.1]$	$[1.1,1.1]$	$[1.21,15.4]$	400	1
MMF12	2	2	$[0,0]$	$[1,1]$	$[1.54,1.1]$	410	1
MMF13	3	2	$[0.1,0.1,0.1]$	$[1.1,1.1,1.1]$	$[1.54,15.4]$	1250	1
MMF14	$N=M$	$M\geq 3$	$[0,\overset{M}{\cdots},0]$	$[1,\overset{M}{\cdots},1]$	$[2.2,\overset{M}{\cdots},2.2]$	1250	2
MMF14_a	$N=M$	$M\geq 3$	$[0,\overset{M}{\cdots},0]$	$[1,\overset{M}{\cdots},1]$	$[2.2,\overset{M}{\cdots},2.2]$	1250	2
MMF15	$N=M$	$M\geq 3$	$[0,\overset{M}{\cdots},0]$	$[1,\overset{M}{\cdots},1]$	$[2.5,\overset{M}{\cdots},2.5]$	1250	1
MMF15_a	$N=M$	$M\geq 3$	$[0,\overset{M}{\cdots},0]$	$[1,\overset{M}{\cdots},1]$	$[2.5,\overset{M}{\cdots},2.5]$	1250	1
Omni-test	3	2	$[0,0,0]$	$[6,6,6]$	$[4.4,4.4]$	600	27
SYM-PART simple	2	2	$[-20,-20]$	$[20,20]$	$[4.4,4.4]$	396	9
SYM-PART rotated	2	2	$[-20,-20]$	$[20,20]$	$[4.4,4.4]$	396	9

The polygon MMMaOPs [76] (defined in Section A.6) are considered with $MaxFES = 10000$ as per [170]. The specifications (p_1 and p_2) for defining n_{dir} reference vectors (as per Section 3.2.1) are mentioned in Table 5.2. The goal is to satisfy $n_{dir} \approxeq n_{pop} = 100N$.

5.5.2 Performance Indicators

In the objective space, Inverted Generational Distance (IGD) [32] and Hypervolume indicator (HV) [9] are noted for CEC 2019 MMMOPs [112] with $M = 2$ to assess the convergence

Table 5.2: Specifications for reference-vector based decomposition for problems with M objectives and N decision variables [140].

M	p_1	p_2	n_{dir}
2	$100N - 1$	0	$100N$
3	23	0	300
5	8	0	495
8	5	2	828
10	4	3	935

and diversity of MOEAs [32]. The size of the reference sets[2] ($N_{IGD} = |\mathcal{H}_{IGD}|$) for IGD evaluation and the reference points ($\mathbf{R}_{HV}$) for HV evaluation are specified in Table 5.1. For polygon MMMaOPs, IGD with $N_{IGD} = 5000$ is used. Convergence Metric (CM) [10] with $\mathcal{H}_{CM} = \mathcal{H}_{IGD}$ and *D_metric* [161] are used for CEC 2019 MMMOPs [112] with $M \geq 3$ to individually assess the convergence and diversity of MOEAs.

In decision space, IGD [198] and Pareto-Set Proximity (PSP) [188] are used for CEC 2019 MMMOPs [112] with $M = 2$ and for polygon MMMaOPs [76] to assess the performance of MOEAs. For CEC 2019 MMMOPs [112] with $M \geq 3$, the fraction of non-contributing solutions (NSX) [173,174] and the convergence metric of this non-contributing set (CM_NSX) are noted. Hereafter, IGDX and IGDF represent IGD values in decision and objective space, respectively, and rHV=1/HV and rPSP=1/PSP are noted such that lower value is the better measure over all the indicators. Brief description of all the performance indicators are provided in Section 1.3.3.

5.5.3 Details of Competitor Algorithms

As DN-NSGA-II[3] [113] and MO_Ring_PSO_SCD[3] [188] use non-dominated sorting with CDX, LORD[3] is compared with these two MMMOEAs. For comparison with a standard MOEA outperforming the former MMMOEAs in the objective space, LORD is also compared with NSGA-II [127, 128]. For MMMOPs with $M \geq 3$, LORD-II[3] is compared with MO_Ring_PSO_SCD [188] and a decomposition-based MOEA (MOEA/DD[3]) [109].

Other MMMOEAs (Omni-Optimizer [51], TriMOEA_TA&R [118], MM-NAEMO [120], DE-TriM [137] and NIMMO [170]) have demonstrated their effectiveness only for certain kinds of test problems. These MMMOEAs are also compared with LORD and LORD-II

[2] Reference sets are obtained from `http://www5.zzu.edu.cn/ecilab/info/1036/1163.htm` for CEC 2019 MMMOPs [112] and from `https://sites.google.com/view/nimmopt/` for polygon MMMaOPs [76].

[3] The MATLAB codes are acquired from `http://www5.zzu.edu.cn/ecilab/info/1036/1163.htm` for MO_Ring_PSO_SCD and DN-NSGA-II, and from `https://github.com/BIMK/PlatEMO` for MOEA/DD. Source code of LORD and LORD-II is available at `http://worksupplements.droppages.com/lord`.

on some CEC 2019 MMMOPs [112] and polygon problems [76, 170].

Most of the hyper-parameters of LORD and LORD-II are adaptive, while the rest of them are set as mentioned in Table 5.3.

Table 5.3: Recommended values of different parameters for LORD and LORD-II.

Parameters	**Values**	**Remarks**
k_{nbr}	$0.2 \times n_{dir}$	Number of non-empty neighboring directions for mating pool formation (line 2, Algorithm 5.3) which is easily within $0.2 \times n_{dir}$ all test cases (except MMF12) have regular PFs
P_{mut}	0.25	Probability of switching among reproduction methods (line 9, Algorithm 5.2) and sensitivity is analyzed in Section 5.5.4
ε_L	α_L times diagonal of $\mathcal{D}$ with $\alpha_L = 0.2$	Threshold on inter-solution distance for formation of nearest neighbor graph (Section 5.4.2) and sensitivity is analyzed in Section 5.5.4
$\mu_{F^{DE},G=1}$, $\mu_{CR,G=1}$ and $\mu_{\eta_c,G=1}$	Initialized as 0.5, 0.2 and 30	Initial mean values of reproduction parameters (line 2, Algorithm 5.1), later adapted per generation

5.5.4 Parameter Sensitivity Studies

Two experiments are presented to study the sensitivity of the following parameters: (1) threshold (ε_L) for nearest neighbor graph formation (Section 5.4.2), and (2) the probability (P_{mut}) of switching between DE/rand/1/bin and SBX crossover (line 9, Algorithm 5.2).

1. **Threshold for Nearest Neighbor Graph**: During spectral clustering (Section 5.4.2) in LORD and LORD-II, the formation of the nearest neighbor graph ($\mathcal{G}$) considers edges between those pairs of solutions (nodes) whose distance is less than the threshold ε_L. This parameter ε_L is set as α_L ($= 0.2$) times the longest distance in the decision space, i.e., diagonal of the box-constrained decision space, $\mathcal{D}$. For validating this value of α_L, it is varied between 0.1 to 0.8 (10% to 80% of the diagonal of $\mathcal{D}$) and the performance of LORD and LORD-II are noted in Table 5.4 for some MMMOPs with $M = 2$ or $M = 3$.

From Table 5.4, the best performance is observed when $\alpha_L = 0.2$. The performance deteriorates for higher α_L as all the candidates in $\mathcal{A}^{nd}$ form a single cluster ($k_{\mathcal{CC}} = 1$) and distinguishability of the multiple subsets in PS is lost. The performance also deteriorates for lower α_L as $k_{\mathcal{CC}} \rightarrow \left|\mathcal{A}^{nd}\right|$ and the candidates become independent (higher randomness).

2. **Probability of Reproduction Switching**: During the probabilistic mutation switching (Algorithm 5.2) in LORD and LORD-II, P_{mut} decides between DE/rand/1/bin [41, 168] and SBX-crossover [44]. However, in either case, polynomial mutation [110] is

Table 5.4: Mean IGDX and IGDF over 51 independent runs for sensitivity study of α_L (parameter of LORD and LORD-II) on some 2- and 3-objective MMMOPs [140].

		IGDX				**IGDF**			
	$\alpha_L \rightarrow$	0.1	0.2	0.5	0.8	0.1	0.2	0.5	0.8
LORD	MMF1	0.0504	0.0431	0.0479	0.0492	0.0028	0.0025	0.0025	0.0028
	MMF2	0.1431	0.0180	0.0304	0.0366	0.0092	0.0070	0.0109	0.0173
	MMF3	0.0459	0.0176	0.0419	0.0458	0.0084	0.0069	0.0103	0.0117
	MMF4	0.0298	0.0251	0.0303	0.0352	0.0021	0.0018	0.0023	0.0024
	MMF5	0.0976	0.0814	0.0943	0.1165	0.0025	0.0024	0.0025	0.0027
	MMF6	0.0812	0.0692	0.0720	0.0890	0.0025	0.0023	0.0024	0.0024
	MMF7	0.0277	0.0218	0.0299	0.0339	0.0024	0.0022	0.0026	0.0028
	MMF8	0.1631	0.0762	0.1299	0.1577	0.0025	0.0025	0.0025	0.0025
LORD-II	MMF14	0.0495	0.0443	0.0522	0.0580	0.0550	0.0540	0.0545	0.0546
	MMF14_a	0.0657	0.0576	0.0665	0.0674	0.0574	0.0561	0.0582	0.0583
	MMF15	0.0295	0.0287	0.0292	0.0293	0.0552	0.0548	0.0552	0.0558
	MMF15_a	0.0369	0.0355	0.0373	0.0379	0.0584	0.0571	0.0589	0.0593

also executed. This parameter P_{mut} is set as 0.25 after investigating the following cases:

1. $P_{mut} = 0.00$: only DE/rand/1/bin is used,
2. $P_{mut} = 0.25$: DE/rand/1/bin is used more often than SBX-crossover,
3. $P_{mut} = 0.50$: DE/rand/1/bin and SBX-crossover are equally-likely to be used,
4. $P_{mut} = 0.75$: SBX-crossover is used more often than DE/rand/1/bin, and
5. $P_{mut} = 1.00$: only SBX-crossover is used.

The performance of LORD and LORD-II are noted in Table 5.5 for some MMMOPs. From Table 5.5, the best performance is observed when $P_{mut} = 0.25$. Hence, for exploration of the search space, DE/rand/1/bin is preferred over SBX-crossover [176] along with a switching scheme to combine the benefits of both these strategies.

Table 5.5: Mean IGDX and IGDF over 51 independent runs for sensitivity study of P_{mut} (parameter of LORD and LORD-II) on some 2- and 3-objective MMMOPs [140].

		IGDX					**IGDF**				
	$P_{mut} \rightarrow$	0.00	0.25	0.50	0.75	1.00	0.00	0.25	0.50	0.75	1.00
LORD	MMF1	0.0529	0.0431	0.0470	0.0472	0.0506	0.0028	0.0026	0.0027	0.0027	0.0027
	MMF2	0.0694	0.0110	0.0169	0.0207	0.0251	0.0100	0.0069	0.0085	0.0097	0.0141
	MMF3	0.0603	0.0275	0.0116	0.0188	0.0217	0.0169	0.0065	0.0070	0.0082	0.0475
	MMF4	0.0283	0.0237	0.0239	0.0287	0.0381	0.0023	0.0021	0.0021	0.0023	0.0025
	MMF5	0.0923	0.0789	0.0738	0.0900	0.0904	0.0027	0.0025	0.0024	0.0026	0.0028
	MMF6	0.1199	0.0693	0.0777	0.0827	0.0976	0.0025	0.0024	0.0025	0.0025	0.0025
	MMF7	0.0240	0.0209	0.0229	0.0228	0.0278	0.0024	0.0023	0.0023	0.0023	0.0025
	MMF8	0.4619	0.1197	0.1085	0.0737	0.1123	0.0026	0.0025	0.0026	0.0025	0.0026
LORD-II	MMF14	0.0490	0.0484	0.0497	0.0494	0.0502	0.0547	0.0545	0.0542	0.0548	0.0554
	MMF14_a	0.0617	0.0609	0.0613	0.0650	0.0671	0.0578	0.0563	0.0580	0.0589	0.0596
	MMF15	0.0296	0.0288	0.0291	0.0292	0.0291	0.0553	0.0551	0.0552	0.0553	0.0560
	MMF15_a	0.0374	0.0370	0.0364	0.0372	0.0378	0.0588	0.0588	0.0593	0.0594	0.0606

5.5.5 Comparison of LORD and LORD-II with Other MMMOEAs

Four sets of experiments are conducted to compare the performance of LORD and LORD-II with other MMMOEAs.

1) **Experiment-I: Comparison on CEC 2019 Test Suite**: For 2-objective MMMOPs, the performance of LORD in decision and objective spaces are presented in Tables 5.6 and 5.7, respectively. For M-objective MMMOPs (with $M \geq 3$), the performance of LORD-II in decision and objective spaces are presented in Tables 5.8 and 5.9, respectively. The estimated PSs and PFs are also plotted in Fig. 5.4 for some of the MMMOPs. All the results are statistically validated using the Wilcoxon's rank-sum test [173] under the null hypothesis (H_0) that the performance of LORD (or LORD-II) is equivalent to other MMMOEAs. The statistical significance is indicated using three signs: + denoting LORD (or LORD-II) is superior, − denoting the competitor MMMOEA is superior, and ∼ indicating the algorithms are equivalent.

Table 5.6: Mean and standard deviation of rPSP and IGDX over 51 independent runs for comparing LORD on 2-objective MMMOPs [140].

Problems	rPSP=IGDX/CoRa				IGDX			
	LORD	**MO_Ring_PSO_SCD**	**DN-NSGA-II**	**NSGA-II**	**LORD**	**MO_Ring_PSO_SCD**	**DN-NSGA-II**	**NSGA-II**
MMF1	0.0441 ± 0.0044	0.0489 ± 0.0018 (+)	0.0957 ± 0.0146 (+)	0.0652 ± 0.0103 (+)	0.0431 ± 0.0044	0.0485 ± 0.0017 (+)	0.0939 ± 0.0141 (+)	0.0645 ± 0.0098 (+)
MMF1_z	0.0356 ± 0.0069	0.0354 ± 0.0019 (∼)	0.0822 ± 0.0166 (+)	0.3892 ± 0.3913 (+)	0.0351 ± 0.0075	0.0352 ± 0.0018 (∼)	0.0805 ± 0.0157 (+)	0.2606 ± 0.1608 (+)
MMF1_e	0.8894 ± 0.1466	0.5501 ± 0.1276 (−)	1.7201 ± 1.2086 (+)	14.0870 ± 8.1289 (+)	0.7499 ± 0.4192	0.4738 ± 0.0847 (−)	1.1536 ± 0.5095 (+)	3.0324 ± 0.7634 (+)
MMF2	0.0219 ± 0.0108	0.0444 ± 0.0113 (+)	0.1356 ± 0.0805 (+)	0.0766 ± 0.0402 (+)	0.0180 ± 0.0093	0.0416 ± 0.0103 (+)	0.1121 ± 0.0525 (+)	0.0650 ± 0.0300 (+)
MMF3	0.0200 ± 0.0105	0.0294 ± 0.0074 (+)	0.1249 ± 0.1291 (+)	0.0785 ± 0.0416 (+)	0.0176 ± 0.0080	0.0276 ± 0.0061 (+)	0.0968 ± 0.0632 (+)	0.0661 ± 0.0311 (+)
MMF4	0.0253 ± 0.0036	0.0274 ± 0.0014 (+)	0.0854 ± 0.0232 (+)	0.1066 ± 0.0468 (+)	0.0251 ± 0.0039	0.0271 ± 0.0014 (+)	0.0849 ± 0.0230 (+)	0.1004 ± 0.0411 (+)
MMF5	0.0814 ± 0.0080	0.0864 ± 0.0045 (+)	0.1788 ± 0.0179 (+)	0.1525 ± 0.0296 (+)	0.0814 ± 0.0074	0.0857 ± 0.0044 (+)	0.1763 ± 0.0165 (+)	0.1478 ± 0.0265 (+)
MMF6	0.0692 ± 0.0104	0.0741 ± 0.0044 (+)	0.1453 ± 0.0176 (+)	0.1410 ± 0.0272 (+)	0.0692 ± 0.0104	0.0736 ± 0.0042 (+)	0.1433 ± 0.0173 (+)	0.1372 ± 0.0251 (+)
MMF7	0.0219 ± 0.0044	0.0264 ± 0.0014 (+)	0.0535 ± 0.0098 (+)	0.0452 ± 0.0132 (+)	0.0218 ± 0.0025	0.0262 ± 0.0014 (+)	0.0524 ± 0.0092 (+)	0.0420 ± 0.0106 (+)
MMF8	0.0745 ± 0.0452	0.0679 ± 0.0049 (∼)	0.2969 ± 0.1120 (+)	0.9348 ± 0.4682 (+)	0.0762 ± 0.0504	0.0673 ± 0.0048 (∼)	0.2860 ± 0.1078 (+)	0.7198 ± 0.3034 (+)
MMF9	0.0047 ± 0.0002	0.0079 ± 0.0005 (+)	0.0229 ± 0.0081 (+)	1.7445 ± 1.9877 (+)	0.0046 ± 0.0002	0.0079 ± 0.0005 (+)	0.0229 ± 0.0081 (+)	0.1783 ± 0.0740 (+)
MMF10	0.0018 ± 0.0007	0.0293 ± 0.0113 (+)	0.1426 ± 0.0834 (+)	0.0398 ± 0.1184 (∼)	0.0018 ± 0.0009	0.0276 ± 0.0092 (+)	0.1295 ± 0.0747 (+)	0.0398 ± 0.1184 (∼)
MMF11	0.0029 ± 0.0002	0.0055 ± 0.0003 (+)	0.0045 ± 0.0003 (+)	0.0027 ± 0.0003 (−)	0.0029 ± 0.0002	0.0054 ± 0.0003 (+)	0.0045 ± 0.0003 (+)	0.0027 ± 0.0003 (−)
MMF12	0.0013 ± 0.0001	0.0038 ± 0.0003 (+)	0.0090 ± 0.0159 (+)	0.0013 ± 0.0002 (∼)	0.0013 ± 0.0001	0.0038 ± 0.0003 (+)	0.0090 ± 0.0159 (+)	0.0013 ± 0.0002 (∼)
MMF13	0.0243 ± 0.0039	0.0317 ± 0.0014 (+)	0.0614 ± 0.0070 (+)	0.1492 ± 0.0652 (+)	0.0242 ± 0.0039	0.0314 ± 0.0013 (+)	0.0609 ± 0.0064 (+)	0.0880 ± 0.0173 (+)
Omni-test	0.0754 ± 0.0242	0.3946 ± 0.0939 (+)	1.4390 ± 0.2069 (+)	1.8176 ± 0.6886 (+)	0.0706 ± 0.0215	0.3907 ± 0.0927 (+)	1.4159 ± 0.1986 (+)	1.4210 ± 0.3726 (+)
SYM-PART simple	0.0556 ± 0.0145	0.1741 ± 0.0301 (+)	4.1590 ± 0.8683 (+)	113.0044 ± 131.2343 (+)	0.0549 ± 0.0130	0.1733 ± 0.0300 (+)	4.0657 ± 0.7040 (+)	6.8332 ± 1.8906 (+)
SYM-PART rotated	0.1730 ± 0.0743	0.3142 ± 0.3533 (+)	5.5941 ± 3.6017 (+)	13.9239 ± 12.8588 (+)	0.1558 ± 0.0760	0.2926 ± 0.2938 (+)	3.7659 ± 1.2478 (+)	5.4249 ± 1.9790 (+)
LORD vs. others (+/ − / ∼)		15/1/2	18/0/0	15/1/2	(+/ − / ∼)	15/1/2	18/0/0	15/1/2

Table 5.7: Mean and standard deviation of rHV and IGDF over 51 independent runs for comparing LORD on 2-objective MMMOPs [140].

Problems	**rHV=1/HV**				**IGDF**			
	LORD	**MO_Ring_PSO_SCD**	**DN-NSGA-II**	**NSGA-II**	**LORD**	**MO_Ring_PSO_SCD**	**DN-NSGA-II**	**NSGA-II**
MMF1	1.0737 ± 0.0008	1.1484 ± 0.0005 (+)	1.1495 ± 0.0014 (+)	1.0738 ± 0.0006 (∼)	0.0025 ± 0.0002	0.0037 ± 0.0002 (+)	0.0043 ± 0.0005 (+)	0.0028 ± 0.0004 (+)
MMF1_z	1.0731 ± 0.0008	1.1483 ± 0.0005 (+)	1.1484 ± 0.0009 (+)	1.1255 ± 0.0615 (+)	0.0022 ± 0.0001	0.0036 ± 0.0002 (+)	0.0036 ± 0.0004 (+)	0.0396 ± 0.0496 (+)
MMF1_e	1.0751 ± 0.0021	1.1861 ± 0.0173 (+)	1.2080 ± 0.0387 (+)	1.1058 ± 0.0180 (+)	0.0029 ± 0.0006	0.0119 ± 0.0017 (+)	0.0276 ± 0.0207 (+)	0.0250 ± 0.0139 (+)
MMF2	1.0817 ± 0.0120	1.1848 ± 0.0059 (+)	1.1944 ± 0.0322 (+)	1.1168 ± 0.0280 (+)	0.0070 ± 0.0031	0.0207 ± 0.0034 (+)	0.0325 ± 0.0238 (+)	0.0300 ± 0.0182 (+)
MMF3	1.0792 ± 0.0322	1.1739 ± 0.0043 (+)	1.1873 ± 0.0398 (+)	1.1089 ± 0.0212 (+)	0.0069 ± 0.0023	0.0154 ± 0.0025 (+)	0.0263 ± 0.0308 (+)	0.0229 ± 0.0126 (+)
MMF4	1.5234 ± 0.0003	1.8620 ± 0.0021 (+)	1.8577 ± 0.0012 (+)	1.5241 ± 0.0004 (+)	0.0018 ± 0.0002	0.0037 ± 0.0004 (+)	0.0032 ± 0.0002 (+)	0.0024 ± 0.0002 (+)
MMF5	1.0734 ± 0.0006	1.1485 ± 0.0006 (+)	1.1488 ± 0.0015 (+)	1.0739 ± 0.0003 (+)	0.0024 ± 0.0001	0.0037 ± 0.0001 (+)	0.0039 ± 0.0007 (+)	0.0028 ± 0.0002 (+)
MMF6	1.0732 ± 0.0003	1.1483 ± 0.0009 (+)	1.1486 ± 0.0016 (+)	1.0738 ± 0.0006 (+)	0.0023 ± 0.0001	0.0035 ± 0.0002 (+)	0.0036 ± 0.0003 (+)	0.0026 ± 0.0002 (+)
MMF7	1.0731 ± 0.0002	1.1484 ± 0.0009 (+)	1.1498 ± 0.0011 (+)	1.0736 ± 0.0003 (+)	0.0022 ± 0.0001	0.0037 ± 0.0003 (+)	0.0039 ± 0.0003 (+)	0.0027 ± 0.0003 (+)
MMF8	1.7915 ± 0.0012	2.4065 ± 0.0164 (+)	2.3813 ± 0.0025 (+)	1.7920 ± 0.0014 (∼)	0.0025 ± 0.0001	0.0048 ± 0.0002 (+)	0.0040 ± 0.0004 (+)	0.0025 ± 0.0001 (∼)
MMF9	0.0820 ± 0.0000	0.1034 ± 0.0000 (+)	0.1034 ± 0.0000 (+)	0.0820 ± 0.0000 (∼)	0.0085 ± 0.0007	0.0160 ± 0.0014 (+)	0.0141 ± 0.0012 (+)	0.0108 ± 0.0007 (+)
MMF10	0.0678 ± 0.0000	0.0679 ± 0.0000 (+)	0.0680 ± 0.0001 (+)	0.0678 ± 0.0001 (∼)	0.0061 ± 0.0009	0.1128 ± 0.0230 (+)	0.1446 ± 0.0660 (+)	0.0074 ± 0.0000 (+)
MMF11	0.0581 ± 0.0000	0.0581 ± 0.0000 (∼)	0.0581 ± 0.0000 (∼)	0.0581 ± 0.0000 (∼)	0.0082 ± 0.0004	0.0176 ± 0.0018 (+)	0.0136 ± 0.0014 (+)	0.0107 ± 0.0008 (+)
MMF12	0.5431 ± 0.0000	0.5452 ± 0.0014 (+)	0.5598 ± 0.0492 (∼)	0.5430 ± 0.0000 (−)	0.0020 ± 0.0001	0.0068 ± 0.0006 (+)	0.0110 ± 0.0187 (+)	0.0020 ± 0.0001 (∼)
MMF13	0.0444 ± 0.0000	0.0444 ± 0.0000 (∼)	0.0444 ± 0.0000 (∼)	0.0444 ± 0.0000 (∼)	0.0063 ± 0.0014	0.0264 ± 0.0076 (+)	0.0121 ± 0.0036 (+)	0.0089 ± 0.0014 (+)
Omni-test	0.0518 ± 0.0000	0.0190 ± 0.0000 (−)	0.0189 ± 0.0000 (−)	0.0518 ± 0.0000 (∼)	0.0091 ± 0.0015	0.0422 ± 0.0034 (+)	0.0080 ± 0.0005 (−)	0.0100 ± 0.0021 (∼)
SYM-PART simple	0.0520 ± 0.0000	0.0605 ± 0.0001 (+)	0.0601 ± 0.0000 (+)	0.0520 ± 0.0000 (∼)	0.0165 ± 0.0039	0.0419 ± 0.0044 (+)	0.0127 ± 0.0014 (−)	0.0109 ± 0.0013 (−)
SYM-PART rotated	0.0520 ± 0.0000	0.0606 ± 0.0001 (+)	0.0601 ± 0.0000 (+)	0.0520 ± 0.0000 (∼)	0.0178 ± 0.0047	0.0467 ± 0.0058 (+)	0.0152 ± 0.0022 (−)	0.0159 ± 0.0040 (∼)
LORD vs. others (+/ − / ∼)		15/1/2	14/1/3	8/1/9	(+/ − / ∼)	18/0/0	15/3/0	13/1/4

From Tables 5.6 and 5.7, the following insights are obtained for LORD:

- LORD is superior to DN-NSGA-II [113] as DN-NSGA-II neglects the solution diversity in the objective space. Thus, the solution distribution also suffers in the decision space by the neighborhood property (Theorem 4.1).

- While NSGA-II is the second-best in the objective space (Table 5.7), it neglects the solution diversity in the decision space and thus, gets outperformed by LORD.

- While MO_Ring_PSO_SCD is the second-best in the decision space (Table 5.6), it often gets trapped in the local optima (Fig. 5.4) leading to poor performance for some MMMOPs (e.g., MMF11 and MMF12). As LORD efficiently addresses the crowding illusion problem (Section 5.3), it has superior performance in most cases.

- The performance of LORD remains consistent (Tables 5.6 and 5.7), even for high #PSs (e.g., Omni-test with #PSs=27 in Fig. 5.4). It can successfully overcome the

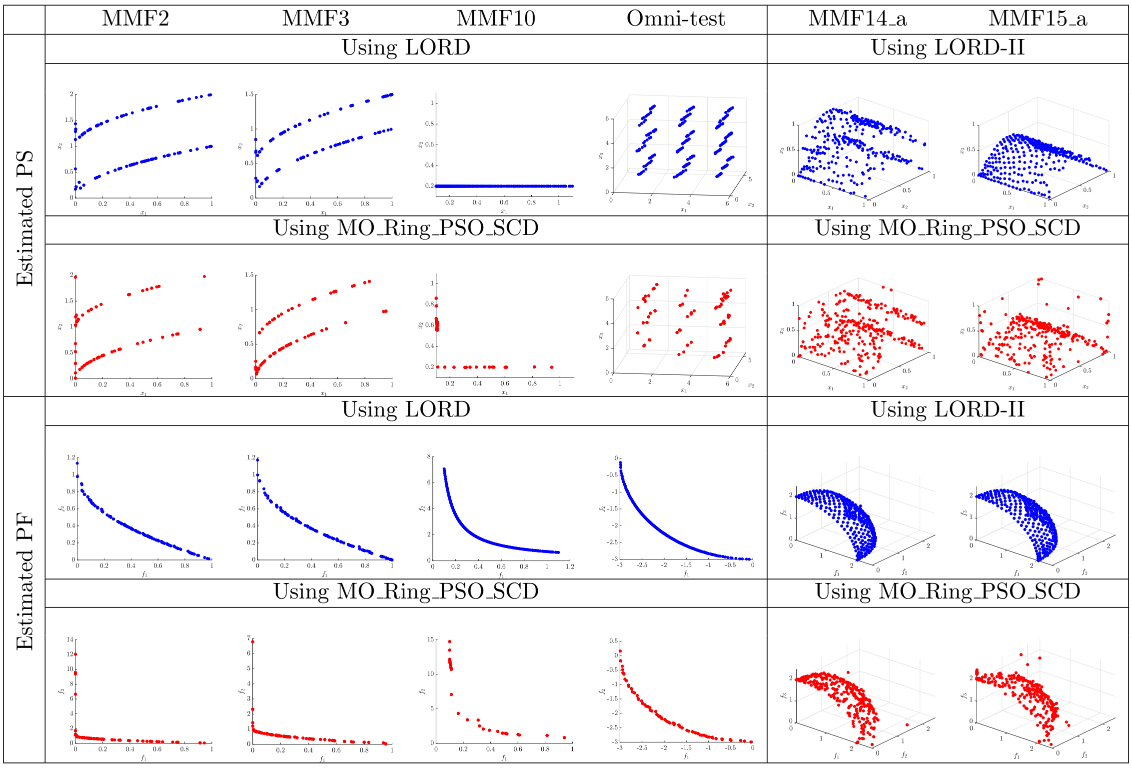


Figure 5.4: Estimated PSs and PFs for some 2- and 3-objective MMMOPs [140].

local optima (e.g., MMF10 in Fig. 5.4) and thus, also, acts as an excellent MOEA. As cover rate (CoRa) is nearly equal to one (ideal value as mentioned in Section 1.3.3), Table 5.6 reflects rPSP (=IGDX/CoRa) to be similar to IGDX.

From Tables 5.8 and 5.9, the following insights are obtained for LORD-II:

- In the objective space (Table 5.9), both LORD-II and MOEA/DD have similar performance for 3-objective problems. For 5-objective problems, LORD-II is marginally outperformed in only one case by MOEA/DD. For 8- and 10-objective problems, LORD-II is superior. In all cases, LORD-II outperforms MO_Ring_PSO_SCD in both convergence (CM) and diversity (*D_metric*) as also seen in Figs. 5.4 and 5.5.

- In the decision space (Table 5.8), LORD-II maintains superiority.

- The estimated PS and PF from LORD-II (Figs. 5.4 and 5.5) demonstrate excellent convergence and diversity. The results from MO_Ring_PSO_SCD deteriorate severely with an increase in dimension (Fig. 5.5). In contrast to MO_Ring_PSO_SCD (Fig. 5.5l), the polar plot [68] from LORD-II (Fig. 5.5f) converges all solutions to a near-global PF, forming a uniformly distributed circle for 8-objective MMMaOPs.

Table 5.8: Mean and standard deviation of NSX and CM_NSX over 51 independent runs for comparing LORD-II on M-objective MMMOPs with $M \geq 3$ [140].

Problems (M)	**NSX**			**CM_NSX**		
	LORD-II	**MO_Ring_PSO_SCD**	**MOEA/DD**	**LORD-II**	**MO_Ring_PSO_SCD**	**MOEA/DD**
MMF14 (3)	0.0068 ± 0.0052	0.2400 ± 0.0262(+)	0.0133 ± 0.0024 (+)	0.0203 ± 0.0018	0.1078 ± 0.0088 (+)	0.0228 ± 0.0019 (+)
MMF14_a (3)	0.0251 ± 0.0290	0.2533 ± 0.0320 (+)	0.1333 ± 0.0259 (+)	0.1498 ± 0.0518	0.1534 ± 0.0147 (+)	0.2481 ± 0.0103 (+)
MMF15 (3)	0.0205 ± 0.0073	0.4033 ± 0.0361 (+)	0.0400 ± 0.0024 (+)	0.0209 ± 0.0010	0.2356 ± 0.0251 (+)	0.0265 ± 0.0039 (∼)
MMF15_a (3)	0.0179 ± 0.0076	0.3400 ± 0.0262 (+)	0.0567 ± 0.0024 (+)	0.0270 ± 0.0180	0.2069 ± 0.0263 (+)	0.0454 ± 0.0098 (+)
MMF14 (5)	0.4838 ± 0.0125	0.5140 ± 0.0168 (+)	0.5152 ± 0.0057 (+)	0.1830 ± 0.0015	0.2363 ± 0.0047 (+)	0.1874 ± 0.0015 (+)
MMF14_a (5)	0.4855 ± 0.0141	0.4960 ± 0.0175 (∼)	0.5232 ± 0.0029 (+)	0.2080 ± 0.0155	0.2809 ± 0.0041 (+)	0.2474 ± 0.0044 (+)
MMF15 (5)	0.4959 ± 0.0076	0.5600 ± 0.0155 (+)	0.5172 ± 0.0014 (+)	0.1559 ± 0.0015	0.4118 ± 0.0152 (+)	0.1602 ± 0.0005 (∼)
MMF15_a (5)	0.4969 ± 0.0086	0.5660 ± 0.0155 (+)	0.5232 ± 0.0071 (+)	0.1783 ± 0.0062	0.3672 ± 0.0128 (+)	0.1938 ± 0.0078 (+)
MMF14 (8)	0.2772 ± 0.0012	0.5663 ± 0.0129 (+)	0.3088 ± 0.0018 (+)	0.4025 ± 0.0011	0.4386 ± 0.0070 (+)	0.4178 ± 0.0011 (+)
MMF14_a (8)	0.2796 ± 0.0008	0.5588 ± 0.0114 (+)	0.2900 ± 0.0062 (+)	0.4222 ± 0.0004	0.4480 ± 0.0045 (+)	0.4335 ± 0.0021 (+)
MMF15 (8)	0.2524 ± 0.0002	0.6438 ± 0.0175 (+)	0.2713 ± 0.0018 (+)	0.3537 ± 0.0052	0.5312 ± 0.0059 (+)	0.3815 ± 0.0015 (+)
MMF15_a (8)	0.2430 ± 0.0137	0.6113 ± 0.0093 (+)	0.2688 ± 0.0027 (+)	0.3797 ± 0.0115	0.4885 ± 0.0078 (+)	0.3886 ± 0.0068 (∼)
MMF14 (10)	0.2595 ± 0.0005	0.5880 ± 0.0101 (+)	0.2620 ± 0.0038 (+)	0.5088 ± 0.0087	0.5584 ± 0.0046 (+)	0.5356 ± 0.0085 (+)
MMF14_a (10)	0.2717 ± 0.0106	0.5930 ± 0.0134 (+)	0.2718 ± 0.0083 (∼)	0.5165 ± 0.0008	0.5699 ± 0.0036 (+)	0.5393 ± 0.0063 (+)
MMF15 (10)	0.2557 ± 0.0023	0.6590 ± 0.0123 (+)	0.2481 ± 0.0030 (−)	0.4516 ± 0.0096	0.6180 ± 0.0071 (+)	0.4850 ± 0.0005 (+)
MMF15_a (10)	0.2664 ± 0.0143	0.6350 ± 0.0119 (+)	0.2652 ± 0.0098 (∼)	0.4731 ± 0.0061	0.5780 ± 0.0048 (+)	0.4903 ± 0.0006 (+)
LORD-II vs. others (+/ − / ∼)		15/0/1	13/1/2	(+/ − / ∼)	16/0/0	13/0/3

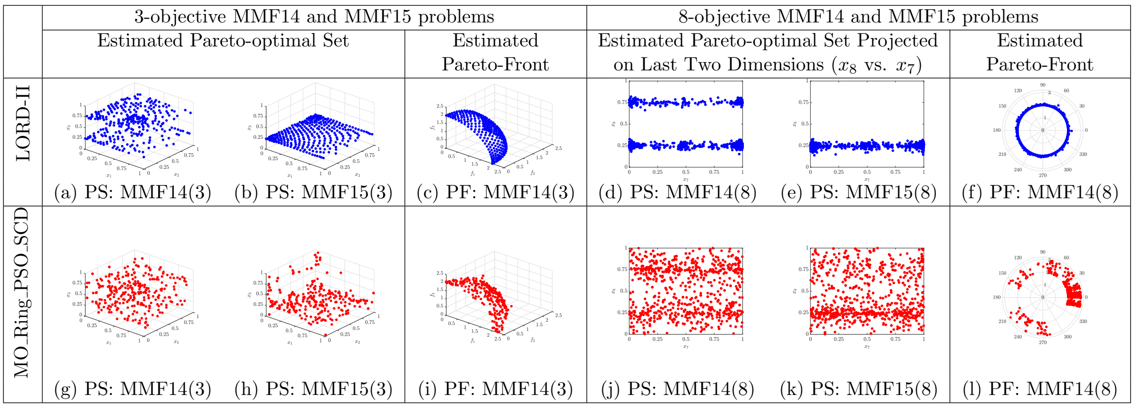


Figure 5.5: Estimated PSs and PFs of MMF14 and MMF15 problems, as both MMF14 and MMF15 problems have similar PFs, only the PFs of MMF14 are shown [140].

2) **Experiment-II: Comparison with Reference Vector Assisted MMMOEAs**: The performance of two recent reference-vector assisted MMMOEAs (DE-TriM [137] and MM-NAEMO [120]) are compared with LORD and LORD-II in Table 5.10 on CEC 2019 MMMOPs [112]. The results of this experiment are also compared with MO_Ring_PSO_SCD

Table 5.9: Mean and standard deviation of D_metric and CM over 51 independent runs for comparing LORD-II on M-objective MMMOPs with $M \geq 3$ [140].

Problems (M)	**D_metric**			**CM**		
	LORD-II	**MO_Ring_ PSO_SCD**	**MOEA/DD**	**LORD-II**	**MO_Ring_ PSO_SCD**	**MOEA/DD**
MMF14 (3)	0.0000 ± 0.0000	21.3266 ± 2.0086 (+)	0.0000 ± 0.0000 (∼)	0.0419 ± 0.0003	0.1083 ± 0.0130 (+)	0.0419 ± 0.0002 (∼)
MMF14_a (3)	0.0000 ± 0.0000	22.2752 ± 2.2816 (+)	0.0000 ± 0.0000 (∼)	0.0435 ± 0.0007	0.0949 ± 0.0149 (+)	0.0438 ± 0.0010 (∼)
MMF15 (3)	0.0000 ± 0.0000	24.4073± 3.8341 (+)	0.0000 ± 0.0000 (∼)	0.0422 ± 0.0004	0.1471 ± 0.0161 (+)	0.0426 ± 0.0004 (∼)
MMF15_a (3)	0.0000 ± 0.0000	22.5315± 3.5845(+)	0.0000 ± 0.0000 (∼)	0.0445 ± 0.0007	0.1322 ± 0.0207 (+)	0.0449 ± 0.0008 (∼)
MMF14 (5)	0.0000 ± 0.0000	43.9023 ± 4.6565 (+)	0.0000 ± 0.0000 (∼)	0.0590 ± 0.0020	0.4121 ± 0.0177 (+)	0.0587 ± 0.0015 (∼)
MMF14_a (5)	0.0000 ± 0.0000	46.7494 ± 2.8893 (+)	0.9428 ± 0.8165 (+)	0.0781 ± 0.0022	0.3659 ± 0.0115 (+)	0.0827 ± 0.0021 (+)
MMF15 (5)	0.0000 ± 0.0000	43.6883 ± 3.9734 (+)	0.0000 ± 0.0000 (∼)	0.0654 ± 0.0017	0.4610 ± 0.0152 (+)	0.0625 ± 0.0025 (−)
MMF15_a (5)	0.0000 ± 0.0000	45.2327 ± 3.4695 (+)	0.9428 ± 0.8165 (+)	0.0954 ± 0.0045	0.4339 ± 0.0162 (+)	0.0961 ± 0.0053 (∼)
MMF14 (8)	0.0000 ± 0.0000	101.1673 ± 4.3256 (+)	5.3833 ± 0.0000 (+)	0.1332 ± 0.0002	0.6277 ± 0.0154 (+)	0.1456 ± 0.0016 (+)
MMF14_a (8)	0.0000 ± 0.0000	106.8644 ± 3.5223 (+)	5.3833 ± 0.0000 (+)	0.1742 ± 0.0044	0.5917 ± 0.0149 (+)	0.1817 ± 0.0044 (+)
MMF15 (8)	0.0000 ± 0.0000	99.5550 ± 3.0343 (+)	5.4810 ± 0.1382 (+)	0.1288 ± 0.0038	0.6767 ± 0.0128 (+)	0.1508 ± 0.0008 (+)
MMF15_a (8)	0.0000 ± 0.0000	103.5948 ± 3.5180 (+)	5.5787 ± 0.0000 (+)	0.2014 ± 0.0025	0.6488 ± 0.0151 (+)	0.2122 ± 0.0284 (+)
MMF14 (10)	14.1331 ± 2.5516	132.4681 ± 4.6720 (+)	38.9838 ± 0.7255 (+)	0.2200 ± 0.0037	0.6575 ± 0.0129 (+)	0.2504 ± 0.0002 (+)
MMF14_a (10)	21.9290 ± 0.4837	137.8081 ± 3.8224 (+)	41.7357 ± 0.5083 (+)	0.2554 ± 0.0065	0.6311 ± 0.0128 (+)	0.2966 ± 0.0041 (+)
MMF15 (10)	17.6340 ± 1.4436	131.8568 ± 4.0914 (+)	36.4571 ± 9.3681 (+)	0.2219 ± 0.0118	0.7072 ± 0.0109 (+)	0.2535 ± 0.0072 (+)
MMF15_a (10)	23.2171 ± 2.4365	133.7104 ± 2.8705 (+)	39.7423 ± 1.7614 (+)	0.2672 ± 0.0037	0.6812 ± 0.0092 (+)	0.3193 ± 0.0007 (+)
LORD-II vs. others (+/ − / ∼)		16/0/0	10/0/6	(+/ − / ∼)	16/0/0	9/1/6

to fairly assess the relative rankings of algorithms. Each of these algorithms (DE-TriM, MM-NAEMO and MO_Ring_PSO_SCD) are set up using the parameters recommended in [137], [120] and [188], respectively.

From Table 5.10, LORD-II is noted to have the best performance in all cases and LORD is noted to have the best or the second-best performance in both objective and decision spaces for most of the cases. Unlike other MMMOEAs [56, 120, 188] which yield poor performance in objective space in order to improve the performance in decision space, LORD and LORD-II perform satisfactorily in both the spaces and competitively outperform the other reference vector assisted MMMOEAs.

3) **Experiment-III: Comparison on Polygon MMMaOPs**: Similar to [170], the mean IGDX and IGDF of LORD-II are compared with NIMMO on Polygon test problems as both the MMMOEAs are designed for MMMaOPs. The results of MO_Ring_PSO_SCD [188], Omni-Optimizer [51] and TriMOEA_TA&R [118] are also compared.

Table 5.10: Mean of IGDX and IGDF over 51 independent runs for comparing reference-vector guided MMMOEAs on 2- and 3-objective MMMOPs [140].

	IGDX				IGDF			
2-objective Problems	**LORD**	**DE-TriM**	**MM-NAEMO**	**MO_Ring_PSO_SCD**	**LORD**	**DE-TriM**	**MM-NAEMO**	**MO_Ring_PSO_SCD**
MMF1	0.0431	0.0465 (+)	0.0486 (+)	0.0485 (+)	0.0025	0.0026 (∼)	0.0040 (+)	0.0037 (+)
MMF1_z	0.0351	0.0503 (+)	0.0347 (∼)	0.0352 (∼)	0.0022	0.0026 (+)	0.0035 (+)	0.0036 (+)
MMF1_e	0.7499	2.8757 (+)	0.4115 (−)	0.4738 (−)	0.0029	0.0029 (∼)	0.0051 (+)	0.0119 (+)
MMF2	0.0180	0.0505 (+)	0.0118 (−)	0.0416 (+)	0.0070	0.0035 (−)	0.0083 (+)	0.0207 (+)
MMF3	0.0176	0.0235 (+)	0.0137 (−)	0.0276 (+)	0.0069	0.0047 (−)	0.0085 (+)	0.0154 (+)
MMF4	0.0251	0.0211 (−)	0.0312 (+)	0.0271 (+)	0.0018	0.0025 (+)	0.0033 (+)	0.0037 (+)
MMF5	0.0814	0.0892 (+)	0.0871 (+)	0.0857 (+)	0.0024	0.0027 (+)	0.0037 (+)	0.0037 (+)
MMF6	0.0692	0.0756 (+)	0.0743 (+)	0.0736 (+)	0.0023	0.0025 (∼)	0.0036 (+)	0.0035 (+)
MMF7	0.0218	0.0201 (−)	0.0229 (+)	0.0262 (+)	0.0022	0.0025 (+)	0.0035 (+)	0.0037 (+)
MMF8	0.0762	0.0989 (+)	0.3348 (+)	0.0673 (∼)	0.0025	0.0029 (+)	0.0037 (+)	0.0048 (+)
MMF9	0.0046	0.0787 (+)	0.0048 (∼)	0.0079 (+)	0.0085	0.0119 (+)	0.0479 (+)	0.0160 (+)
MMF10	0.0018	0.0018 (∼)	0.0121 (+)	0.0276 (+)	0.0061	0.0080 (+)	0.0639 (+)	0.1128 (+)
MMF11	0.0029	0.0036 (+)	0.0418 (+)	0.0054 (+)	0.0082	0.0109 (+)	0.0931 (+)	0.0176 (+)
MMF12	0.0013	0.0013 (∼)	0.0050 (+)	0.0038 (+)	0.0020	0.0021 (∼)	0.0196 (+)	0.0068 (+)
MMF13	0.0242	0.0368 (+)	0.1878 (+)	0.0314 (+)	0.0063	0.0094 (+)	0.1059 (+)	0.0264 (+)
Omni-test	0.0706	0.0732 (+)	0.1511 (+)	0.3907 (+)	0.0091	0.0125 (+)	0.0130 (+)	0.0422 (+)
SYM-PART-simple	0.0549	0.0740 (+)	0.1115 (+)	0.0300 (−)	0.0165	0.0101 (−)	0.0472 (+)	0.0419 (+)
SYM-PART-rotated	0.1558	0.1885 (+)	0.7586 (+)	0.2926 (+)	0.0178	0.0125 (−)	0.0395 (+)	0.0467 (+)
LORD vs. others (+/ − / ∼)		14/2/2	13/3/2	14/2/2	(+/ − / ∼)	10/4/4	18/0/0	18/0/0
3-objective Problems	**LORD-II**	**DE-TriM**	**MM-NAEMO**	**MO_Ring_PSO_SCD**	**LORD-II**	**DE-TriM**	**MM-NAEMO**	**MO_Ring_PSO_SCD**
MMF14	0.0443	0.0558 (+)	0.0465 (+)	0.0539 (+)	0.0540	0.0749 (+)	0.0808 (+)	0.0801 (+)
MMF14_a	0.0576	0.0676 (+)	0.0663 (+)	0.0613 (+)	0.0561	0.0809 (+)	0.0791 (+)	0.0789 (+)
MMF15	0.0287	0.0361 (+)	0.0518 (+)	0.0419 (+)	0.0548	0.0787 (+)	0.1113 (+)	0.0854 (+)
MMF15_a	0.0355	0.0503 (+)	0.0848 (+)	0.0452 (+)	0.0571	0.0951 (+)	0.1263 (+)	0.0841 (+)
LORD-II vs. others (+/ − / ∼)		4/0/0	4/0/0	4/0/0	(+/ − / ∼)	4/0/0	4/0/0	4/0/0

Table 5.11: Specifications for the experiment in [140] conducted on polygon and rotated polygon problems according to specifications of [170].

Parameters		Used by LORD-II in [140]	Used by NIMMO, TriMOEA_TA&R, MO_Ring_PSO_SCD, and Omni-Optimizer in [170]
n_{pop}	3-obj	210	210
	5-obj	210	210
	8-obj	156	156
	10-obj	230	230
#runs		31	31
$MaxFES$		10000	10000
N_{IGD}		5000	5000

The performance of these MMMOEAs are noted on M-objective polygon and rotated polygon MMMaOPs [76] in Tables 5.12 (using IGDX) and 5.13 (using IGDF). This experiment considers the specifications mentioned in Table 5.11 as per [170]. The performance values of the other MMMOEAs (except LORD-II) are also noted from [170]. The remaining parameters of LORD-II are set up as specified in Table 5.3.

From Tables 5.12 and 5.13, LORD-II is observed to be superior in both decision and objective spaces, respectively. The performance of all MMMOEAs (except TriMOEA_TA&R) are unaffected due to rotation. However, the IGDX values of TriMOEA_TA&R are widely different (poorer) for rotated polygon problems from those of the polygon problems (Table 5.12). This difference arises as TriMOEA_TA&R considers only the number of solutions as the diversity criteria and neglects the solution distribution in the decision space [118].

Table 5.12: Mean IGDX over 31 independent runs for comparing LORD-II on M-objective polygon and rotated polygon (RPolygon) problems [140].

M-Problems	**LORD-II**	**NIMMO**	**TriMOEA _TA&R**	**MO_Ring_ PSO_SCD**	**Omni-Optimizer**
3-Polygon	0.0054	0.0056 (+)	0.0063 (+)	0.0091 (+)	0.0083 (+)
3-RPolygon	0.0064	0.0059 (−)	0.0295 (+)	0.0090 (+)	0.0085 (+)
5-Polygon	0.0055	0.0070 (+)	0.0162 (+)	0.0113 (+)	0.0110 (+)
5-RPolygon	0.0062	0.0074 (+)	0.0400 (+)	0.0113 (+)	0.0110 (+)
8-Polygon	0.0046	0.0089 (+)	0.0136 (+)	0.0143 (+)	0.0140 (+)
8-RPolygon	0.0051	0.0093 (+)	0.0747 (+)	0.0144 (+)	0.0138 (+)
10-Polygon	0.0044	0.0072 (+)	0.0123 (+)	0.0120 (+)	0.0112 (+)
10-RPolygon	0.0053	0.0076 (+)	0.0404 (+)	0.0118 (+)	0.0112 (+)
LORD-II vs. others (+/ − / ∼)		7/1/0	8/0/0	8/0/0	8/0/0

Table 5.13: Mean IGDF over 31 independent runs for comparing LORD-II on M-objective polygon and rotated polygon (RPolygon) problems [140].

M-Problems	**LORD-II**	**NIMMO**	**TriMOEA _TA&R**	**MO_Ring_ PSO_SCD**	**Omni-Optimizer**
3-Polygon	0.0023	0.0025 (+)	0.0040 (+)	0.0034 (+)	0.0028 (+)
3-RPolygon	0.0023	0.0025 (+)	0.0046 (+)	0.0034 (+)	0.0028 (+)
5-Polygon	0.0031	0.0044 (+)	0.0149 (+)	0.0057 (+)	0.0051 (+)
5-RPolygon	0.0030	0.0044 (+)	0.0149 (+)	0.0058 (+)	0.0052 (+)
8-Polygon	0.0031	0.0069 (+)	0.0180 (+)	0.0092 (+)	0.0082 (+)
8-RPolygon	0.0032	0.0069 (+)	0.0190 (+)	0.0093 (+)	0.0083 (+)
10-Polygon	0.0033	0.0064 (+)	0.0204 (+)	0.0087 (+)	0.0074 (+)
10-RPolygon	0.0034	0.0064 (+)	0.0185 (+)	0.0086 (+)	0.0075 (+)
LORD-II vs. others (+/ − / ∼)		8/0/0	8/0/0	8/0/0	8/0/0

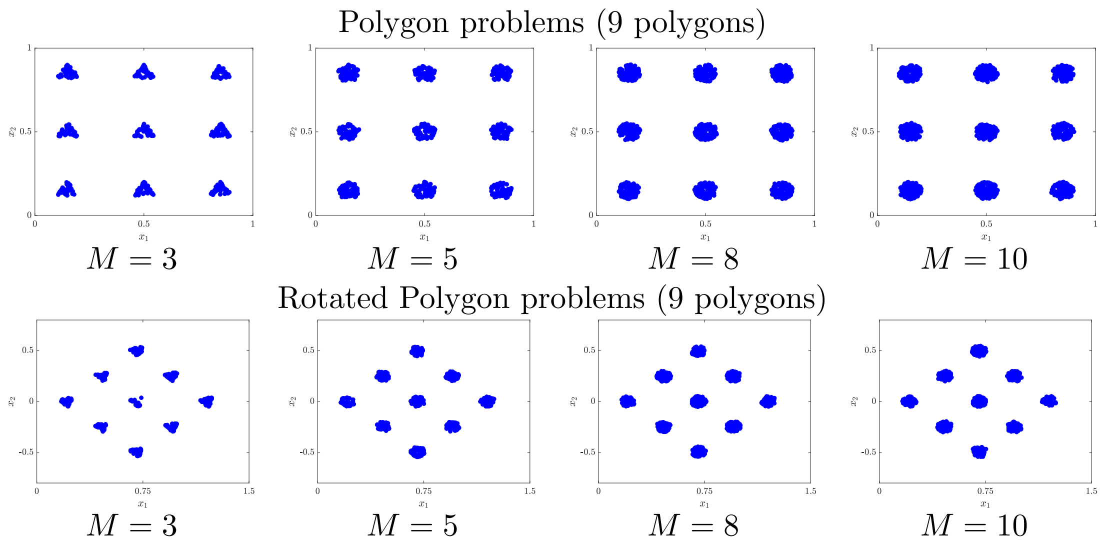
The estimated PSs from LORD-II are shown in Table 5.14 from which the following observations are noted:

- For all the 8 instances, LORD-II converges to global surfaces without any outliers.
- The number of solutions per subset is relatively uniform over the 9 subsets in PS.
- For both polygon and rotated polygon problems, the shape of the polygon is properly replicated for 3- and 5-objective problems. For 8- and 10-objective problem, a near-spherical blob (of unidentifiable shape) is formed at each of the subsets in PS.

4) **Experiment-IV: Comparison by Variation in Population Size**: While a large population size (n_{pop}) is a necessity for MMMOPs (as mentioned in Section 1.3.5), standard MOEAs such as MOEA/DD may have poor performance due to a large n_{pop}. For a fair assessment on the superiority of LORD-II, this experiment compares LORD-II with MOEA/DD using both small n_{pop} (as per the optimal setting of MOEA/DD in [109]) and large n_{pop} (= $100 \times N$ as per the recommendation of CEC 2019 MMMOPs [112]) in Table 5.15, from which the following insights are obtained:

- LORD-II is superior even for small n_{pop}.

Table 5.14: Estimated PSs from LORD-II for M-objective polygon and rotated polygon problems [140].



- While MOEA/DD never outperforms LORD-II in the decision space, the former is marginally superior for a few cases (one out of 16 cases for small n_{pop} and two out of 16 cases for large n_{pop}) in the objective space.

- A large n_{pop} improves IGDX and IGDF (Eq. (1.11)) regardless of the effectiveness of the underlying algorithm [170], as also observed in Table 5.15 for both LORD-II and MOEA/DD. However, since the superiority of LORD-II against MOEA/DD is also established for a small n_{pop}, these results indeed reflect the efficient synergism of various strategies in the evolutionary framework of LORD-II.

Table 5.15: Mean of IGDX and IGDF over 51 independent runs with different population sizes (n_{pop}) for M-objective MMMOPs [140].

Problems	M	**Recommended Population Size for MOEA/DD in [109]**					**Recommended Population Size for MMMOPs in [112]**				
			IGDX		**IGDF**			**IGDX**		**IGDF**	
		n_{pop}	**LORD-II**	**MOEA/DD**	**LORD-II**	**MOEA/DD**	n_{pop}	**LORD-II**	**MOEA/DD**	**LORD-II**	**MOEA/DD**
MMF14	3	91	0.0832	0.2150 (+)	0.1044	0.1045 (∼)	300	0.0443	0.0671 (+)	0.0540	0.0555 (+)
MMF14_a	3	91	0.1150	0.2076 (+)	0.1044	0.1045 (∼)	300	0.0576	0.0780 (+)	0.0561	0.0568 (+)
MMF15	3	91	0.0514	0.0522 (∼)	0.1055	0.1056 (∼)	300	0.0287	0.0295 (+)	0.0548	0.0562 (+)
MMF15_a	3	91	0.0638	0.0705 (+)	0.1056	0.1144 (+)	300	0.0355	0.0357 (∼)	0.0571	0.0607 (+)
MMF14	5	210	0.3070	0.3314 (+)	0.3125	0.3136 (+)	495	0.2448	0.2554 (+)	0.0564	0.0598 (+)
MMF14_a	5	210	0.3283	0.4083 (+)	0.3129	0.3135 (∼)	495	0.2670	0.2846 (+)	0.0752	0.0839 (+)
MMF15	5	210	0.2460	0.2652 (+)	0.3155	0.3167 (+)	495	0.1960	0.2032 (+)	0.0602	0.0645 (+)
MMF15_a	5	210	0.2695	0.2963 (+)	0.3181	0.3168 (−)	495	0.2155	0.2230 (+)	0.0895	0.0999 (+)
MMF14	8	156	0.6864	0.7006 (+)	0.7233	0.7244 (+)	828	0.5621	0.5857 (+)	0.1445	0.1494 (+)
MMF14_a	8	156	0.6851	0.7363 (+)	0.7225	0.7241 (+)	828	0.5725	0.5936 (+)	0.1776	0.1905 (+)
MMF15	8	156	0.6086	0.6263 (+)	0.7270	0.7291 (+)	828	0.5146	0.5586 (+)	0.1503	0.1486 (−)
MMF15_a	8	156	0.6543	0.6539 (∼)	0.7277	0.7289 (+)	828	0.5315	0.5498 (+)	0.2195	0.2159 (∼)
MMF14	10	275	0.8404	0.8847 (+)	0.6811	0.6864 (+)	935	0.7088	0.7373 (+)	0.3463	0.3102 (∼)
MMF14_a	10	275	0.8374	0.8972 (+)	0.6839	0.6907 (+)	935	0.6869	0.7241 (+)	0.4296	0.4317 (∼)
MMF15	10	275	0.7787	0.8105 (+)	0.6864	0.6903 (+)	935	0.6469	0.6731 (+)	0.3561	0.2984 (−)
MMF15_a	10	275	0.8074	0.8246 (+)	0.6913	0.6940 (+)	935	0.6712	0.6848 (+)	0.4375	0.4384 (∼)
LORD-II vs. MOEA/DD (+/ − / ∼)				14/0/2	(+/ − / ∼)	11/1/4	(+/ − / ∼)		15/0/1	(+/ − / ∼)	10/2/4

Thus, it is evident that the improved performance is an attribute of the algorithmic framework of LORD-II and not of the large n_{pop}.

5.5.6 Scalability Study on LORD-II framework

As most of the MMMOEAs are not tested on scalable problems [118], the scalability of LORD-II is established by studying its performance in Table 5.16 with variations in the candidate dimension (N) of a 3-objective MMF14 problem for $N = \{3, 10, 30, 50, 100\}$.

Table 5.16: Mean of rPSP, IGDX, rHV and IGDF for 3-objective MMF14 (with different candidate dimensions, N) over 51 Independent Runs of LORD-II [140].

N	rPSP	IGDX	rHV	IGDF
3	0.0449	0.0443	1.0395	0.0540
10	0.5928	0.5838	1.0414	0.0013
30	2.8270	1.5038	1.0402	0.0001
50	2.1513	2.1258	1.0405	0.0001
100	3.2476	3.1807	1.0406	0.0000

From Table 5.16, the following observations are noted:

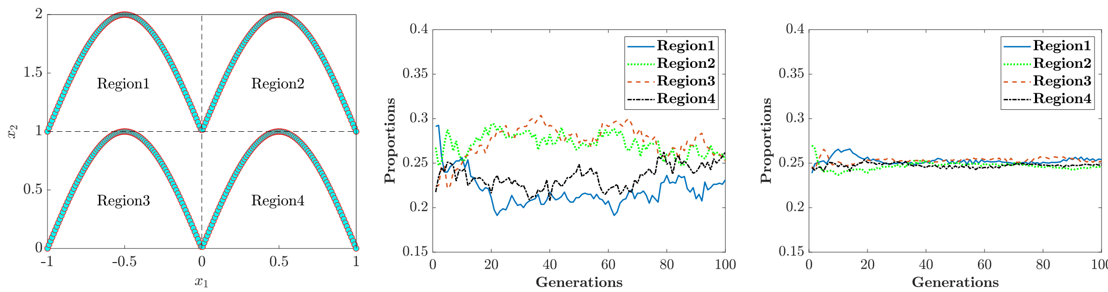
- As the number of objectives (M) does not change, the performance of LORD-II remains unaffected in the objective space as noted from the absence of any significant increase in rHV and IGDF.
- For small N, the performance in the decision space deteriorates only linearly (not exponentially) with an increase in N. For example, IGDX increases 34 times when N is increased from 3 to 30. However, with further increase in N, the deterioration in performance is even less drastic. For example, IGDX only doubles when N is increased from 30 to 100.

Thus, LORD-II, using decomposition of decision and objective spaces, works efficiently even for high-dimensional MMMOPs, i.e., LORD-II is scalable with problem size.

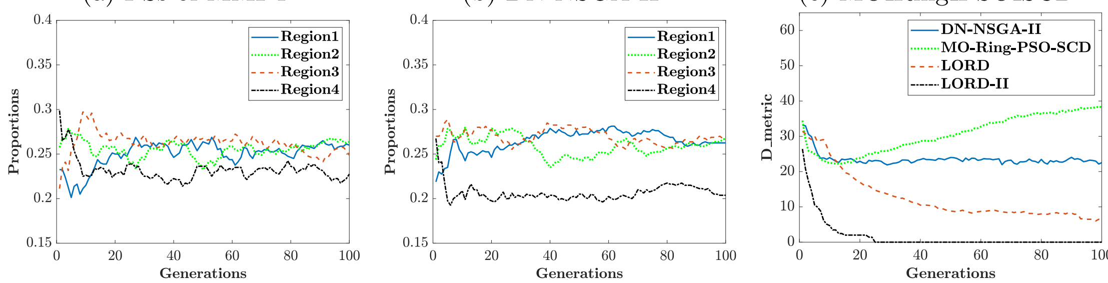
5.5.7 Population Dynamics in Decision and Objective Space

In this sub-section, the diversity attainment rates in the decision and objective spaces are analyzed for several MMMOEAs (LORD [140], LORD-II [140], MO_Ring_PSO_SCD [188] and DN-NSGA-II [113]).

For comparing the diversity attainment rate in the decision space, the time-evolution of the proportion of solutions in each of the four distinct regions of MMF4 (Fig. 5.6a) is considered with $n_{pop} = 800$ and $G_{max} = 100$ (as done in [188]). The mean proportions over 5 independent runs are plotted for the MMMOEAs (Fig. 5.6b to 5.6e). Ideally, these proportions should saturate at 25%. As seen from Figs. 5.6d and 5.6e, the proportions of solutions in 3 regions are steady near 25-27% and the proportion in region 4 fluctuates between 20 to 25% for LORD, whereas it is steady around 19-21% for LORD-II. Thus, the diversity attainment rate in the decision space of LORD and LORD-II is intermediate between that of MO_Ring_PSO_SCD and DN-NSGA-II.



(a) PSs of MMF4 (b) DN-NSGA-II (c) MO_Ring_PSO_SCD



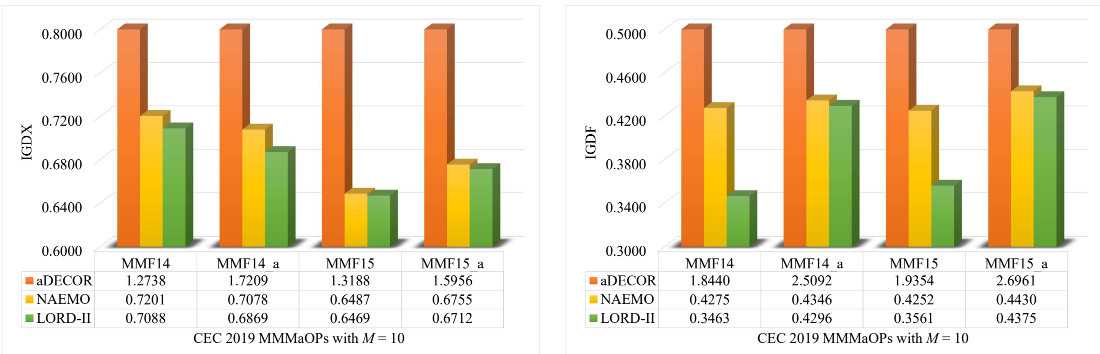
(d) LORD (e) LORD-II (f) D_metric

Figure 5.6: (a) True solution distribution of MMF4 problem in the decision space, (b-e) convergence behavior in the decision space for four algorithms: (b) DN-NSGA-II, (c) MO_Ring_PSO_SCD, (d) LORD and (e) LORD-II, (f) diversity attainment rate in the objective space using *D_metric* [140].

For comparing the diversity attainment rate in the objective space, the time-evolution of *D_metric* (Eq. (4.20)) is considered with $n_{dir} = 800$. The mean *D_metric* over 5 independent runs are plotted in Fig. 5.6f for all the four MMMOEAs. Ideally, *D_metric* should saturate at 0. As seen in Fig. 5.6f, *D_metric* for MO_Ring_PSO_SCD severely deteriorates with generations. This may be a result of the crowding illusion problem (Section 5.3). For LORD, a decreasing trend in *D_metric* is observed. For LORD-II, the *D_metric* has reached the ideal value roughly by 25 generations. These observations support the enhanced diversity preservation of the LORD and LORD-II in the objective space without sacrificing too much on the distribution in the decision space.

5.5.8 Comparing Partitioning Strategies for Many-Objective Problems

Over the previous chapters various partitioning strategies have been analyzed for addressing MaOO problems. Specifically, Chapter 2 presents DECOR [142] which consists of an objective reduction strategy to handle MaOO problems and Chapters 3 and 4 present ESOEA [138] and NAEMO [160], respectively, which explores decomposition of objective space for MaOO problems. As LORD-II [140] is based in decompostion in both objective and decision spaces for many-objective optimization problems, Fig. 5.7 compares its performance on 10-objective MMMaOPs against aDECOR [142] and NAEMO [160], in terms of IGDX and IGDF with $N_{IGD} = 1250$ points uniformly sampled from the true optimal surfaces. For fair comparison under the same function evaluation budget, DECOR, NAEMO and LORD-II are realized with $n_{pop} = n_{dir} = 935$ and $MaxFES = 50,000$.



(a) Decision Space (b) Objective Space

Figure 5.7: Mean IGDF and IGDX over 51 independent runs to compare objective reduction in aDECOR, decomposition of objective space in NAEMO and decomposition of decision space in LORD-II on 10-objective MMMaOPs where for better scaling the maximum limit of IGDX is considered as 0.8 and that of IGDF is considered as 0.5.

From Fig. 5.7, it is observed that the decomposition of objective space (used by both NAEMO and LORD-II) is hugely beneficial for MaOO problems. Additionally, decomposition of decision space (used by LORD-II) is beneficial for dealing with multi-modal problems like MMMF14 and MMF14_a (where #PSs > 1 as shown in Table 5.4). Nonetheless, across all the problems in this experiment, LORD-II demonstrates its superiority.

Thus, all the experiments, presented in this chapter, establish the efficacy of LORD and LORD-II for addressing a wide range of MOO problems (MMMOPs or otherwise).

5.6 Conclusion

As most of the existing MMMOEAs use crowding distance over the entire decision space, its analysis exhibits a major disadvantage which is identified as the crowding illusion problem. To mitigate the adverse effects of this problem for MMMOPs, a novel evolutionary framework is presented in this chapter. It is the first MMMOEA to consider the decomposition of decision space using graph Laplacian based clustering for maintaining the diversity of solutions in that space. It uses reference vectors to partition the objective space for maintaining diversity in the objective space. This algorithmic framework has two different versions to impart and explore the convergence attribute. The first version (LORD) is for MMMOPs with a small number of objectives, which eliminates the maximally crowded solution from the last non-dominated rank. The second version (LORD-II) is for problems with a high number of objectives which eliminates the candidate with maximal PBI, from the maximally large cluster. During the elimination of a candidate, LORD and LORD-II try to ensure that the removal does not occur from the sub-spaces (defined by reference vectors) with only one associated candidate. These frameworks have been tested over several MMMOPs and MMMaOPs and their performance have been compared with recent state-of-the-art algorithms to establish their efficacy.

While any multi- or many-objective optimization algorithm explores the search space to generate a set of multiple alternative solutions for different trade-offs among the objectives, an application problem can implement only one of these solutions. This selection of one of the solutions from the estimated set of Pareto-optimal solutions is often governed by domain knowledge or decision-maker's preferences. In the absence of such preferences or to resolve conflict among preferences from multiple decision-makers, formulating certain strategies is essential to guide the decision-making process for yielding the most relevant solution for a real-world optimization problem. This application-driven necessity motivates the research work in the next chapter, which is vital from the practical perspective of decision-making.

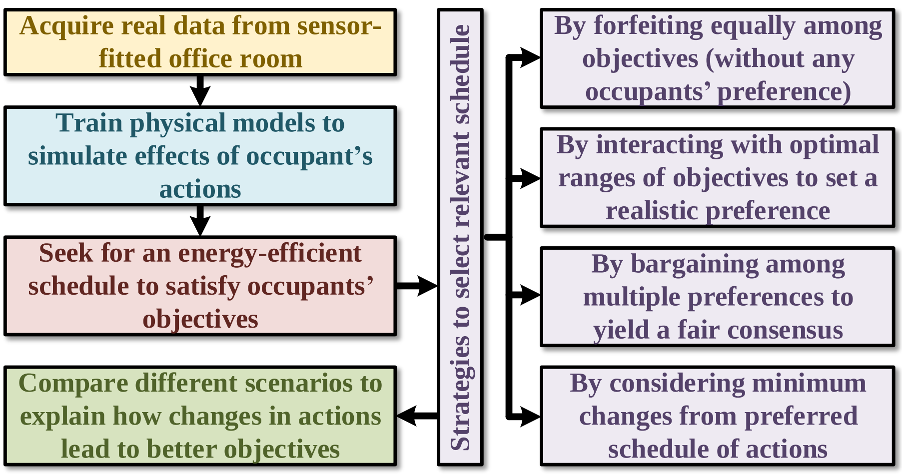
Chapter 6

Unmasking the Causal Relationships Latent in the Interplay Between Occupant's Actions and Room Ambience [133]

Outline

Objective: *To present various decision-making strategies for the real-world many-objective building energy management problem through a framework developed to associate the occupants (users) with building energy systems.*

Workflow:



6.1 Introduction

The Many-Objective Optimization (MaOO) problem of building energy management (Section 1.4) is explored in this chapter to study the research direction (identified in the previous chapter) of integrating decision-making strategies with the evolutionary algorithms for selecting the context-relevant implementable solution from the Pareto-optimal Set (PS).

It is crucial to satisfy an ever-growing energy demand with limited resources [132] in the building sector, as it occupies nearly 40% of the global energy consumption [183]. Building energy management through regulation of occupant behavior [147] is an effective strategy, even for existing (non-green) buildings. Preliminary works [53, 143] show that scheduling occupants' actions, like opening/closing doors and windows, could save energy. Thus, the physical factors stimulating these actions and the consequent effects [2] should be explained to the occupants, as these users are not domain-experts.

The crucial goals of building energy management [72] are to achieve premium indoor thermal comfort and finest indoor air quality without increasing energy consumption. This problem (Section 1.4) therefore generates several Pareto-optimal schedules of occupants' actions of which the most relevant one is chosen for implementation. This optimal schedule can explain the causal phenomenon leading up to the differences in actions (between optimal and actual schedules) and convey the temporal importance of a particular action.

Section 6.2 analyzes the state-of-the-art of building energy management approaches, Section 6.3 outlines the concerned framework followed by its four components: the experimental platform in Section 6.4, the building simulation models in Section 6.5, the optimization problem (with novel decision-making strategies) in Section 6.6 and the explanation generating framework in Section 6.7. Finally, Section 6.8 summarizes this chapter.

6.2 Research Gap Analysis

For the concerned approach of building energy management (Fig. 1.6), a solution [72] is characterized by (i) types of occupants' actions [147], (ii) specifications of building simulation models [64], and (iii) explaining the impacts of occupants' actions [2]. Brief surveys on each of these aspects are presented next.

Occupants' actions can greatly influence the indoor ambience. A few such actions include opening/closing windows [53] and doors [20], adjusting window blinds [130, 156], switching on/off heater [180] and lights [177], and plug load energy consumption [181].

For simulating the occupants' actions given the contextual information, various cognitive [104,194,196] and stochastic models [64,156] have been studied. Towards the middle of this decade, Multi-Objective Optimization (MOO) has also been considered [7, 57, 145] for placement of windows and solar panels. However, studies are scarce for identifying optimal occupants' actions concerning conflicting goals like comfort, economy, and ecology.

The vital effects of occupants' actions are thermal comfort [39, 145] and air-quality comfort [2, 132]. However, quantifying the influence of energy savings incurred through occupants' actions [72] and analysis with hourly granularity of decision variables for a holistic occupant-building interaction profile [145] continue to be challenging tasks.

This chapter contributes to the many-objective building energy management problem [133] by discussing various schedule selection strategies and causal explanation generating strategies to assist the occupants in learning and adopting the energy-efficient schedules.

6.2.1 Contributions of the Case Study

This case study [133] contributes to the domain of energy buildings in the following ways:

1. It aims to attain the occupants' action schedule for minimal indoor thermal dissatisfaction, minimal indoor air quality dissatisfaction, and minimal energy expenditure, while providing a minimal number of changes in recommended actions.

2. For a holistic exploration, it considers hourly granularity of data acquisition and system-generated recommendations along with performance analysis at different granularities (hourly, daily and seasonal).

3. It presents various decision-making strategies considering different types of occupants' preferences for selecting the most relevant energy-efficient schedule.

4. It identifies the driving forces behind the recommended actions. Such explanations are summarized to obtain the cause-and-effect chart. These justifications bridge the gap between recommendations and cognitive adaptations for the occupants.

Thus, this novel case study [133] assists the occupants to interact with the building energy system regarding their preferences and helps them to embrace the recommended energy-efficient schedules.

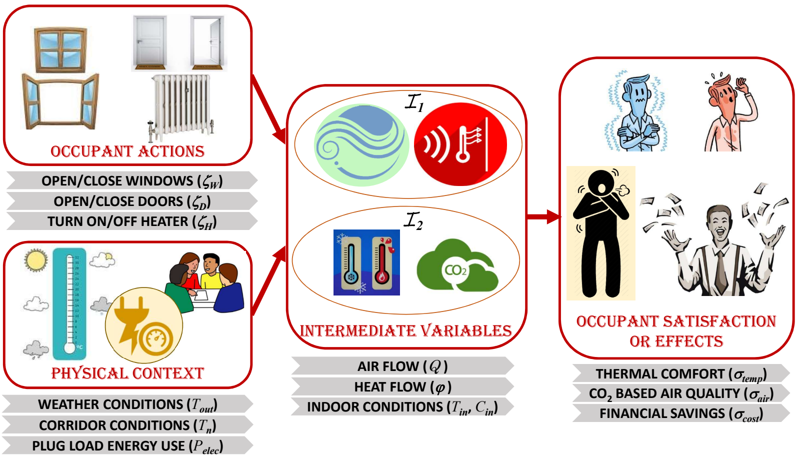


Figure 6.1: General schema for studying the impact of occupants' actions [133].

6.3 General Schema for Obtaining Explanations

Due to the complexity and formalism of building models and unconsciously varying schedules of occupants, it is highly challenging for the occupants to understand the underlying causal relations between their actions and effects. The concerned framework (Fig. 1.6) aims to simplify the recommended actions by associating them with contextual explanations. For this case study [133], the physical variables, involved in the building simulation models, can be grouped into the following categories (Fig. 6.1):

1. *Occupant actions* ($\mathcal{X}_B$): At time t, $\mathcal{X}_B$ contains variables directly controllable by the occupants, like opening/closing doors ($\zeta_D(t)$) and windows ($\zeta_W(t)$), and switching on/off a room heater ($\zeta_H(t)$).

2. *Physical context* ($\mathcal{P}_B$): At time t, $\mathcal{P}_B$ contains variables, which cannot be controlled by the occupants, like outdoor temperature ($T_{out}(t)$), wind speed, humidity, illuminance, temperature of neighboring zones ($T_n(t)$), number of occupants ($n(t)$) and electric power consumption from work-associated appliances ($P_{elec}(t)$).

3. *Occupant satisfaction* ($\mathcal{F}_B$): At time t, $\mathcal{F}_B$ contains variables desired by the occupants, like indicators of thermal discomfort ($\sigma_{temp}(t)$), aeraulic discomfort ($\sigma_{air}(t)$), the heater energy cost ($\sigma_{cost}(t)$), and changes in successive recommendations ($\delta_{WD}(t)$).

4. *Intermediate variables* ($\mathcal{I}_B$): At time t, the set of auxiliary variables $\mathcal{I}_B$ contains some model-estimated parameters ($\mathcal{I}_1$), like airflow ($Q(t)$) and heat flow ($\varphi(t)$), along with some sensor-recorded parameters ($\mathcal{I}_2$), like indoor temperature ($T_{in}(t)$) and indoor CO_2 concentration ($C_{in}(t)$).

Thus, this causal relationship (Fig. 6.1) is denoted as $\mathcal{X}_B, \mathcal{P}_B \xrightarrow{\mathcal{I}_B} \mathcal{F}_B$. According to Fig. 1.6, the chosen set of optimal actions $\mathcal{X}_B^\star$ under the same $\mathcal{P}_B$ leads to $\mathcal{X}_B^\star, \mathcal{P}_B \xrightarrow{\mathcal{I}_B^\star} \mathcal{F}_B^\star$. For conveying the impact of this change from usual to optimal plan, the difference between the usual values ($\tilde{x}$) and the optimal values ($x^\star$) is translated as follows:

$$
\begin{gathered}
\Pi(\Delta x, v_{-3}, v_{-2}, v_{-1}, v_1, v_2, v_3) : \\
\begin{array}{cc}
\Delta x < v_{-3} \rightarrow \text{big fall } (\downarrow\downarrow\downarrow), & \Delta x \geq v_3 \rightarrow \text{big rise } (\uparrow\uparrow\uparrow), \\
v_{-2} \leq \Delta x < v_{-3} \rightarrow \text{medium fall } (\downarrow\downarrow), & v_2 \leq \Delta x < v_3 \rightarrow \text{medium rise } (\uparrow\uparrow), \\
v_{-3} \leq \Delta x < v_{-1} \rightarrow \text{small fall } (\downarrow), & v_1 \leq \Delta x < v_2 \rightarrow \text{small rise}(\uparrow),
\end{array} \\
v_{-1} \leq \Delta x < v_1 \rightarrow \text{no significant change (no arrows)}, \\
\text{where } \Delta x = x^\star - \tilde{x},\ x^\star \in \{\mathcal{X}_B^\star, \mathcal{I}_B^\star, \mathcal{F}_B^\star\} \text{ and } \tilde{x} \in \{\tilde{\mathcal{X}}_B, \tilde{\mathcal{I}}_B, \tilde{\mathcal{F}}_B\}.
\end{gathered}
\tag{6.1}
$$

6.4 Experimental Testbed and its Description

The experimental testbed (Fig. 6.2) is an office room[1] at Grenoble Institute of Technology, France, shared among four researchers [133]. The office draws power from a fixed tariff power supply at the rate (E_{elec}) of 0.15 Euros per kilowatt-hour (kWh). Although Heating, Ventilation and Air Conditioning (HVAC) is absent, the office has a room heater. Its fuel consumption cost is at the rate (E_{fuel}) of 0.089 Euros per kWh. The metabolism of the occupants (φ_{bodies}) is assumed to be constant at 129 watts (W) per person.

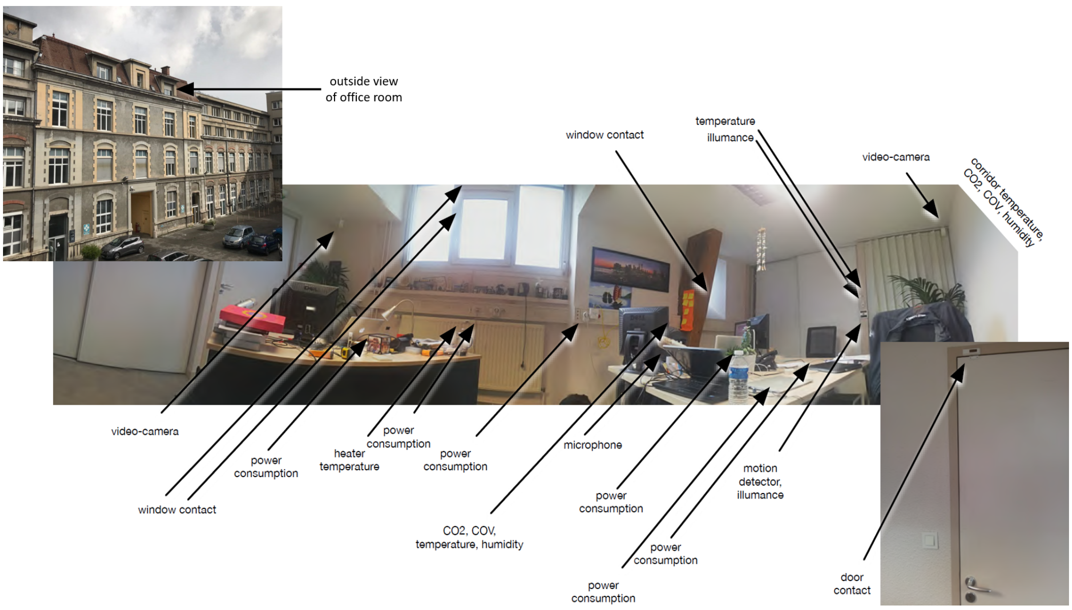
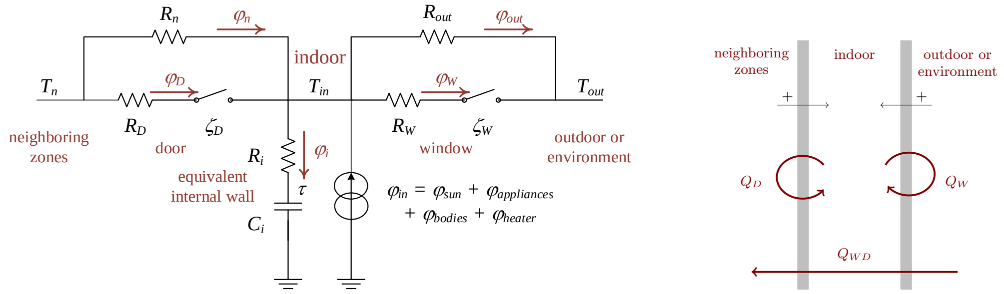


Figure 6.2: Panoramic view (from door) of the office and its outside view [133].

[1]This work is partially supported by the Indo-French project (DST-INRIA/2015-02/BIDEE/0978).



(a) Thermal model [159] (b) Aeraulic model [158]

Figure 6.3: Simulation models based on occupants' actions fitted to the office room.

This testbed is fitted with 27 sensors for acquiring data like door and window openings, acoustic pressure, illuminance, indoor and corridor physical variables like temperature, CO_2 concentration, humidity and volatile organic compounds (COV). Hourly data from sensor recordings and weather conditions are stored in the historical database ($\mathcal{H}_{DB}$) from April 2015 to October 2016 for creating the context during simulations. This data is also utilized to dynamically estimate the occupancy [3, 4].

6.5 Physical Knowledge Models

Developing the electro-mechanical equivalent of the office room is a challenging task as there is a multitude of ongoing building-occupant-environment interactions, like heat flow and air flow. The specifications of the physical models, fitted to this office room, is found from [158, 159]. Using the parameters mentioned in Table 6.1, the equivalent thermal model (Fig. 6.3a) is described as follows:

$$\begin{aligned}
T_{in} &= \frac{R_{eq}}{R_i}\tau + R_{eq}\varphi_{in} + R_{eq}\left(\frac{1}{R_{out}} + \frac{\zeta_W}{R_W}\right)T_{out} + R_{eq}\left(\frac{1}{R_n} + \frac{\zeta_D}{R_D}\right)T_n,\\
&\text{where } \frac{1}{R_{eq}} = \frac{1}{R_i} + \frac{1}{R_{out}} + \frac{\zeta_W}{R_W} + \frac{1}{R_n} + \frac{\zeta_D}{R_D},\ R_D = \frac{1}{\rho_{air}c_{p,air}Q_D},\\
R_W &= \frac{1}{\rho_{air}c_{p,air}Q_W},\ Q_W = Q_W^0 + \zeta_W Q_W^1 \text{ and } Q_D = Q_D^0 + \zeta_D Q_D^1\\
&\text{with time-invariant } R_n,\ R_{out} \text{ and } R_i.
\end{aligned} \tag{6.2}$$

Similarly, the equivalent model (Fig. 6.3b) representing the CO_2 based aeraulic characteristics of the office room [158] is described as follows:

$$C_{in}(t) = C_{out} + \frac{\left(Q_n^0 + \zeta_D(t)Q_D\right)C_n(t)}{Q_{out}^0 + Q_n^0 + \zeta_W(t)Q_W + \zeta_D(t)Q_D} + \frac{S_{CO_2}n(t)}{Q_{out}^0 + Q_n^0 + \zeta_W(t)Q_W + \zeta_D(t)Q_D},$$

$$\text{where } Q_{out}(t) = Q_{out}^0 + \zeta_W(t)Q_W \text{ and } Q_n(t) = Q_n^0 + \zeta_D(t)Q_D. \tag{6.3}$$

Finally, the heater-related energy consumption (P_{fuel}) is described as follows:

$$P_{fuel}\,(t) = \zeta_H\,(t) \times P_{heater}^{max}. \tag{6.4}$$

Thus, besides the physical context, ζ_D, ζ_W and ζ_H can influence the indoor ambience. The next section discusses an approach to search ζ_W, ζ_D and ζ_H for optimal effects.

6.6 Obtaining Optimal Actions of Occupants

Using the discretized simulation models (Section 6.5), the optimal ζ_W, ζ_D and ζ_H are recommended (Fig. 1.6). The continuous-time variables ($x(t)$) are transformed into discrete-time variables (x^k) using the average value over the k^{th} time quantum.

The concerned problem (Section 1.4) is characterized by the solution vector encoding in Eq. (1.19). As it considers an hourly granularity, the solution vector ($\mathbf{X}_B$) is 72-dimensional ($N = 72 = 24$ hours $\times 3$ actions). To simplify intepretations, the optimization problem considers only binary-valued (open(1)/close(0) or on(1)/off(0)) ζ_W^k, ζ_D^k and ζ_H^k.

Table 6.1: Description of building simulation model parameters [133]

Parameters	Meaning	Remarks
τ	Average temperature of the building envelope	Data from $\mathcal{H}_{DB}$
R_n, R_{out}, R_W, R_D	Thermal resistance of neighboring zones, outdoor, window and door	Data from $\mathcal{H}_{DB}$
R_i, C_i	Equivalent resistance and capacitance due to inertia	Data from $\mathcal{H}_{DB}$
R_{eq}	Equivalent resistance	By Eq. (6.2)
T_{in}, T_n, T_{out}	Temperatures inside, with adjacent corridor and outside	Data from $\mathcal{H}_{DB}$
φ_{in}	Total indoor energy gains	Data from $\mathcal{H}_{DB}$
ρ_{air}	Air density	Typical value is 1.204m^3
$c_{p,air}$	Specific heat of air at room temperature	Typical value is 1.004 kJ.kg^{-1}.K^{-1}
C_{in}, C_n, C_{out}	CO_2 concentration indoor, with neighboring zone and outdoor	Data from $\mathcal{H}_{DB}$, $C_{out} = 395 \times 10^{-6}$ mol per mol of air (constant)
Q_n, Q_{out}, Q_W, Q_D, Q_{WD}	Air flow with adjacent corridor, outdoor, through window, through door, through window and door (cross-ventilation)	Data from $\mathcal{H}_{DB}$
S_{CO_2}	Breath production in CO_2 from each occupant	Typical value is 8.73×10^{-6} mol.m^3.s^{-1} per person per mol of air
P_{elec} or $\varphi_{appliances}$	Power drawn from electric supply or net heat flow from appliances	Data from $\mathcal{H}_{DB}$
P_{heater}^{max}	Maximum energy consumption associated with water circulation for hourly heater usage	Typical value is 2000W

When there is at least one occupant ($n^k \neq 0$), the optimization problem (Section 1.4) minimizes the occupants' dissatisfaction ($\mathbf{F}_B$), defined by Eq. (1.20). These effects of occupants' actions ($M = 4$ objectives) are given in Table 1.2, which represents the thermal discomfort (σ^k_{temp}), CO_2 based air quality discomfort (σ^k_{air}), an indicator for energy cost (σ^k_{cost}) and the dissatisfaction from changes in actions at successive hours (δ^k_{WD}).

These objectives necessitate the use of MaOO algorithms to address this building energy management problem. The result of such MaOO algorithms is a set of trade-off solutions. However, a single solution from this set can be chosen for implementation. The next part discusses various approaches to select the final context-relevant solution.

6.6.1 Decision-making Strategies

Earlier optimization approaches [6, 57] combined multiple objectives into a single objective and thus, resulted in a single solution. However, after termination (i.e., after G_{max} generations), MaOO algorithms result in a set of solutions to represent the estimated Pareto-optimal Set (PS: $\mathcal{A}_{G_{max}}$) and the estimated Pareto-Front (PF: $\mathcal{A}_{\mathbf{F},G_{max}}$). A multi-objective formulation for the building energy management problem [145] mentions the requirement of expert's knowledge for selecting the most relevant solution. However, in absence of an expert's knowledge, for allowing convenient decision-making by the occupants, this chapter subsequently outlines four strategies.

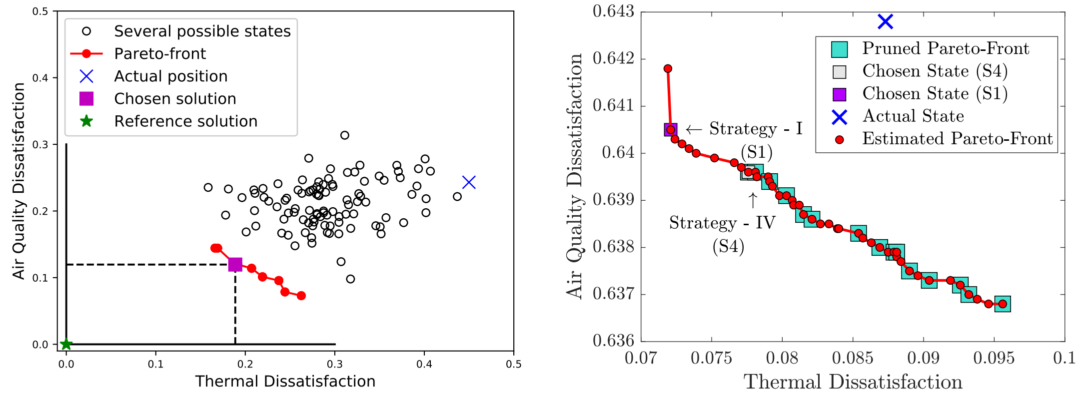
6.6.2 Strategy I: In absence of user preferences

Assuming the objectives of Eq. (1.20) have a preference weighting of $w_{b,1}$, $w_{b,2}$, $w_{b,3}$ and $w_{b,4}$, the distance to a solution ($D_B(.)$) and the best compromise ($\mathbf{X}^\star_B$) are as follows:

$$D_B\left(\mathbf{X}_B\right) = \sum_{i=1}^{4} \left(w_{b,i} \times \left| f_{B,i}\left(\mathbf{X}_B\right) - f^{ide}_{B,i} \right|\right), \text{ where } \sum_{i=1}^{4} w_{b,i} = 1, \tag{6.5}$$

$$\mathbf{X}^\star_B = \underset{\mathbf{X}_B \in \mathcal{A}_{G_{max}}}{\arg\min}\; D_B\left(\mathbf{X}_B\right) \text{ and } \mathbf{F}^\star_B = \left[f_{B,1}\left(\mathbf{X}^\star_B\right), \cdots, f_{B,4}\left(\mathbf{X}^\star_B\right)\right]. \tag{6.6}$$

As the objectives in Eq. (1.20) have nearly the same scale, $\mathbf{F}^\star_B$ from Eq. (6.6) represents the point from the estimated PF closest to the ideal objective vector ($\mathbf{F}^{ide}_B = [0, 0, 0, 0]$). Thus, $\mathbf{F}^\star_B$ corresponds to a schedule with the minimum net (global) occupant dissatisfaction. As constraining changes in actions will prohibit the other optimization objectives to evolve leading to poor exploration of the search space, the objective weights are considered as $w_{b,1} = 33.22\%$, $w_{b,2} = 33.22\%$, $w_{b,3} = 33.22\%$ and $w_{b,4} = 0.34\%$. For a 2-objective problem (minimizing σ_{temp} and σ_{air}), this strategy is shown in Fig. 6.4a.



(a) Strategy-I (05-May-2015) [133]

(b) Strategy-IV (08-April-2015) [139]

Figure 6.4: Choosing a single solution from the estimated PF.

6.6.3 Strategy II: Setting a reasonable preference

For allowing occupants (users) to interact with the building energy system, a slider prototype [1, 133] is developed. It has co-dependent horizontal bars (per objective), which are divided into infeasible (red), non-optimal (gray) and optimal (white) regions (e.g., Fig. 6.5 shows the sliders from a real interface [105]). When the user voluntarily navigates one of the sliders, the other sliders can simultaneously adjust the respective optimal or non-optimal regions. When all the sliders are set at desired positions representing occupants' preference $\mathbf{F}_B^{pref}$, the objective vector closest to it is selected from the PF using Eq. (6.6) and the corresponding schedule ($\mathbf{X}_B^{\star}$) is recommended.

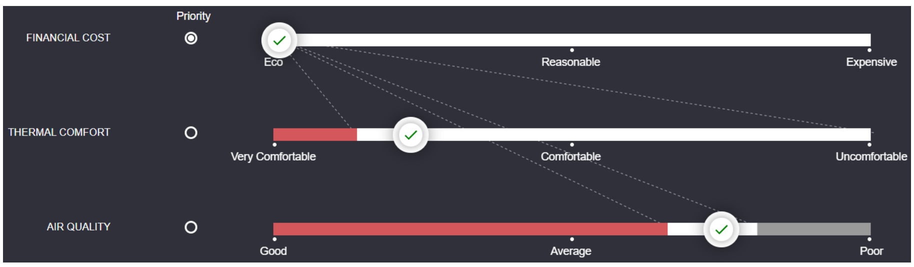


Figure 6.5: Screenshot of the slider (Strategy-II of decision-making) from the real user interface (`https://pareto-sliders.firebaseapp.com/`).

6.6.4 Strategy III: Multiple subjective preferences [135]

When the subjective preferences of multiple occupants are considered with equal priority, i.e., when there is no hierarchy among the occupants, this strategy provides a fair consensus solution [135]. The situation must satisfy the following conditions:

1. The occupants are rational, intelligent, and cooperative individuals.
2. The occupants can conclude on a mutually beneficial state (solution) of comfort.
3. There is a conflict of interest about which state (solution) should be preferred.

This situation is called the bargaining problem [129] which is commonly addressed using the α-fairness criteria for yielding the generalized Nash bargaining solution [175]. Inspired from this approach, the fair consensus criterion [135] is developed to yield the Fair Consensus Schedule (FCS) where the following assumptions are made:

- The i^{th} occupant in the office sets the preference $\mathbf{F}_{B,i}^{pref}$ using the sliders (Fig. 6.5).
- In the objective space, a point $\mathbf{F}_B$ belongs to a set of alternatives ($\mathcal{A}_{\mathbf{F}}$).
- The utility function $U\left(\mathbf{F}_B, \mathbf{F}_{B,i}^{pref}\right)$ indicates the degree of unfairness (disagreement) of the comfort state ($\mathbf{F}_B$) as compared to the preference of the i^{th} occupant ($\mathbf{F}_{B,i}^{pref}$).
- A parameter α_B regulates the kind of fairness sought among the occupants. Thus, considering $U(.)$, α_B and a small number ϵ_B, the estimated unfairness $C(\mathbf{F}_B, \mathbf{F}_{B,i}^{pref})$ of a comfort state $\mathbf{F}_B$ to the preferred state $\mathbf{F}_{B,i}^{pref}$ is given as follows:

$$C\left(\mathbf{F}_B, \mathbf{F}_{B,i}^{pref}\right) = \frac{\left(U\left(\mathbf{F}_B, \mathbf{F}_{B,i}^{pref}\right) + \epsilon_B\right)^{(\alpha_B - 1)}}{(\alpha_B - 1)}, \text{ where } i = 1, \cdots, n. \tag{6.7}$$

- The unfairness of a solution (having a comfort state $\mathbf{F}_B$) to the preferences of all the n occupants is denoted by $D_F(\mathbf{F}_B)$. Thus, the comfort state $\mathbf{F}_B^{\star}$ having the minimum value of $D_F(.)$ is the least unfair to all the occupants and the corresponding schedule $\mathbf{X}_B^{\star}$, mapping to $\mathbf{F}_B^{\star}$, is FCS [135]. The state $\mathbf{F}_B^{\star}$ is obtained as follows:

$$\mathbf{F}_B^{\star} = \underset{\mathbf{F}_B \in \mathcal{A}_{\mathbf{F}}}{\arg\min}\, D_F(\mathbf{F}_B), \text{ where } D_F(\mathbf{F}_B) = \sum_{i=1}^{n} C\left(\mathbf{F}_B, \mathbf{F}_{B,i}^{pref}\right). \tag{6.8}$$

Thus, the evaluation of FCS is governed by $U(.)$ and α_B. As $U(.)$ signifies how much a comfort state $\mathbf{F}_B$ is different from a preference $\mathbf{F}_B^{pref}$, it is evaluated as follows:

$$U\left(\mathbf{F}_B, \mathbf{F}_B^{pref}\right) = \sum_{k=1}^{M} \left(f_{B,k} - f_{B,k}^{pref}\right)^2, \text{ where } M = \text{number of objectives}. \tag{6.9}$$

The variations of $U(.)$ (Eq. (6.9)), $C(.)$ (Eq. (6.7)) and $D_F(.)$ (Eq. (6.8)) over the solutions in the Pareto-Front are shown in Fig. 6.6.

6.6. OBTAINING OPTIMAL ACTIONS OF OCCUPANTS

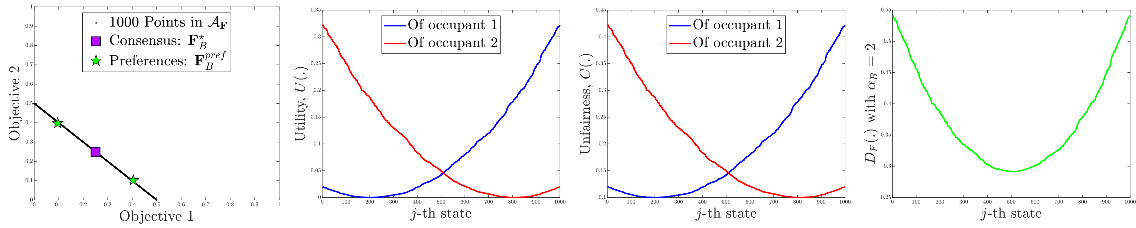


Figure 6.6: Variations in different functions for Fair Consensus Schedule [135].

The parameter α_B for FCS evaluation (used in Eq. (6.7)) is defined for values in the range $(1, \infty)$. Different values of α_B signify different perspectives of fairness among all the occupants as described in Table 6.2. Thus, $\mathbf{X}^{\star}_{B} = \arg \mathbf{F}^{\star}_{B}$ (where $\mathbf{F}^{\star}_{B}$ is obtained using Eq. (6.8)) is the chosen schedule based on the fair consensus criterion. For the concerned problem, $\alpha_B = 3$ is used as a trade-off between minimizing the group disagreement and the maximum disagreement, as specified in [135].

Table 6.2: Implications of fair consensus criterion for different values of α_B.

Parameters	**Implications**	**Contextual Significance**
$\alpha_B \to 1$	Proportional Fairness Criteria	Consensus highly prioritizes the solutions with lower unfairness and thus, FCS settles in favour of majority.
$\alpha_B = 2$	Group Disagreement	As $D_F(.)$ (Eq. (6.8)) becomes sum of $U(.)$, FCS minimizes average disagreement over the group of occupants.
$\alpha_B \to \infty$	Min-Max Criteria	It magnifies large disagreements and diminishes small disagreements. Thus, consensus is obtained by minimizing the maximum disagreement such that FCS is not oblivious to an occupant with a very different preference.

6.6.5 Strategy IV: Preference in decision space [139]

The occupants have a usual/preferred schedule ($\tilde{\mathbf{X}}_B$ from $\mathcal{H}_{DB}$) and are more likely to embrace a recommended schedule ($\mathbf{X}^{\star}_{B}$) which has a smaller deviation from $\tilde{\mathbf{X}}_B$. By exploring the multi-modality (Section 5.1) of the concerned MaOO problem, different alternative schedules for the same objective values can be discovered, which further assists in finding a schedule $\mathbf{X}^{\star}_{B}$ with the least deviation from $\tilde{\mathbf{X}}_B$.

Once a Multi-Modal Multi-Objective Evolutionary Algorithm (MMMOEA) estimates the PF and the equivalent subsets within the PS, this decision-making strategy filters out the subset of the most relevant schedules ($\mathcal{A}^{sch}_{\mathbf{F},G_{max}}$). Thereafter, Strategy-I, II or III of decision-making is used in the objective space to finally obtain the schedule $\mathbf{X}^{\star}_{B}$ for recommendation, as illustrated in Fig. 6.4b.

If the least deviation (Δ^{sch}_{min}) is the minimum change in schedules ($\mathbf{X}_B$) of $\mathcal{A}_{G_{max}}$ from the usual schedule ($\tilde{\mathbf{X}}_B$), the pruned Pareto-Front $\mathcal{A}^{sch}_{\mathbf{F},G_{max}}$ is estimated as follows:

$$\begin{aligned} \mathcal{A}^{sch}_{\mathbf{F},G_{max}} &= \{\mathbf{F}_B(\mathbf{X}_B) \mid \Delta\mathbf{X}_B = \Delta^{sch}_{min}\}, \\ \text{where } \Delta^{sch}_{min} &= \underset{\mathbf{X}_B \in \mathcal{A}_{G_{max}}}{\arg\min}\ \Delta\mathbf{X}_B = \underset{\mathbf{X}_B \in \mathcal{A}_{G_{max}}}{\arg\min} \sum_{j=1}^{N} |\tilde{x}_{B,j} - x_{B,j}| \,. \end{aligned} \tag{6.10}$$

6.6.6 Optimization Results and Discussions

The purpose of the optimization module is to yield the set of best trade-offs and to allow the users to browse through this set for obtaining a desirable schedule. For addressing this MaOO problem, the concerned framework (Fig. 1.6) is implemented for the office room of the Grenoble Institute of Technology on a computer with 8 GB RAM and Intel Core i7 processor (having 2.20 GHz clock speed) using Python 3.4 and its performance is analyzed through following experiments.

Recommending an Optimization Algorithm

The algorithms, investigated for solving the MaOO problem, are specified as follows:

1. *Using SA*: Similar to existing works [6, 57], a weighted combination of multiple objectives ($\sum_{i=1}^{4} w_{b,i} \times f_{B,i}$) is considered to generate an equivalent single-objective problem and then solved using Simulated Annealing (SA) [102]. For defining the neighborhood in SA, the changes in variables are restricted to 10%, and the radius is attenuated by 1 in each of the 1000 iterations. The temperature is considered to be linearly decreasing over the iterations. The best solution over 100 runs is noted.

2. *Using NSGA-II*: Being a popular choice, NSGA-II [47, 145] is used along with the reproduction scheme in Fig. 6.7. It uses binary tournament for selecting the parent solutions, a mutation probability of $3/N$ and n_{pop} of 36.

3. *Using AGE-II*: For approximate estimation of PS and PF, Approximation-Guided Evolutionary algorithm (AGE-II) [133, 179] is used along with the reproduction scheme in Fig. 6.7 and n_{pop} of 36. It incorporates the formal notion of additive approximation with a degree of approximation of 0.01.

4. *Using NAEMO*: To estimate a well-diverse PF, a decomposition-based algorithm (NAEMO [160]) is also considered. It is implemented using Algorithm 4.1 with

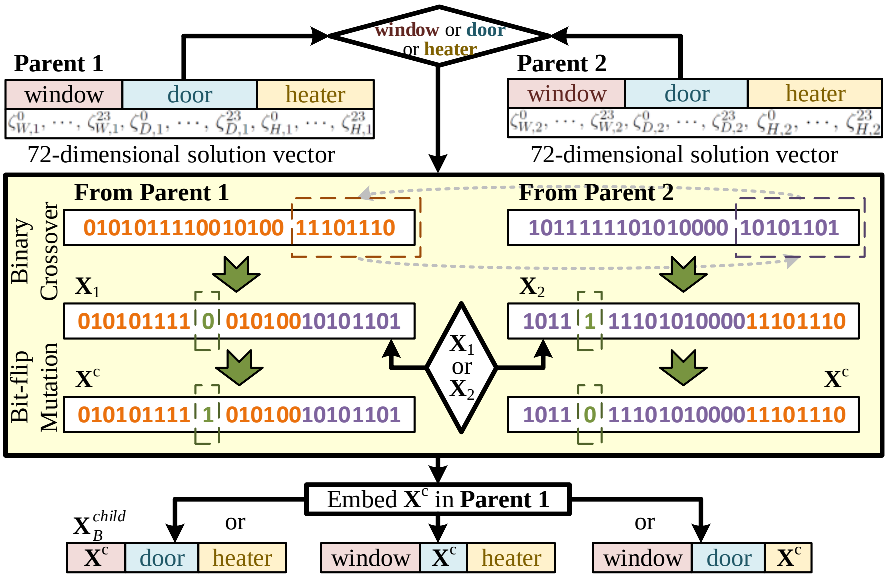


Figure 6.7: With all decisions equally-likely, reproduction of a solution vector $\mathbf{X}_B^{child}$.

$n_{dir} = 35$ and the reproduction scheme in Fig. 6.7 instead of Algorithm 4.2.

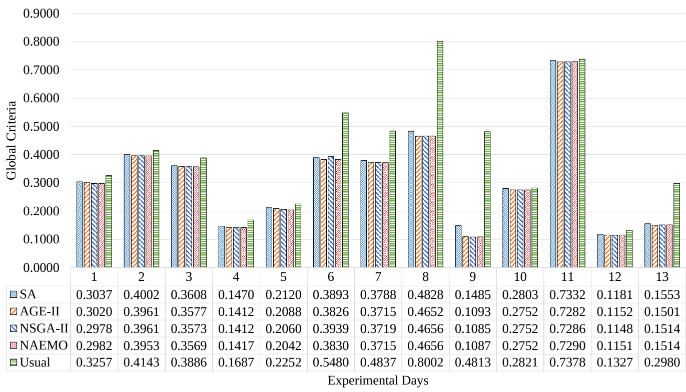
For NSGA-II, AGE-II, and NAEMO, a maximum of 300 iterations is used (i.e., $G_{max} = 300$). For each algorithm (except SA), the result over 5 runs is considered using Eq. (1.13) with $K_{PF} = 5$. Combining the small solution sets obtained over multiple runs to generate the final Pareto-Front, help in balancing the trade-off between small population size and better performance of a MOEA [170]. For 20 randomly sampled days over the experimental duration, the results are noted from the optimization algorithms using Strategy-I of decision-making. For detailed results, kindly refer to [133, 139].

For denoting the global minima attained by an algorithm, $D_B\left(\mathbf{X}_B^\star\right)$ is noted in Fig. 6.8. The corresponding value for the usual schedule of the occupants, i.e., $D_B\left(\tilde{\mathbf{X}}_B\right)$, is also noted for comparison. A higher deviation between the global and the usual dissatisfaction (i.e., $\Delta D_B = D_B\left(\tilde{\mathbf{X}}_B\right) - D_B\left(\mathbf{X}_B^\star\right)$) denotes better energy management. Such deviations are noted in Table 6.3 along with p-values from the t-test corresponding to 95% confidence interval under the null hypothesis that the mean ΔD_B is zero (insignificant).

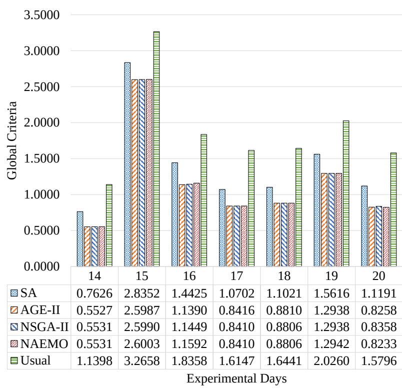
By analyzing the results, the following insights are obtained:

1. From Fig. 6.8, $D_B\left(\mathbf{X}_B^\star\right) < D_B\left(\tilde{\mathbf{X}}_B\right)$ for all cases. This is supported by the positive values of ΔD_B and p-values ≥ 0.05 (rejecting the null hypothesis) in Table 6.3.

2. In all cases, SA yields the worst $D_B\left(\mathbf{X}_B^\star\right)$ (Fig. 6.8) and the least ΔD_B (Table 6.3) as transforming multiple objectives into a single objective neglects their conflict. Thus, such transformations are not recommended during the optimization.



(a)



(b)

Figure 6.8: Comparing the performance of optimizers when global criteria obtained with respect to usual schedule of occupant's actions ($D_B(\tilde{\mathbf{X}}_B)$) is (a) ≤ 1, (b) > 1 [133].

Table 6.3: Difference in the global dissatisfaction of the optimal schedule from that of the historical schedule ($\Delta D_B = D_B(\tilde{\mathbf{X}}_B) - D_B(\mathbf{X}^\star_B)$) [133, 139].

Day Number	Date	Deviation ΔD_B				Execution time (in seconds)			
		SA	**AGE-II**	**NSGA-II**	**NAEMO**	**SA**	**AGE-II**	**NSGA-II**	**NAEMO**
1	01-Apr-2015	0.0220	0.0237	0.0279	0.0275	61.8668	91.6743	101.3238	65.2527
2	20-May-2015	0.0142	0.0183	0.0183	0.0191	59.3280	97.6341	103.1781	68.9027
3	30-Sep-2015	0.0278	0.0309	0.0314	0.0318	60.9106	84.8661	100.1119	73.1182
4	08-Oct-2015	0.0217	0.0275	0.0275	0.0270	62.3862	89.8696	101.8488	70.2369
5	03-Nov-2015	0.0132	0.0164	0.0192	0.0210	64.7919	90.9098	102.4084	70.1459
6	07-Dec-2015	0.1587	0.1654	0.1541	0.1650	61.0254	86.0383	100.6460	66.0193
7	26-Jan-2016	0.1049	0.1122	0.1118	0.1122	61.6363	83.7820	101.3354	67.7403
8	01-Feb-2016	0.3173	0.3349	0.3345	0.3345	62.4322	85.8027	101.1520	70.8845
9	17-Mar-2016	0.3329	0.3720	0.3728	0.3726	60.5009	84.0765	100.3140	68.2200
10	13-Apr-2016	0.0018	0.0069	0.0069	0.0069	62.8810	85.6809	102.4386	65.1108
11	23-May-2016	0.0046	0.0096	0.0092	0.0088	41.6105	67.8174	71.4416	65.1858
12	02-Jun-2016	0.0146	0.0175	0.0180	0.0176	40.9297	67.6796	74.3269	70.1280
13	20-Oct-2016	0.1427	0.1479	0.1467	0.1467	62.0120	101.5243	100.8283	66.5560
14	16-Jun-2015	0.3773	0.5871	0.5867	0.5867	62.4832	89.1091	103.2391	66.1912
15	07-Jul-2015	0.4306	0.6671	0.6668	0.6655	62.8645	88.0524	100.6763	66.0063
16	01-Sep-2015	0.3933	0.6968	0.6909	0.6765	63.3529	94.0166	105.1845	73.7089
17	30-Jun-2016	0.5444	0.7731	0.7737	0.7737	42.0875	68.1006	73.8499	64.0145
18	26-Jul-2016	0.5420	0.7631	0.7635	0.7635	42.3909	70.1146	72.1790	64.0524
19	31-Aug-2016	0.4644	0.7321	0.7321	0.7317	42.2508	65.3673	72.9211	66.4241
20	08-Sep-2016	0.4605	0.7539	0.7438	0.7563	42.9969	63.3149	74.5413	68.6838
Mean		**0.2194**	**0.3128**	**0.3118**	**0.3122**	**56.0369**	**82.7716**	**93.1973**	**67.8291**
***p*-value**		**0.000136**	**0.000306**	**0.000305**	**0.000148**	–	–	–	–

3. The remaining MaOO algorithms (NSGA-II, AGE-II and NAEMO) have similar bar heights in Fig. 6.8 and similar ΔD_B in Table 6.3. This indicates all the MaOO algorithms are equally capable of finding the approximation of an optimal schedule.

4. Besides $D_B(\mathbf{X}^\star_B)$ and ΔD_B, the execution time (in seconds) of the algorithms are noted in Table 6.3, using which the algorithms are ranked as follows: SA, NAEMO, AGE-II and NSGA-II. The lower speed of NSGA-II [47] is due to the computationally expensive non-dominated sorting step whereas the higher speed of SA [102] is due to its simpler solution comparisons for being a single-objective optimization algorithm. The speed of AGE-II [179] and NAEMO [160] are intermediate.

5. During summer (mid-June to mid-September), the outside weather is less favorable to attain occupants' comfort. It is observed from higher values of $D_B(\mathbf{X}_B^{\star})$ and $D_B\left(\tilde{\mathbf{X}}_B\right)$ (Fig. 6.8b) for these days (day number 14 to 20) as compared to other days. Thus, future case studies in summer may benefit from using cooling devices.

Since NAEMO is fastest (Table 6.3) among the MaOO algorithms, it is recommended for the concerned building energy management problem.

Importance of Each Objective

On a random day (03-November-2015), the hourly variations in T_{in}, C_{in}, ζ_W, ζ_D and P_{fuel} are plotted in Fig. 6.9, for both the optimal schedule (green dotted curves) and the actual schedule (blue dashed curves). The final schedule is recommended using Strategy-I of decision-making. The following observations are noted from these plots:

1. The decision-making and the results [2], with data from 8 am to 8 pm, are illustrated in Fig. 6.9 (first column) for two objectives (σ_{temp} and σ_{air}). The plots for both T_{in} and C_{in} from the optimal schedule have a lower trend than those from the actual schedule. Being an autumn day, the doors and windows are usually closed whereas the optimal schedule recommends that occasionally opening them can be beneficial.

2. The above problem is extended to a 3-objective problem by introducing σ_{cost} (i.e., the heater operation with P_{fuel}) and the associated plots are shown in Fig. 6.9 (second column). Although the overall trends of T_{in} and C_{in} are similar to the 2-objective problem, T_{in} is closer to 21°C (294.15K) by using the heater. Thus, introducing the heater in the 3-objective problem allows a more regulated control over T_{in}.

3. The above problem is extended to a 4-objective problem by introducing δ_{WD} and the associated plots are shown in Fig. 6.9 (third column). The overall trends of T_{in} and C_{in} continue to be similar to the 3-objective problem with fewer changes in ζ_W and ζ_D. Hence, building energy management can occur without too much interference with the daily work of the occupants.

Thus, all four objectives are essential for the concerned MaOO problem.

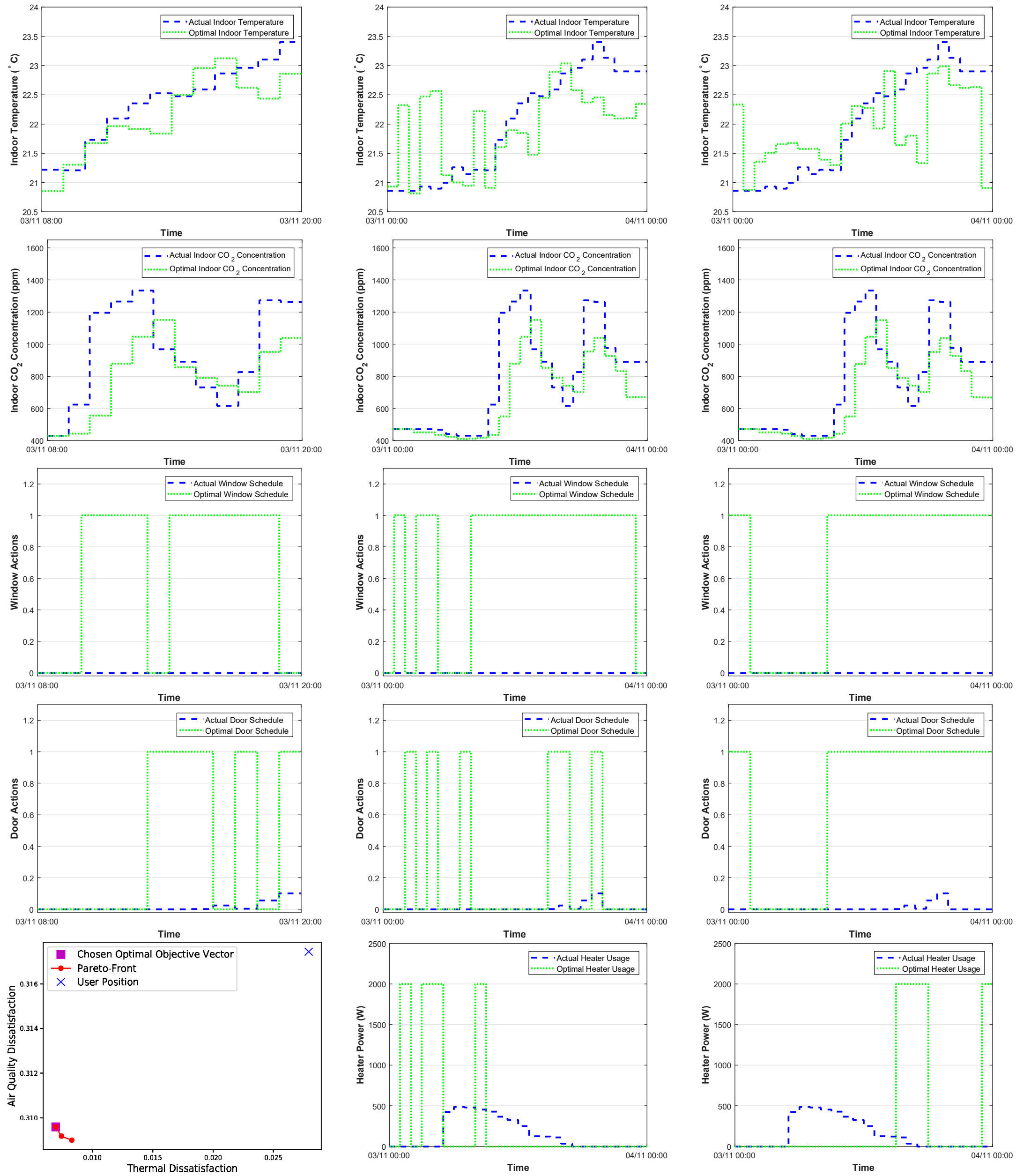


Figure 6.9: Indoor physical parameters along optimal actions for 2-objective problem with σ_{temp} and σ_{air} (first column); for 3-objective problem with σ_{temp}, σ_{air} and σ_{cost} (second column); and for 4-objective problem with σ_{temp}, σ_{air}, σ_{cost} and δ_{WD} (third column) [133].

Analyzing the Fair Consensus Criterion [135]

For this experiment, the 2-objective problem is revisited with door and window actions (ζ_D and ζ_W) between 8 am to 8 pm while minimizing thermal and aeraulic discomfort (σ_{temp} and σ_{air}). In this experiment, the FCS (Eq. (6.8)) is obtained post-optimization (Approach-I) over the estimated PF ($\mathcal{A}_{\mathbf{F},G_{max}}$). However, the search for FCS can also be integrated during optimization (Approach-II) by considering minimization of $D_F(.)$ (Eq. (6.8)) as the $(M+1)^{\text{th}}$ optimization objective. Thus, this extended optimization problem considers the objective vector ($\mathbf{F}'_B$) as follows:

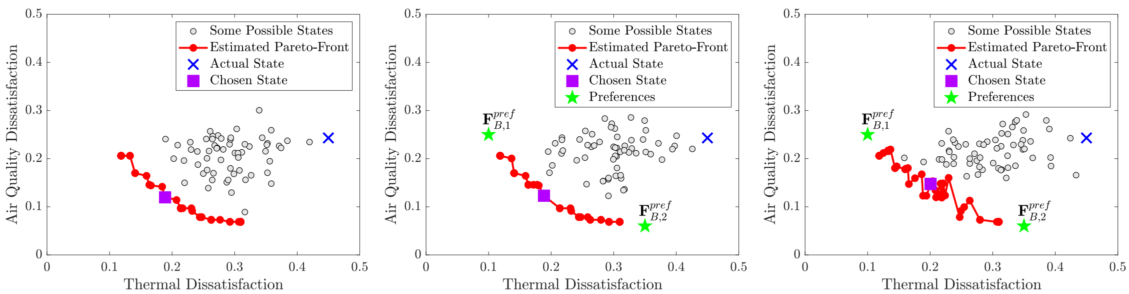
$$\mathbf{F}'_B = [f_{B,1}, \cdots, f_{B,M}, D_F(\mathbf{F}_B)], \text{ where } D_F(.) \text{ is given by Eq. (6.8).} \tag{6.11}$$

After estimating the PF by this approach, $\mathbf{X}^\star_B$ is obtained similar to Approach-I.

For this experiment, two occupants specify their preferences in terms of average T_{in} and C_{in} over the entire day and the physical context of 05-May-2015 is considered for investigating the performance in Table 6.4. Additionally, Strategy-I of decision-making is also considered for comparison of results.

Table 6.4: Comparison of the consensus searching approaches for the building energy management problem [135]

Items		Average of C_{in} (ppm)	Average of T_{in} (° C)	σ_{air}	σ_{temp}
Preference of occupant 1 ($\mathbf{F}^{pref}_{B,1}$)		675.0000	23.3000	0.2500	0.1000
Preference of occupant 2 ($\mathbf{F}^{pref}_{B,2}$)		466.0000	24.0500	0.0600	0.3500
Actual State ($\tilde{\mathbf{F}}_B$) Using Historical Schedule		667.4100	24.3497	0.2431	0.4499
Chosen State ($\mathbf{F}^\star_B$)	**Strategy-I**	531.5600	23.5664	0.1196	0.1888
	Approach-I (Strategy-III)	535.3000	23.5658	0.1230	0.1886
	Approach-II (Strategy-III)	562.0300	23.6021	0.1473	0.2007



(a) Strategy-I (b) Approach-I (Strategy-III) (c) Approach-II (Strategy-III)

Figure 6.10: Comparison of different approaches for obtaining consensus from two occupants with subjective preferences of comfort criteria [135].

Although the chosen states ($\mathbf{F}^\star_B$) from the Strategy-I and Approach-I (Strategy-III) appear to overlap in Figs. 6.10a and 6.10b, the small differences in σ_{temp} and σ_{air} lead to dissimilar average values of T_{in} and C_{in} (Table 6.4). Unlike these results, the chosen state ($\mathbf{F}^\star_B$) from Approach-II (Strategy-III with Eq. (6.11)) is less biased to either of the preferences. Instead of highly satisfying one of the decision-makers [129], a bargaining is more likely when intermediate choices exist between the multiple preferences [135]. Thus, the fair consensus criterion is a more practical decision-making approach where the fairness can be regulated through the utility function $U(.)$ and the parameter α_B.

Exploring the Multi-Modality of the Problem [139]

In the previous experiments with standard MOEAs, multiple schedules ($\mathbf{X}_B^\star$) were observed to have the same $\mathbf{F}_B^\star$, out of which any random schedule was reported. To further explore this multi-modality, the 4-objective MaOO problem (Section 1.4) is optimized using LORD [140]. To implement LORD (Algorithm 5.1) for the concerned building energy management problem, the following enhancements are considered:

- LORD uses the reproduction scheme outlined in Fig. 6.7, instead of Algorithm 5.2, as single-point binary crossover and bit-flip mutation address the binary nature of the action variables whereas the decision tree deals with the multi-view nature (action variables from multiple domains: window, door and heater) of the problem.

- The following changes (Fig. 6.11) are considered to customize the filtering step of LORD (Algorithm 5.4):

  1. Due to the binary nature of the action variables, the cosine distance measures the node similarity (domain-wise). The binary symmetric matrices ($\mathcal{G}^{window}$, $\mathcal{G}^{door}$ and $\mathcal{G}^{heater}$) are generated by placing edges between those pairs of nodes where the distances are less than $\varepsilon_L = 0.4$ (more similar).

     Subsequently, the symmetric normalized graph Laplacians ($\mathcal{L}_{sym}^{window}$, $\mathcal{L}_{sym}^{door}$ and $\mathcal{L}_{sym}^{heater}$) are obtained using Eq. (5.1) and the eigendecomposition is performed separately on the Laplacians. In the respective domains, the smallest non-zero eigenvalue or the Fiedler value [59] (λ_2^{window}, λ_2^{door} and λ_2^{heater}) represents the quality of the graph partitioning [59]. Thus, these Fiedler values can determine the influence of the multiple domains on the overall Laplacian ($\mathcal{L}_{sym}^{comb}$) as follows:

     $$\mathcal{L}_{sym}^{comb} = \frac{\lambda_2^{window} \times \mathcal{L}_{sym}^{window} + \lambda_2^{door} \times \mathcal{L}_{sym}^{door} + \lambda_2^{heater} \times \mathcal{L}_{sym}^{heater}}{\lambda_2^{window} + \lambda_2^{door} + \lambda_2^{heater}}. \tag{6.12}$$

     Similar to Section 5.4.2, the algebraic multiplicity of 0 eigenvalue of $\mathcal{L}_{sym}^{comb}$ gives the number of connected components (k_{CC}) of the overall cluster structure and the eigenvectors of $\mathcal{L}_{sym}^{comb}$ from the second smallest to the k_{CC}^{th} eigenvalue are clustered ($\mathcal{C}_1, \cdots, \mathcal{C}_{k_{CC}}$) using the k-means algorithm. This clustering of the action schedules is performed in line 5 of Algorithm 5.4 for this study.

  2. Due to the assignment of high crowding distances to extreme values along each

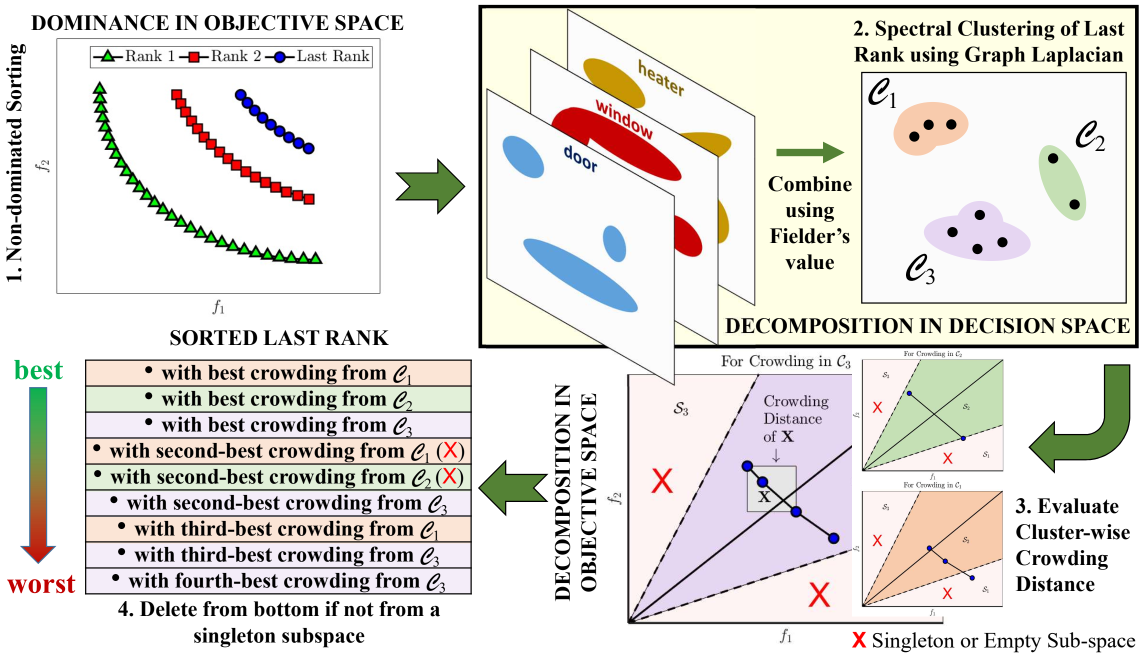


Figure 6.11: Modified filtering step for LORD [139].

dimension (Section 2.3.3), all binary action schedules will have equal crowding in the decision space (CDX). Hence, only crowding distance in the objective space (CDF) is considered in lines 6 and 7 of Algorithm 5.4 instead of Special Crowding Distance (SCD: a combination of CDX and CDF).

As NAEMO is recommended based on its performance in Table 6.3, the performance of the customized version of LORD is compared with that of NAEMO using $n_{dir} = 35$ and $n_{pop} = 105$. This experiment notes the amount of change from the usual schedule ($\Delta\mathbf{X}^\star_B/N$) required for a deviation of ΔD_B in the global criteria. As multiple distinct schedules are noted for the same $\mathbf{F}^\star_B$ (Eq. (6.6)), among these schedules the schedule $\mathbf{X}^\star_B$ having the least deviation from $\tilde{\mathbf{X}}_B$ can be recommended (Approach S1). Otherwise, the subset $\mathcal{A}^{sch}_{\mathbf{F},G_{max}}$ with the least deviation over the estimated PS (Eq. (6.10)) is obtained, from which $\mathbf{X}^\star_B$ is estimated using Eq. (6.6) (Approach S4).

The results obtained from the above two approaches for the same 20 days (as in Table 6.3) are noted in Table 6.5. The following insights are obtained from this experiment:

- For both LORD and NAEMO, as S1 prioritizes $D_B(\mathbf{X}^\star_B)$ over $\Delta\mathbf{X}^\star_B$, better ΔD_B values are obtained by S1 approaches. Similarly, as S4 prioritizes in the reverse order, better $\Delta\mathbf{X}^\star_B/N$ values are obtained by S4 approaches.
- Although S4 yields a poorer ΔD_B value, it is numerically very close to S1 at a much

Table 6.5: Amount of change in schedule ($\Delta\mathbf{X}_B^\star/N = \left(\sum_{j=1}^{N}\left|\tilde{\mathbf{x}}_{B,j} - \mathbf{x}_{B,j}^\star\right|\right)/N$) required for a deviation of ΔD_B $(= D_B(\tilde{\mathbf{X}}_B) - D_B(\mathbf{X}_B^\star))$ in global criteria [139].

Day Number	Date	Change in Schedule ($\Delta\mathbf{X}_B^\star/N$)				Change in Global Criteria (ΔD_B)			
		NAEMO		LORD		NAEMO		LORD	
		(S1)	(S4)	(S1)	(S4)	(S1)	(S4)	(S1)	(S4)
1	01-Apr-2015	0.4583	0.4583	0.3889	0.3889	0.0275	0.0275	0.0275	0.0275
2	20-May-2015	0.6806	0.6806	0.4167	0.4167	0.0191	0.0191	0.0183	0.0183
3	30-Sep-2015	0.5139	0.4167	0.3472	0.3333	0.0318	0.0303	0.0320	0.0307
4	08-Oct-2015	0.4167	0.4167	0.3889	0.3889	0.0270	0.0270	0.0270	0.0270
5	03-Nov-2015	0.7778	0.6667	0.7639	0.4861	0.0210	0.0170	0.0211	0.0179
6	07-Dec-2015	0.7083	0.5833	0.6111	0.5694	0.1650	0.1625	0.1650	0.1545
7	26-Jan-2016	0.7917	0.6111	0.6528	0.6250	0.1122	0.1112	0.1126	0.1116
8	01-Feb-2016	0.7778	0.7500	0.6806	0.5417	0.3345	0.3340	0.3349	0.3348
9	17-Mar-2016	0.7222	0.6250	0.7083	0.6528	0.3726	0.3706	0.3732	0.3728
10	13-Apr-2016	0.6944	0.5833	0.7778	0.6250	0.0069	0.0061	0.0073	0.0069
11	23-May-2016	0.6250	0.6250	0.5694	0.5694	0.0088	0.0088	0.0092	0.0092
12	02-Jun-2016	0.4444	0.4444	0.3472	0.3472	0.0176	0.0176	0.0176	0.0176
13	20-Oct-2016	0.7917	0.7917	0.4444	0.4444	0.1467	0.1467	0.1463	0.1463
14	16-Jun-2015	0.4028	0.4028	0.3611	0.3611	0.5867	0.5867	0.5872	0.5872
15	07-Jul-2015	0.5000	0.3889	0.2917	0.2917	0.6655	0.6610	0.6668	0.6668
16	01-Sep-2015	0.4306	0.2083	0.4028	0.1667	0.6765	0.5469	0.6972	0.5306
17	30-Jun-2016	0.3889	0.3611	0.2917	0.2917	0.7737	0.7347	0.7737	0.7737
18	26-Jul-2016	0.3611	0.3611	0.3333	0.3333	0.7635	0.7635	0.7635	0.7635
19	31-Aug-2016	0.3611	0.3611	0.2222	0.2222	0.7317	0.7317	0.7319	0.7319
20	08-Sep-2016	0.3472	0.2639	0.3472	0.1806	0.7563	0.7356	0.7563	0.7230
Mean		**0.5597**	**0.5000**	**0.4674**	**0.4118**	**0.3122**	**0.3019**	**0.3134**	**0.3026**

better $\Delta\mathbf{X}_B^\star/N$ value. For example, on 26-Jan-2016, NAEMO attains similar ΔD_B value with only a 61% change in the schedule using S4 as opposed to 79% using S1. Thus, S4 is a better approach than S1 for choosing the schedule to be recommended.

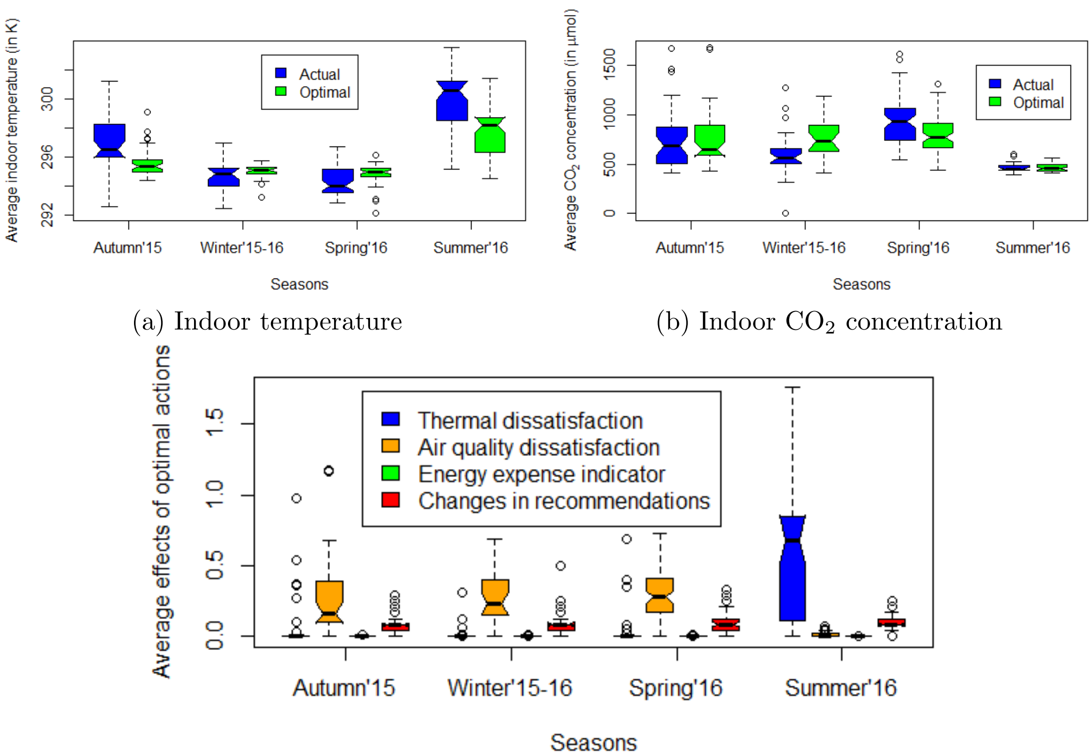
- LORD (S4) is noted to perform as good as or better than NAEMO (S4) in 15 out of 20 cases. Moreover, for five days (numbered 2, 5, 8, 13, and 19), $\Delta\mathbf{X}_B^\star/N$ from LORD (S4) is at least 10% better than those from NAEMO (S4). This superiority of LORD is due to its efficacy for multi-modal optimization problems (Chapter 5).

Thus, recommending a relevant and Pareto-optimal schedule can be beneficial for energy management. For example, on 01-Sep-2015, with only a 16% change from $\tilde{\mathbf{X}}_B$, LORD (S4) has obtained a better (Pareto-optimal) schedule $\mathbf{X}_B^\star$ than the usual schedule $\tilde{\mathbf{X}}_B$.

Analyzing the Seasonal Variations with Year-Round Results

The distribution of daily averages of the actual and the optimal T_{in} and C_{in} along with that of the four objectives are presented over the four seasons: Autumn'15, Winter'15-16, Spring'16 and Summer'16 in Fig. 6.12, from which the following observations are noted:

- In winter and spring, mean $T_{in}^\star$ is higher than $\tilde{T}_{in}$, and vice-versa during autumn and



(a) Indoor temperature

(b) Indoor CO_2 concentration

(c) Effects of optimal actions

Figure 6.12: Seasonal variations in daily average values of physical variables affected by occupants' actions [133].

summer (Fig. 6.12a) as σ_{temp} (Table 1.2) brings T_{in} between 294.15K and 296.15K.

- Mean $C^{\star}_{in}$ is lower than $\tilde{C}_{in}$, except in winter (Fig. 6.12b) when closed windows ($\zeta_W = 0$) reduces heat exchange (φ_{out}) and adversely affects airflow (Q_{out}) and C_{in}.
- The seasonal variations of the effects (optimization objectives) of the optimal actions (Fig. 6.12c) maintain near minimal values, except thermal dissatisfaction (σ_{temp}) in summer [132]. This raised σ_{temp} agrees with similar findings from Fig. 6.8b.

The recommended schedule of actions is beneficial for building energy management but difficult for users to learn, especially when the underlying phenomena are formally presented as in Fig. 6.9. Thus, generating simpler explanations is considered next.

6.7 Generating explanations

Changing the entire schedule may not be acceptable for an occupant [1]. Hence, the effects of each action should be explained so that the occupants can learn the priority of the optimal actions. Thus, the explanation generation is considered in this section.

6.7.1 Differential Explanations

A scenario includes these four groups of variables (Fig. 6.1). The relations between these groups is pre-determined as expert's abstract knowledge [172]. Differential explanations are constructed by analyzing the difference between the two scenarios: the usual scenario recorded in $\mathcal{H}_{DB}$ and the optimal scenario obtained from the estimated PF as follows:

$$\begin{aligned}&\left(\text{Optimal scenario: } \mathcal{X}_B^\star, \mathcal{P}_B \overset{\mathcal{I}_B^\star}{\to} \mathcal{F}_B^\star\right) - \left(\text{Usual scenario: } \tilde{\mathcal{X}}_B, \mathcal{P}_B \overset{\tilde{\mathcal{I}}_B}{\to} \tilde{\mathcal{F}}_B\right)\\ &\quad = \Delta \text{ actions}, \Delta \text{ effects}, \Delta \text{ intermediates (Translate using Eq. (6.14)).}\end{aligned} \tag{6.13}$$

It is necessary to transform the quantitative values into qualitative information for occupants (not domain-experts). These transformations are done using Eq. (6.1), which divides the value domain of a variable into 7 levels, using the following specifications:

$$\begin{aligned}&\Pi\left(\Delta\zeta_W^k, -0.7, -0.5, -0.2, 0.2, 0.5, 0.7\right),\\ &\Pi\left(\Delta\zeta_D^k, -0.7, -0.5, -0.2, 0.2, 0.5, 0.7\right),\\ &\Pi\left(\Delta\sigma_{temp}^k, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25\right),\\ &\Pi\left(\Delta\sigma_{air}^k, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25\right),\\ &\Pi\left(\Delta Q_{in}^k, -0.2, -0.1, -0.05, 0.05, 0.1, 0.2\right),\\ &\Pi\left(\Delta\varphi_{in}^k, -600, -400, -200, 200, 400, 600\right).\end{aligned} \tag{6.14}$$

For example, two scenarios from 05-May-2015 are analyzed over a period ranging from 8 am to 8 pm as shown in Fig. 6.13a. Thus, differential explanations inform how the occupants should change their schedule and the gain they can expect from this change. However, it does not explain which action is responsible for a particular effect. So, a deeper explanation process is considered next.

6.7.2 Differential Explanations with Influence

The optimal actions at different hours, suggested by the system, do not have the same importance in terms of impact. Some of them should necessarily be performed because of their strong influence on a particular criterion. To evaluate the influence of an action at the j^{th} hour, the difference is computed between the following two scenarios: (1) the optimal scenario, and (2) a modified scenario from the schedule ($\hat{\mathbf{X}}_B^j$) hypothesized by

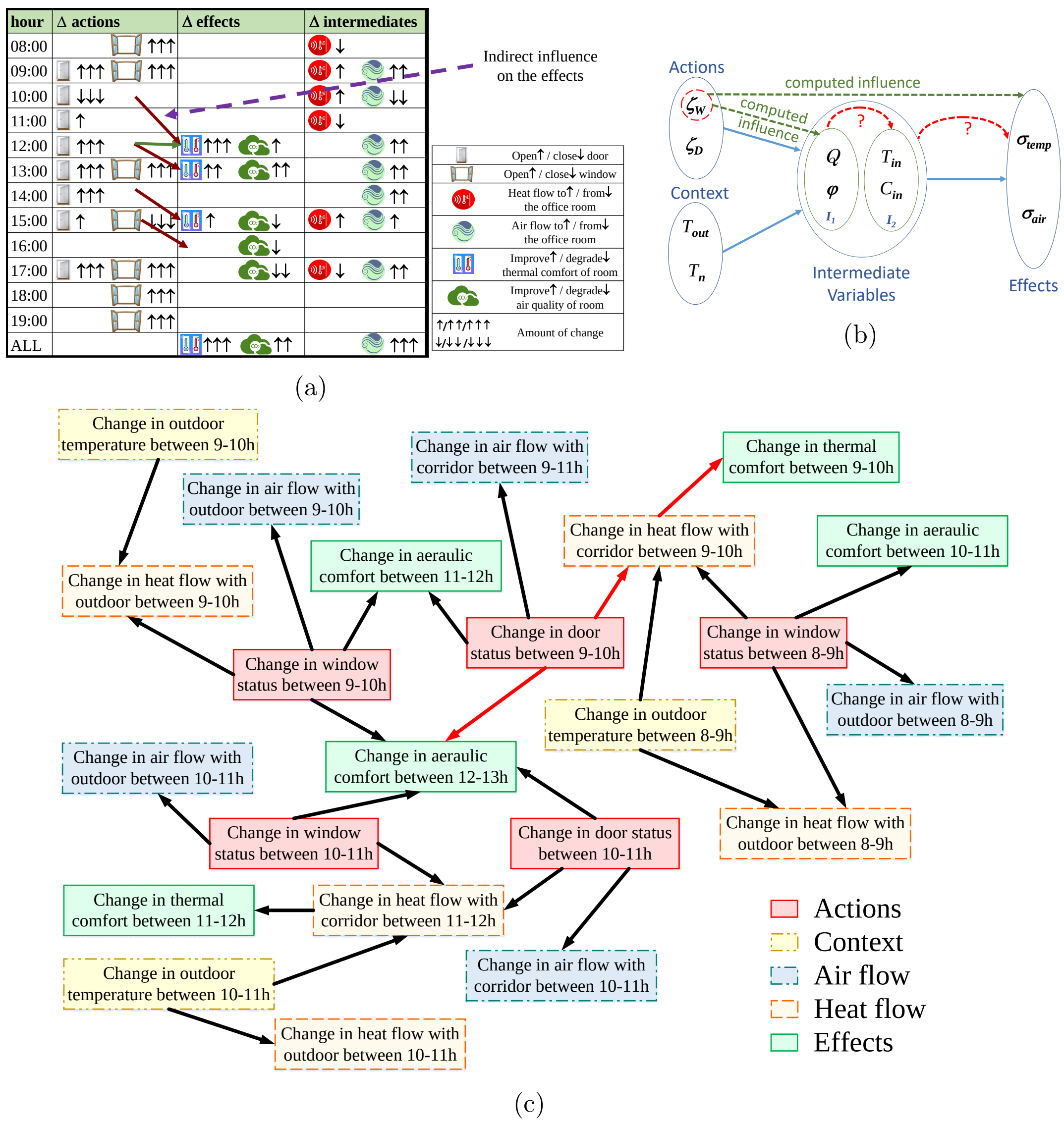


Figure 6.13: (a) Differential explanations and differential explanations with influence [133], (b) inexplicit information flow among the different categories of variables [133], and (c) causal graph to demonstrate the impact of actions on 05-May-2015 [133].

replacing an optimal action $\zeta^{\star,j}$ in $\mathbf{X}_B^{\star}$ with the actual action $\tilde{\zeta}^j$ as follows:

$$\hat{\mathbf{X}}_B^j = \left\{ \bigcup_{k \neq j} \zeta^{\star,k} \right\} \cup \left\{ \tilde{\zeta}^j \right\}. \tag{6.15}$$

Thus, the difference in actions at the j^{th} hour ($\Delta \mathbf{X}_B^j$) can be isolated as follows:

$$\Delta \mathbf{X}_B^j = \mathbf{X}_B^{\star} - \hat{\mathbf{X}}_B^j = \begin{cases} 0, & \forall k \neq j \\ \zeta^{\star,j} - \tilde{\zeta}^j, & k = j. \end{cases} \tag{6.16}$$

Now, using the same explanation generating approach the influence of ignoring the recommended action $\zeta^{\star,j}$ at the j^{th} hour can be obtained as follows:

$$\begin{aligned}&\left(\text{Optimal scenario: } \mathcal{X}_B^\star, \mathcal{P}_B \overset{\mathcal{I}_B^\star}{\to} \mathcal{F}_B^\star\right) - \left(\text{Modified scenario: } \hat{\mathcal{X}}_B^j, \mathcal{P}_B \overset{\hat{\mathcal{I}}_B^j}{\to} \hat{\mathcal{F}}_B^j\right)\\ &= \Delta\mathbf{X}_B^j \text{ influence } \Delta \text{ effects via } \Delta \text{ intermediates (Translate using Eq. (6.14))}\end{aligned} \tag{6.17}$$

A few computed influences are shown in Fig. 6.13a using arrows from some Δ actions to some Δ effects. However, the causality between the different groups of intermediate variables and the effects are unreachable (Fig. 6.13b) as their changes cannot be monitored with the physical knowledge models [159]. Such relations can only be injected using expert knowledge of potential and impossible causalities. For example, heat flow may influence air temperature but not CO_2 concentration. Thus, by integrating the computed influences and the potential influences, a full causal graph for the whole system can be obtained. Part of this graph is shown in Fig. 6.13c. Thus, the occupants can learn from these explanations whether a recommended action is important based on how various variables are affected by it.

6.7.3 Using the Building Energy Management Framework

The concerned framework [133] (Fig. 1.6) is essentially a human-machine-interaction interface which can take input from the occupants' to set various preferences and accordingly output an energy-efficient schedule of actions, equipped with simple explanations. The working of this interface (Fig. 6.14) is described as follows:

1. For a certain day (past or future), the context variables ($\mathcal{P}_B$, recorded or forecasted) and the action variables ($\tilde{\mathcal{X}}_B$, performed or planned) simulate the physical knowledge models, which assist in effects (objectives) evaluation ($\tilde{\mathcal{F}}_B$).

2. Using the same context ($\mathcal{P}_B$) and the physical models, the Pareto-optimal Set of schedules ($\mathcal{A}_{G_{max}}$) is obtained by minimizing occupants' discomfort (thermal and aeraulic), energy expenses and the number of recommended changes. Thereafter, the selection of the most-relevant scenario ($\mathcal{X}_B^\star$ with $\mathcal{F}_B^\star$) is guided by the appropriate decision-making strategy (Sections 6.6.1 to 6.6.5).

3. The recommended scenario ($\mathcal{X}_B^\star$ with $\mathcal{F}_B^\star$) is compared with the actual scenario ($\tilde{\mathcal{X}}_B$ with $\tilde{\mathcal{F}}_B$) to generate explanations. The occupants may opt-out of an optimal action

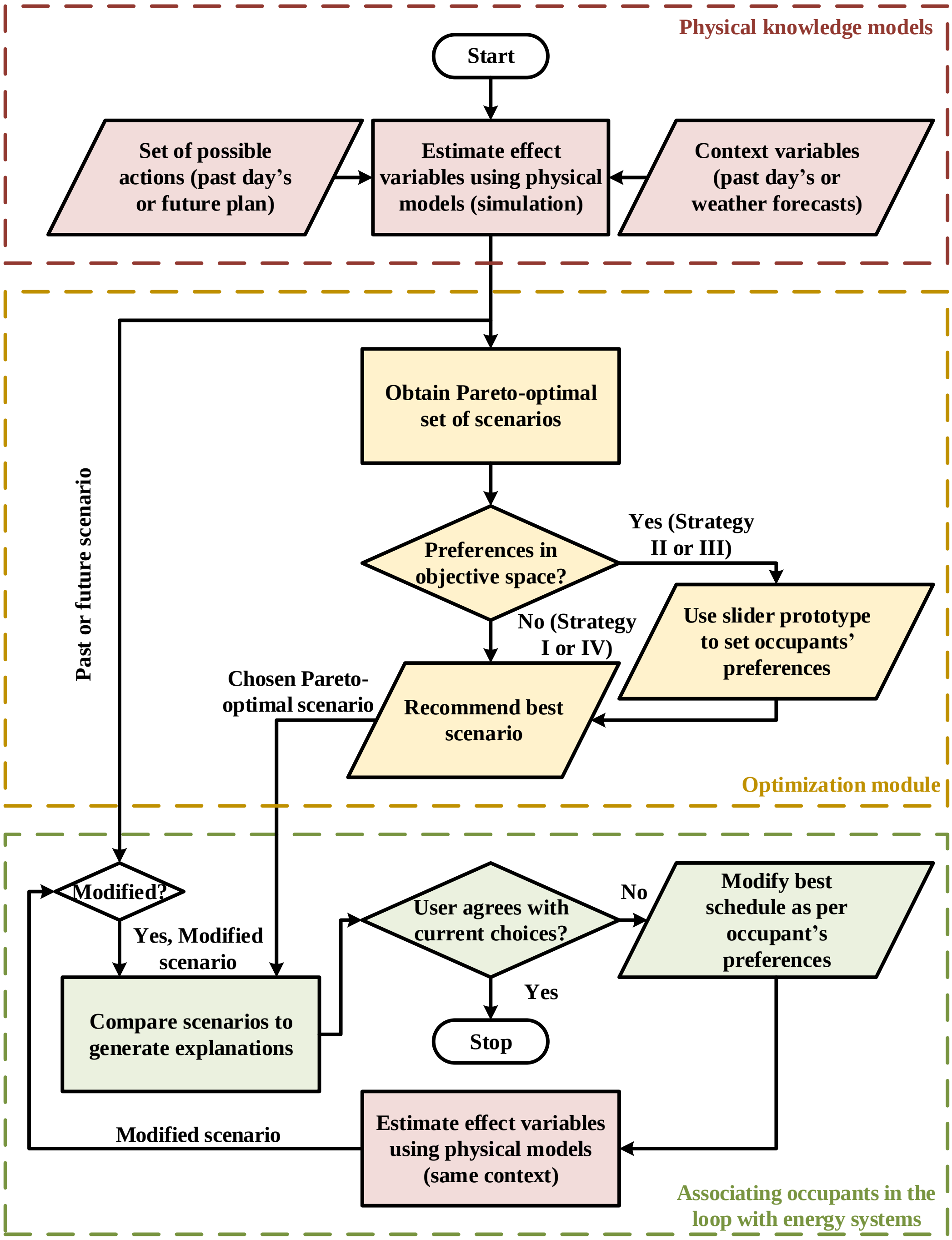


Figure 6.14: Flowchart of the energy manangement scheme for the office room [133].

at a certain hour. The modified scenario ($\hat{\mathcal{X}}_B$ with $\hat{\mathcal{F}}_B$ and $\mathcal{P}_B$) is again compared to the optimal scenario ($\mathcal{X}_B^\star$ with $\mathcal{F}_B^\star$) to yield the associated impacts of the change. This step loops until the occupants are satisfied with the recommendations.

Thus, the framework recommends the optimal actions to the occupants and associates them with the building energy systems for adapting their actions towards a greener future.

6.8 Conclusion

This chapter presents a framework for building energy management which recommends a relevant Pareto-optimal approximation of a schedule of occupants' actions by minimizing occupants' discomfort (thermal and aeraulic), energy expenses and the number of recommended changes. Furthermore, to motivate the occupants to adopt such a framework, it explains how the optimal plan determines the maximum gain in occupants' comfort by revealing the embedded physical phenomena. The primary application of this framework is to find contextually similar days in the recorded database, based on forecasted data for the next day and accordingly recommend the optimal plan of actions to the occupants.

In this chapter, the associated MaOO problem (search for the optimal schedule) is addressed using two algorithms developed in this thesis: NAEMO (for fast optimization) and LORD (for exploring the multi-modality). However, a practical application can implement only one relevant solution from the set of Pareto-optimal solutions, resulting from these algorithms. Hence, based on the different characterization of the occupants' preferences, various decision-making strategies are presented in this chapter for selecting the relevant schedule of occupants' actions.

While the strategies developed across all the previous chapters are capable of addressing various real-world optimization problems, there remain several avenues open for further research. A summary of the issues addressed in this thesis and various open areas are described in the concluding chapter, presented next.

Chapter 7

Conclusions and Scope of Further Research

7.1 Conclusions

This thesis is a comprehensive attempt to develop several computational strategies for improving the performance of EAs while tackling a variety of the MaOO problems. To deal with such problems, EAs are integrated with strategies like objective reduction and reference vector assisted decomposition of objective space, which aid in improving the solution distribution and selection pressure. Additionally, graph Laplacian based clustering of solutions in the decision space is performed to address the multi-modality of the optimization problems. Some of the developed EAs are applied to address the real-world many-objective building energy management problem (Chapter 6). This chapter also presents a few decision-making strategies (for varied scenarios) to recommend a context-relevant solution from the estimated set of Pareto-optimal solutions.

Chapter 2 presents IDEMO [142] with a revised elitist selection and ranking scheme (using a combination crowding distance with distance from the ideal point) to improve the selection pressure, convergence and diversity of the solutions. By integrating IDEMO in an online objective reduction framework, DECOR [142] is developed with a novel decision indicator for cohesive and distinctive clusters. When compared to several other EAs, DECOR shows superior convergence to PF on 10- and 20-objective DTLZ problems. However, the scope of improvement in its diversity characteristics motivates designing further better EAs for MaOO problems.

To improve the diversity attainment behavior of the EAs, Chapter 3 considers the reference vector assisted decomposition of objective space and presents ESOEA [138]. It uses an ensemble of SaNSDE [185] with PBI-based scalarization of MaOO problems. ESOEA adapts to the problem characteristics by adjusting the sub-population sizes in accordance with their contribution towards the global population. Furthermore, its regulated elitist scheme with $d2$-based sorting promotes further exploration. Results exhibit good convergence and superior diversity of ESOEA for test problems with attributes like multiple modalities, biased solution densities, disconnected PFs and PFs with sharp-tails, imbalance difficulties and variable linkage difficulties. However, the lack of any theoretical analysis hinders the understanding of the search behavior of such decomposition-based strategies.

To understand the working of such reference vector assisted decomposition based algorithms, Chapter 4 identifies the neighborhood property for MaOO problems [160]. It is used to develop NAEMO [160], where the neighborhood property dictates the solution mating for generating new solutions. Moreover, NAEMO monotonically improves the diversity through its periodic filtering module (proven using the novel *D_metric* [161]). Results establish the efficacy of NAEMO for several problem characteristics like unimodality, multi-modality, biased solution density, meta-variable mapping, imbalance mapping difficulty and variable linkage difficulty. While such algorithmic designs perform exceptionally well in the objective space, it does not consider the solution distribution in the decision space. Such an analysis of solution distribution in both the objective and decision space forms the basis of developing EAs for MMMOPs [171].

In Chapter 5, the crowding illusion problem for MMMOPs is identified and LORD is devised to deal with the challenges of MMMOPs. It uses graph Laplacian based clustering to maintain the solution diversity in the decision space and reference vector based decomposition to maintain the solution diversity in the objective space. The filtering module of LORD eliminates the maximally crowded solution from the last non-dominated rank. To avoid dominance resistance in problems with a large number of objectives, LORD-II is presented, which eliminates the candidate with maximal PBI from the maximally large cluster. Both LORD and LORD-II retain solutions from singleton sub-spaces during filtering for diversity maintenance. The efficacies of these EAs are established on CEC 2019 test suite [112] and polygon problems [76] (multi-modal or otherwise).

In Chapter 6, NAEMO (Chapter 4) and LORD (Chapter 5) are used to address the MaOO problem involved in building energy management. It aims at recommending an optimal schedule of actions by minimizing the thermal discomfort, air quality discomfort, energy-related cost and successive changes in recommended actions [133]. However, from the set of trade-off solutions resulting from the EAs, only one relevant solution could be recommended for implementation. Chapter 6 presents four distinct schedule selection strategies (considering equal compromise in all objectives, considering a slider prototype to interact with optimal objective ranges, considering different comfort preferences of multiple occupants [135] and considering minimum changes from the occupants' usual schedule [139]). Thereafter, the causal impact of the recommended changes is explained by comparing the recommended and the usual scenario. These explanations guide the occupants to adopt an energy-efficient schedule of actions. The next section presents the scope of extending the computational strategies developed in this thesis.

7.2 Limitations and Future Scope

This thesis develops several computational strategies beneficial for obtaining solutions from various kinds of MaOO problems. However, there remain a few areas open with scope for future study. Such areas are enlisted as follows:

- Performance analyses of these EAs on recent [107, 108] and minus problems [84] are necessary to investigate their search behavior for other problem characteristics.
- Some parameters (such as th in DECOR [142], P_{mut} in NAEMO [160], LORD and LORD-II [140], etc.) are required to be tuned. Hence, future studies can investigate the adaptation of such parameters to the fitness landscape.
- An advantage of ESOEA is its inherent parallelism (Fig. 3.2b), which could be further exploited to obtain much faster results. Similarly, the parallel implementation of the other EAs developed in this thesis can be explored for faster execution.
- With the recent spike in research works for MMMOPs, the possibility of better mating operators and performance indicators in the decision space could be foreseen. As the LORD variants [139, 140] use spectral clustering, inter- and intra-cluster mating along with correlation among the equivalent solution subsets can be studied.

- From the decision-making perspective, selecting one out of multiple equivalent solutions from the PS mapping to the same solution in the PF (i.e., decision-making without preference in decision space) is an important direction. Integrating imprecise preferences with the developed strategies is another vital future direction.

- Two major caveats surfaced while deploying the developed building energy management framework [133] in an extended human-machine interface [1] (Fig. 7.1). For any geographical location, extensive research on simulation models for various physical variables (like humidity and pollutants) and selection of necessary contextual parameters are necessary. Thus, hybridizing EAs with machine learning models (such as ensemble of neural networks [99]) could estimate the relevant physical variables.

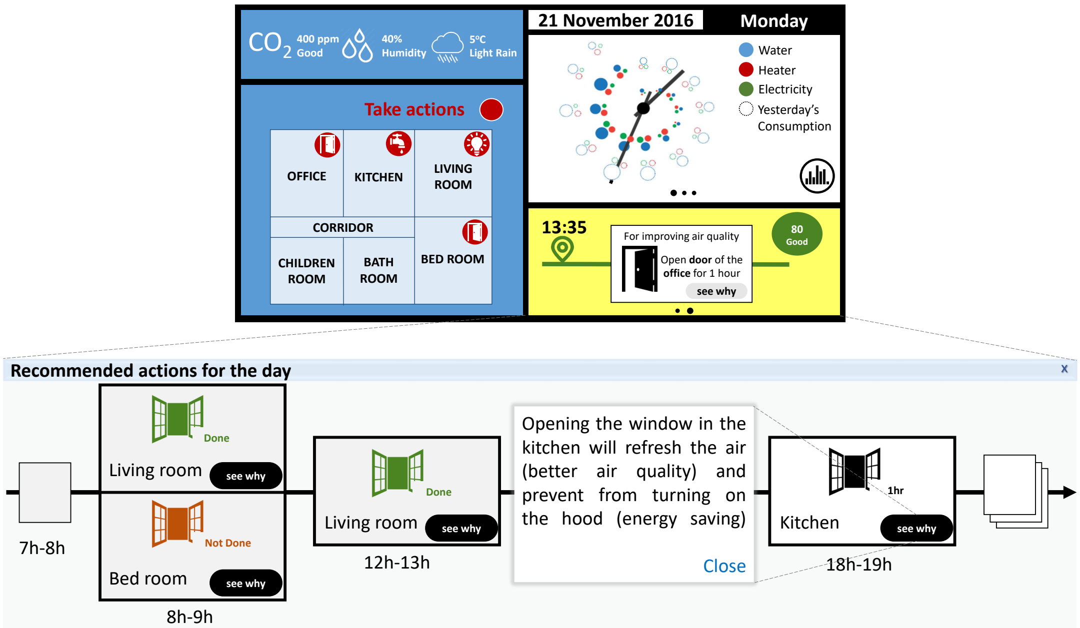


Figure 7.1: A real building energy management interface [1] using the developed approach of generating explanations for changes in occupants' actions.

This thesis develops several EAs for addressing a wide range of MaOO problems. These EAs can be utilized for enhancing the performance in different aspects (like dimensionality reduction [136, 189], model tuning [158], scheduling [131], and many more) of any real-life application. This research also contributes towards simplifying the decision-making process by selecting the relevant trade-off solutions. Although a significant amount of work has been done, as expected there is scope to do a lot more. Research in Many-Objective Optimization and its application in real-life domains will remain important in the coming years.

Appendices

Appendix A

Benchmark Test Problems

A.1 Deb, Thiele, Laumanns, Zitzler (DTLZ) Test Suite [50]

The performance of several algorithms (presented in this thesis) are tested on DTLZ test problems [50,141] among DTLZ1 (multimodal), DTLZ2 (unimodal), DTLZ3 (multimodal), DTLZ4 (biased, unimodal) and DTLZ7 (disconnected) [50, 74, 138]. These test problems are described below while illustrating their true Pareto-Fronts (PFs) in Fig. A.1.

A.1.1 DTLZ1 problem

This M-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = \frac{1}{2}x_1x_2\cdots x_{M-1}\left(1+h\left(\mathbf{X}_M\right)\right)\\
&\text{Minimize: } f_2(\mathbf{X}) = \frac{1}{2}x_1x_2\cdots\left(1-x_{M-1}\right)\left(1+h\left(\mathbf{X}_M\right)\right)\\
&\vdots\\
&\text{Minimize: } f_{M-1}(\mathbf{X}) = \frac{1}{2}x_1\left(1-x_2\right)\left(1+h\left(\mathbf{X}_M\right)\right)\\
&\text{Minimize: } f_M(\mathbf{X}) = \frac{1}{2}\left(1-x_1\right)\left(1+h\left(\mathbf{X}_M\right)\right)\\
&\text{subjected to } 0 \le x_i \le 1, \ \text{for } i = 1, 2, \cdots, N\\
&\text{where, } h\left(\mathbf{X}_M\right) = 100\left[|\mathbf{X}_M| + \sum_{x_i\in\mathbf{X}_M}\left\{(x_i-0.5)^2 - \cos(20\pi(x_i-0.5))\right\}\right]
\end{aligned} \tag{A.1}$$

Optimal PF of Eq. (A.1) is linear and corresponds to $x_i^\star = 0.5$ where $x_i^\star \in \mathbf{X}_M$ and $\sum_{i=1}^{M} f_i(\mathbf{X}) = 0.5$. According to the literature [50, 138, 142], $k_D = |\mathbf{X}_M| = 5$ and the number of variables defining the decision space is $N = M + k_D - 1 = M + 4$.

A.1.2 DTLZ2 problem

This M-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}_M)) \cos\left(x_1 \frac{\pi}{2}\right) \cdots \cos\left(x_{M-1} \frac{\pi}{2}\right) \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}_M)) \cos\left(x_1 \frac{\pi}{2}\right) \cdots \sin\left(x_{M-1} \frac{\pi}{2}\right) \\
&\vdots \\
&\text{Minimize: } f_{M-1}(\mathbf{X}) = (1 + h(\mathbf{X}_M)) \cos\left(x_1 \frac{\pi}{2}\right) \sin\left(x_2 \frac{\pi}{2}\right) \\
&\text{Minimize: } f_M(\mathbf{X}) = (1 + h(\mathbf{X}_M)) \sin\left(x_1 \frac{\pi}{2}\right) \\
&\text{subjected to } 0 \leq x_i \leq 1, \text{ for } i = 1, 2, \cdots, N \\
&\text{where, } h(\mathbf{X}_M) = \sum_{x_i \in \mathbf{X}_M} (x_i - 0.5)^2
\end{aligned} \tag{A.2}$$

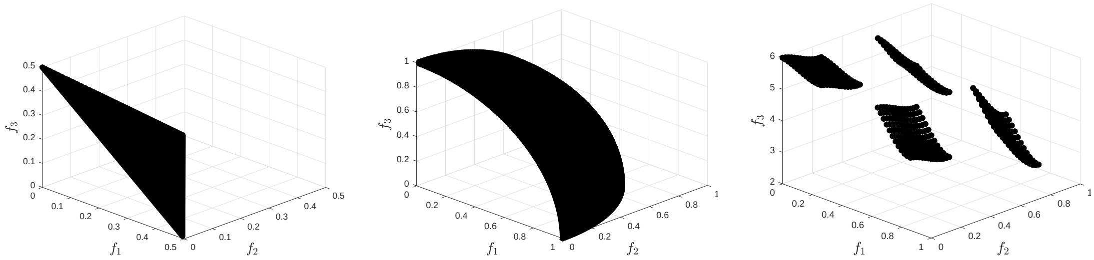
Optimal PF of Eq. (A.2) corresponds to $x_i^\star = 0.5$ where $x_i^\star \in \mathbf{X}_M$ and $\sum_{i=1}^{M} f_i^2(\mathbf{X}) = 1$. According to the literature [50, 138, 142], $k_D = |\mathbf{X}_M| = 10$ and the number of variables defining the decision space is $N = M + k_D - 1 = M + 9$.

A.1.3 DTLZ3 problem

This M-objective problem is defined similar to DTLZ2 except that the $h(.)$ function from DTLZ1 is used. Optimal PF of this problem corresponds to $x_i^\star = 0.5$ where $x_i^\star \in \mathbf{X}_M$ and $\sum_{i=1}^{M} f_i^2(\mathbf{X}) = 1$. According to the literature [50, 138, 142], $k_D = |\mathbf{X}_M| = 10$ and the number of variables defining the decision space is $N = M + k_D - 1 = M + 9$.

A.1.4 DTLZ4 Problem

This M-objective problem is a modification of the DTLZ2 problem. The modification involves a meta-variable mapping: $x_i \to x_i^{\alpha_D}$ while leads to biased density of solutions towards the f_1-f_M plane. Optimal PF of this problem corresponds to $x_i^\star = 0.5$ where $x_i^\star \in \mathbf{X}_M$ and $\sum_{i=1}^{M} f_i^2(\mathbf{X}) = 1$. According to the literature [50,138,142], $k_D = |\mathbf{X}_M| = 10$, $\alpha_D = 100$ and the number of variables defining the decision space is $N = M + k_D - 1 = M + 9$.



(a) DTLZ1 (b) DTLZ2, DTLZ3, DTLZ4 (c) DTLZ7

Figure A.1: Cartesian coordinate plots of true PFs for 3-objective DTLZ test instances.

A.1.5 DTLZ7 problem

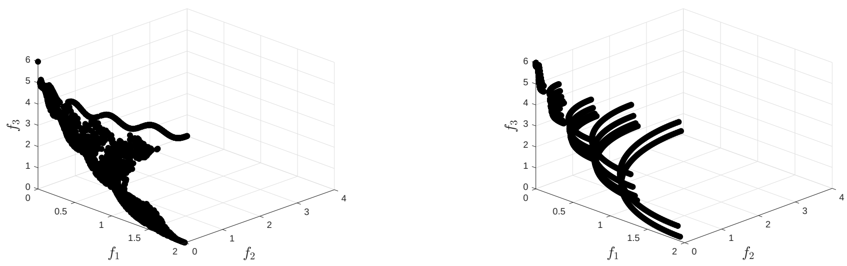
This M-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = x_2 \\
&\vdots \\
&\text{Minimize: } f_{M-1}(\mathbf{X}) = x_{M-1} \\
&\text{Minimize: } f_M(\mathbf{X}) = (1 + h_1(\mathbf{X}_M))h_2(f_1, f_2, \cdots, f_{M-1}, h_1) \\
&\text{subjected to } 0 \leq x_i \leq 1, \text{ for } i = 1, 2, \cdots, N \\
&\text{where, } h_1(\mathbf{X}_M) = 1 + \frac{9}{|\mathbf{X}_M|} \sum_{x_i \in \mathbf{X}_M} \text{ and} \\
&h_2(f_1, f_2, \cdots, f_{M-1}, h_1) = M - \sum_{i=1}^{M-1} \left[\frac{f_i}{1 + h_1} \left(1 + sin\left(3\pi f_i\right)\right) \right]
\end{aligned}
\tag{A.3}
$$

Pareto-optimal solutions of Eq. (A.3) correspond to $\mathbf{X}_M = \mathbf{0}$ and the optimal PF has 2^{M-1} disconnected regions. According to the literature [50,138], $k_D = |\mathbf{X}_M| = 20$ and the number of variables defining the decision space is $N = M + k_D - 1 = M + 19$.

A.2 Walking Fish Group (WFG) Test Suite [74]

The performance of several algorithms (presented in this thesis) are tested on WFG1 and WFG2 test problems [74]. These test problems are described below while illustrating their true PFs in Fig. A.2.



(a) WFG1 (b) WFG2

Figure A.2: Cartesian coordinate plots of true PFs for 3-objective WFG1-2 problems.

A.2.1 WFG1 problem

This M-objective problem is defined as:

$$
\begin{aligned}
\text{Given: } \mathbf{Z} =& \{z_1/2, z_2/4, \cdots, z_N/(2N)\} \\
\text{Minimize: } f_1(\mathbf{X}) =& \prod_{i=1}^{M-1} (1 - cos(x_i\pi/2)) \\
\text{Minimize: } f_2(\mathbf{X}) =& \prod_{i=1}^{M-2} (1 - cos(x_i\pi/2))\,(1 - sin(x_{M-1}\pi/2)) \\
&\vdots \\
\text{Minimize: } f_{M-1}(\mathbf{X}) =& (1 - cos(x_1\pi/2))\,(1 - sin(x_2\pi/2)) \\
\text{Minimize: } f_M(\mathbf{X}) =& \left(1 - x_1 - \frac{cos(10\pi x_1 + \pi/2)}{10\pi}\right) \\
\text{where } x_{i=1:M-1} =& r\_sum\left(\{y_{(i-1)k_W/(M-1)+1}, \cdots, y_{ik_W/(M-1)}\},\right. \\
& \left.\{2((i-1)k_W/(M-1)+1), \cdots, 2ik_W/(M-1)\}\right) \\
x_M =& r\_sum\left(\{y_{k_W+1}, \cdots, y_N\}, \{2(k_W+1), \cdots, 2N\}\right) \\
y_{i=1:N} =& b\_poly\left(y'_i, 0.02\right) \\
y'_{i=1:k_W} =& y''_i \\
y'_{i=(k_W+1):N} =& b\_flat(y''_i, 0.8, 0.75, 0.85) \\
y''_{i=1:k_W} =& z_i \\
y''_{i=(k_W+1):N} =& s\_linear(z_i, 0.35) \\
\text{with } r\_sum\left(|\mathbf{y}|, |\mathbf{w}|\right) =& \frac{\sum_{i=1}^{|\mathbf{y}|} w_i y_i}{\sum_{i=1}^{|\mathbf{y}|} w_i} \\
b\_poly(y, \alpha_W) =& y^{\alpha_W}
\end{aligned}
$$

$$\begin{aligned} b\_flat(y, A_W, B_W, C_W) =& A_W + min(0, \lfloor y - B_W \rfloor)\frac{A_W(B_W - y)}{B_W} \\ & - min(0, \lfloor C_W - y \rfloor)\frac{(1 - A_W)(y - C_W)}{1 - C_W} \\ s\_linear(y, D_W) =& \frac{|y - D_W|}{|\lfloor D_W - y \rfloor + D_W|} \\ & \text{subjected to } 0 \le z_i \le 1, \text{ for } i = 1, 2, \cdots, N \end{aligned} \tag{A.4}$$

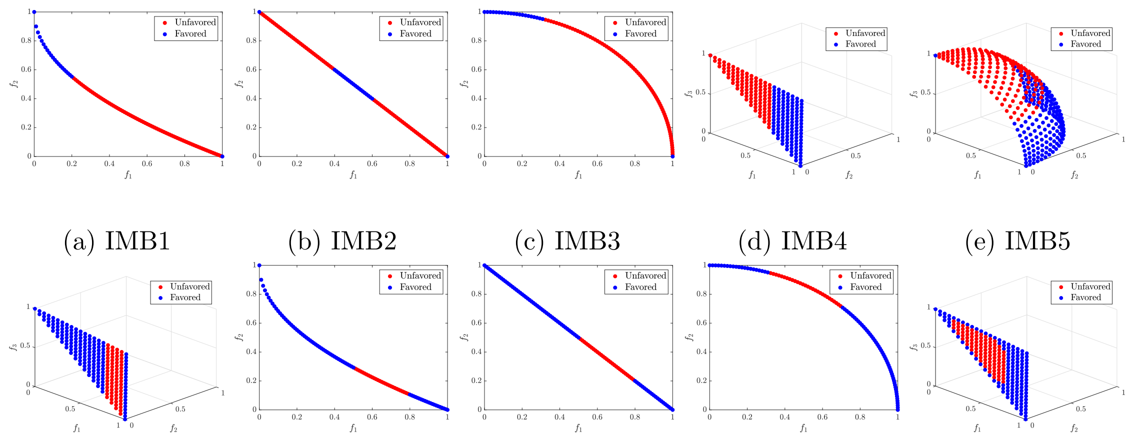
This problem is characterized as unimodal and has convex PF. A solution of WFG1 is Pareto-optimal iff $z_i = 0.35$, for $i = (k_W + 1), \cdots, N$. As per [187], WFG1 is realized with $k_W = N - M + 1$ distance related variables and $(M - 1)$ position related variables.

A.2.2 WFG2 problem

This M-objective problem is defined as:

$$\begin{aligned} \text{Given: } \mathbf{Z} =& \{z_1/2, z_2/4, \cdots, z_N/(2N)\} \\ \text{Minimize: } f_1(\mathbf{X}) =& \prod_{i=1}^{M-1} (1 - cos(x_i\pi/2)) \\ \text{Minimize: } f_2(\mathbf{X}) =& \prod_{i=1}^{M-2} (1 - cos(x_i\pi/2))\,(1 - sin(x_{M-1}\pi/2)) \\ & \vdots \\ \text{Minimize: } f_{M-1}(\mathbf{X}) =& (1 - cos(x_1\pi/2))\,(1 - sin(x_2\pi/2)) \\ \text{Minimize: } f_M(\mathbf{X}) =& \left(1 - x_1 cos^2(5\pi x_1)\right) \\ \text{where } x_{i=1:M-1} =& r\_sum\left(\{y_{(i-1)k_W/(M-1)+1}, \cdots, y_{ik_W/(M-1)}\}, \{1, \cdots, 1\}\right) \\ x_M =& r\_sum\left(\{y_{k_W+1}, \cdots, y_{k_W+l/2}\}, \{1, \cdots, 1\}\right) \\ y_{i=1:k_W} =& y'_i \\ y_{i=(k_W+1):(k_W+l_W/2)} =& r\_nonsep(\{y'_{k_W+2(i-k_W)-1}, y_{k_W+2(i-k_W)}\}, 2) \\ y'_{i=1:k_W} =& z_i \\ y'_{i=(k_W+1):N} =& s\_linear(z_i, 0.35) \\ & \text{with } r\_sum\,(.) \text{ and } s\_linear\,(.) \text{ same as Eq. (A.4)} \\ & \text{subjected to } 0 \le z_i \le 1, \text{ for } i = 1, 2, \cdots, N \end{aligned} \tag{A.5}$$

This problem is characterized as unimodal, non-separable and has convex and dis-



(a) IMB1 (b) IMB2 (c) IMB3 (d) IMB4 (e) IMB5

(f) IMB6 (g) IMB7 (h) IMB8 (i) IMB9 (j) IMB10

Figure A.3: Cartesian coordinate plots of true PFs for imbalanced multi-objective test instances showing favored parts in blue and unfavored parts in red.

connected regions in the PF. A solution of WFG2 is Pareto-optimal iff $z_i = 0.35$, for $i = (k_W + 1), \cdots, N$. As per [187], WFG2 is realized with $k_W = N - M + 1$ distance related variables and $l_W = (M-1)$ position related variables. It also requires even number of distance related variables due to its nature of non-separable reductions ($r\_nonsep\,(.)$).

A.3 Imbalanced Multi-objective Test Suite [115]

Imbalanced test suite [115] consists of 10 unconstrained (box-constrained) multi-objective problems characterized by regions which are difficult to find as illustrated in their true PF in Fig. A.3. In this test suite, IMB1 to IMB6 problems demonstrate imbalance mapping difficulties whereas IMB7 to IMB10 problems demonstrate variable linkage difficulties. For analyzing the versatility of the algorithms presented in this thesis, their performance is noted on various instances from this test suite.

A.3.1 IMB1 problem

This 2-objective problem is defined as:

$$\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1$$

$$\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, \sqrt{1 - x_1}$$

$$\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 0 \leq x_1 \leq 0.2 \\ \sum_{j=2}^{N} 0.5 \times \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases}$$

$$\text{with } u_j = x_j - sin(0.5\pi x_1), \text{ for } j = 2, \cdots, N$$
$$\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, \cdots, N \tag{A.6}$$

Optimal PF of Eq. (A.6) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. The favored part of the PF is within $0 \leq f_1(\mathbf{X}) \leq 0.2$ and the $(1, 0)$ point in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $0.2 \leq x_1 \leq 1$ and $j = 2, \cdots, N$.

A.3.2 IMB2 problem

This 2-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, (1 - x_1) \\
&\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 0.4 \leq x_1 \leq 0.6 \\ \sum_{j=2}^{N} 0.5 \times \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - sin(0.5\pi x_1), \text{ for } j = 2, \cdots, N \\
&\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.7}$$

Optimal PF of Eq. (A.7) corresponds to $f_2(\mathbf{X}) = 1 - f_1(\mathbf{X})$ with $0 \leq f_1(\mathbf{X}) \leq 1$. The favored part of the PF is within $0.4 \leq f_1(\mathbf{X}) \leq 0.6$ along with the $(0, 1)$ and $(1, 0)$ points in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $x_1 \in [0, 0.4) \cup (0.6, 1]$ and $j = 2, \cdots, N$.

A.3.3 IMB3 problem

This 2-objective problem is defined as:

$$\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, cos\left(\frac{\pi x_1}{2}\right)$$
$$\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, sin\left(\frac{\pi x_1}{2}\right)$$

$$
\begin{aligned}
&\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 0.8 \le x_1 \le 1 \\ \sum_{j=2}^{N} 0.5 \times \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - sin(0.5\pi x_1),\ \text{for } j = 2, \cdots, N \\
&\text{subjected to } 0 \le x_j \le 1,\ \text{for } j = 1, \cdots, N
\end{aligned} \tag{A.8}
$$

Optimal PF of Eq. (A.8) corresponds to $\{f_1(\mathbf{X})\}^2 + \{f_2(\mathbf{X})\}^2 = 1$. The favored part of the PF is within $0 \le f_1(\mathbf{X}) \le 0.309$ and the $(1, 0)$ point in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $0 \le x_1 \le 0.8$ and $j = 2, \cdots, N$.

A.3.4 IMB4 problem

This 3-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 x_2 \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 (1 - x_2) \\
&\text{Minimize: } f_3(\mathbf{X}) = (1 + h(\mathbf{X}))\, (1 - x_1) \\
&\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 2/3 \le x_1 \le 1 \\ 2cos\left(\frac{\pi x_1}{2}\right) \sum_{j=3}^{N} \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - (x_1 + x_2)/2,\ \text{for } j = 3, \cdots, N \\
&\text{subjected to } 0 \le x_j \le 1,\ \text{for } j = 1, \cdots, N
\end{aligned} \tag{A.9}
$$

Optimal PF of Eq. (A.9) corresponds to $f_1(\mathbf{X}) + f_2(\mathbf{X}) + f_3(\mathbf{X}) = 1$ with $0 \le (f_1(\mathbf{X}), f_2(\mathbf{X}), f_3(\mathbf{X})) \le 1$. The favored part of the PF is within $0 \le f_3(\mathbf{X}) \le 1/3$ and the $(0, 0, 1)$ point in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from linear relationships among variables: $x_j = (x_1 + x_2)/2$ with $0 \le (x_1, x_2) \le 1$ and $j = 3, \cdots, N$.

A.3.5 IMB5 problem

This 3-objective problem is defined as:

$$
\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)
$$

$$\begin{aligned}
&\text{Minimize: } f_2(\mathbf{X}) = (1+h(\mathbf{X}))\, cos\left(\frac{\pi x_1}{2}\right) sin\left(\frac{\pi x_2}{2}\right)\\
&\text{Minimize: } f_3(\mathbf{X}) = (1+h(\mathbf{X}))\, sin\left(\frac{\pi x_1}{2}\right)\\
&\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 0 \le x_1 \le 0.5 \\ 2cos\left(\frac{\pi x_1}{2}\right)\sum_{j=3}^{N}\left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases}\\
&\text{with } u_j = x_j - (x_1+x_2)/2, \text{ for } j = 3, \cdots, N\\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.10}$$

Optimal PF of Eq. (A.10) corresponds to $\{f_1(\mathbf{X})\}^2 + \{f_2(\mathbf{X})\}^2 + \{f_3(\mathbf{X})\}^2 = 1$ with $0 \le (f_1(\mathbf{X}), f_2(\mathbf{X}), f_3(\mathbf{X})) \le 1$. The favored part of the PF is within $0 \le f_3(\mathbf{X}) \le \sqrt{2}/2$ and the point $(0,0,1)$ in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from linear relationships among variables: $x_j = (x_1+x_2)/2$ with $0 \le (x_1, x_2) \le 1$ and $j = 3, \cdots, N$.

A.3.6 IMB6 problem

This 3-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1+h(\mathbf{X}))\, x_1 x_2\\
&\text{Minimize: } f_2(\mathbf{X}) = (1+h(\mathbf{X}))\, x_1(1-x_2)\\
&\text{Minimize: } f_3(\mathbf{X}) = (1+h(\mathbf{X}))\,(1-x_1)\\
&\text{where } h(\mathbf{X}) = \begin{cases} 0, & \text{if } 0 \le x_1 \le 0.75 \\ 2cos\left(\frac{\pi x_1}{2}\right)\sum_{j=3}^{N}\left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{otherwise} \end{cases}\\
&\text{with } u_j = x_j - (x_1+x_2)/2, \text{ for } j = 3, \cdots, N\\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.11}$$

Optimal PF of Eq. (A.11) corresponds to $f_1(\mathbf{X}) + f_2(\mathbf{X}) + f_3(\mathbf{X}) = 1$ with $0 \le (f_1(\mathbf{X}), f_2(\mathbf{X}), f_3(\mathbf{X})) \le 1$. The favored part of the PF is within $0.25 \le f_3(\mathbf{X}) \le 1$ and along $f_3 = 1$ in the objective space. The remaining of the PF is the unfavored (difficult to explore) part which comes from linear relationships among variables: $x_j = (x_1+x_2)/2$ with $0 \le (x_1, x_2) \le 1$ and $j = 3, \cdots, N$.

A.3.7 IMB7 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, (1 - \sqrt{x_1}) \\
&\text{where } h(\mathbf{X}) = \begin{cases} \sum_{j=2}^{N} \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{if } 0.5 \leq x_1 \leq 0.8 \\ \sum_{j=2}^{N} |x_j - 0.5|^{0.6}, & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - sin(0.5\pi x_1), \text{ for } j = 2, \cdots, N \\
&\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.12}
$$

Optimal PF of Eq. (A.12) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. The favored part of the PF is within $f_1(\mathbf{X}) \in [0, 0.5] \cup [0.8, 1]$. The remaining of the PF is the unfavored (difficult to explore) part which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $0.5 < x_1 < 0.8$ and $j = 2, \cdots, N$.

A.3.8 IMB8 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, (1 - x_1) \\
&\text{where } h(\mathbf{X}) = \begin{cases} \sum_{j=2}^{N} \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{if } 0.5 \leq x_1 \leq 0.8 \\ \sum_{j=2}^{N} |x_j - 0.5|^{0.6}, & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - sin(0.5\pi x_1), \text{ for } j = 2, \cdots, N \\
&\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.13}
$$

Optimal PF of Eq. (A.13) corresponds to $f_2(\mathbf{X}) = 1 - f_1(\mathbf{X})$ with $0 \leq f_1(\mathbf{X}) \leq 1$. The favored and unfavored part of the PF for IMB8 problem is same as that of IMB7 problem except that for the for the parts lie on a linear PF whereas for the latter the parts lie on a parabolic PF.

A.3.9 IMB9 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, cos\left(\frac{\pi x_1}{2}\right) \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, sin\left(\frac{\pi x_1}{2}\right) \\
&\text{where } h(\mathbf{X}) = \begin{cases} \sum_{j=2}^{N} \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{if } 0.5 \le x_1 \le 0.8 \\ \sum_{j=2}^{N} |x_j - 0.5|^{0.6}, & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - sin(0.5\pi x_1), \text{ for } j = 2, \cdots, N \\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.14}
$$

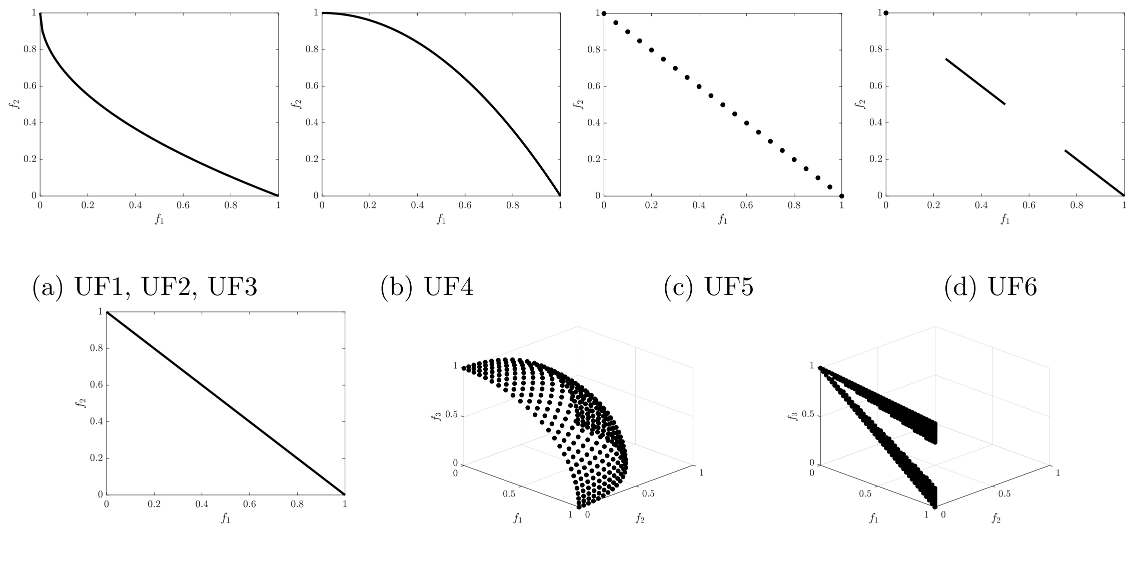
Optimal PF of Eq. (A.14) corresponds to $\{f_1(\mathbf{X})\}^2 + \{f_2(\mathbf{X})\}^2 = 1$ with $0 \le (f_1(\mathbf{X}), f_2(\mathbf{X})) \le 1$. The favored part of the PF is within $f_1(\mathbf{X}) \in [0, 0.309] \cup [0.707, 1]$. The remaining of the PF is the unfavored (difficult to explore) part which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $0.5 < x_1 < 0.8$ and $j = 2, \cdots, N$.

A.3.10 IMB10 problem

This 3-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 x_2 \\
&\text{Minimize: } f_2(\mathbf{X}) = (1 + h(\mathbf{X}))\, x_1 (1 - x_2) \\
&\text{Minimize: } f_3(\mathbf{X}) = (1 + h(\mathbf{X}))\, (1 - x_1) \\
&\text{where } h(\mathbf{X}) = \begin{cases} \sum_{j=2}^{N} \left(-0.9u_j^2 + |u_j|^{0.6}\right), & \text{if } 0.2 \le (x_1, x_2) \le 0.8 \\ \sum_{j=2}^{N} |x_j - x_1 x_2|^{0.6}, & \text{otherwise} \end{cases} \\
&\text{with } u_j = x_j - (x_1 + x_2)/2, \text{ for } j = 3, \cdots, N \\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, \cdots, N
\end{aligned} \tag{A.15}
$$

Optimal PF of Eq. (A.15) corresponds to $f_1(\mathbf{X}) + f_2(\mathbf{X}) + f_3(\mathbf{X}) = 1$ with $0 \le (f_1(\mathbf{X}), f_2(\mathbf{X}), f_3(\mathbf{X})) \le 1$. The unfavored (difficult to explore) part of the PF is within $0.04 \le f_1(\mathbf{X}) \le 0.64$ and $0.2 \le f_3(\mathbf{X}) \le 0.8$ which comes from non-linear combination of variables: $x_j = sin(0.5\pi x_1)$ with $0.2 < (x_1, x_2) < 0.8$ and $j = 3, \cdots, N$.



(a) UF1, UF2, UF3 (b) UF4 (c) UF5 (d) UF6

(e) UF7 (f) UF8, UF10 (g) UF9

Figure A.4: Cartesian coordinate plots of true PFs for multi-objective test instances from CEC 2009 session.

A.4 CEC 2009 Multi-objective Problems [191]

As the ensemble based algorithms from literature [101, 182, 195] have been tested on this test suite, the performance of ESOEA (presented in this thesis) is also analyzed on this test suite for comparison. The CEC 2009 test suite [191] consists of 10 unconstrained (box-constrained) multi-objective problems whose true PF are illustrated in Fig. A.4. In this thesis, $N = 30$ is considered as per [138, 191].

A.4.1 UF1 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 + \frac{2}{|J_1|}\sum_{j\in J_1} u_j^2 \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt{x_1} + \frac{2}{|J_2|}\sum_{j\in J_2} u_j^2 \\
&\text{where, } J_1 = \{j|j \text{ is odd and } 2 \le j \le N\} \text{ and } J_2 = \{j|j \text{ is even and } 2 \le j \le N\} \\
&\text{and } u_j = x_j - sin\left(6\pi x_1 + \frac{j\pi}{N}\right),\ \text{for } j = 2, \cdots, N, \\
&\text{subjected to } 0 \le x_1 \le 1 \text{ and } -1 \le x_j \le 1,\ \text{for } j = 2, \cdots, N
\end{aligned}
\tag{A.16}
$$

Optimal PF of Eq. (A.16) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$.

A.4.2 UF2 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 + \frac{2}{|J_1|}\sum_{j\in J_1} u_j^2 \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt{x_1} + \frac{2}{|J_2|}\sum_{j\in J_2} u_j^2 \\
&\text{where, } J_1 \text{ and } J_2 \text{ are same as in Eq. (A.16)} \\
&\text{and } u_j = \begin{cases} x_j - \left[0.3x_1^2 cos\left(24\pi x_1 + \frac{4j\pi}{N}\right) + 0.6x_1\right] cos\left(6\pi x_1 + \frac{j\pi}{N}\right) & \text{if } j \in J_1 \\ x_j - \left[0.3x_1^2 cos\left(24\pi x_1 + \frac{4j\pi}{N}\right) + 0.6x_1\right] sin\left(6\pi x_1 + \frac{j\pi}{N}\right) & \text{if } j \in J_2 \end{cases} \\
&\text{subjected to } 0 \le x_1 \le 1 \text{ and } -1 \le x_j \le 1, \text{ for } j = 2, \cdots, N
\end{aligned}
\tag{A.17}
$$

Optimal PF of Eq. (A.17) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$.

A.4.3 UF3 problem

This 2-objective problem is defined as:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 + \frac{2}{|J_1|}\left(4 + \sum_{j\in J_1} u_j^2 - 2\prod_{j\in J_1} cos\left(\frac{20\pi u_j}{\sqrt{j}}\right) + 2\right) \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt{x_1} + \frac{2}{|J_2|}\left(4 + \sum_{j\in J_2} u_j^2 - 2\prod_{j\in J_2} cos\left(\frac{20\pi u_j}{\sqrt{j}}\right) + 2\right) \\
&\text{where, } J_1 \text{ and } J_2 \text{ are same as in Eq. (A.16)}
\end{aligned}
\tag{A.18}
$$

$$
\begin{aligned}
&\text{and } u_j = x_j - x_1^{0.5\left(1+\frac{3(j-2)}{N-2}\right)}, j = 2, \cdots, N \\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, 2, \cdots, N
\end{aligned}
\tag{A.19}
$$

Optimal PF of Eq. (A.19) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$.

A.4.4 UF4 problem

This 2-objective problem is defined as:

$$
\text{Minimize: } f_1(\mathbf{X}) = x_1 + \frac{2}{|J_1|}\sum_{j\in J_1} h\left(u_j\right)
$$

$$\begin{aligned}
&\text{Minimize: } f_2(\mathbf{X}) = 1 - x_1^2 + \frac{2}{|J_2|}\sum_{j\in J_2} h(u_j) \\
&\text{where, } J_1, J_2 \text{ and } u_j \text{ are same as in Eq. (A.16), and } h(u_j) = \frac{|u_j|}{1+e^{2|u_j|}} \\
&\text{subjected to } 0 \le x_1 \le 1 \text{ and } -2 \le x_j \le 2, \text{ for } j = 2, \cdots, N
\end{aligned} \tag{A.20}$$

Optimal PF of Eq. (A.20) corresponds to $f_2(\mathbf{X}) = 1 - \{f_1(\mathbf{X})\}^2$ with $0 \le f_1(\mathbf{X}) \le 1$.

A.4.5 UF5 problem

This 2-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 + \left(\frac{1}{2k_U} + \varepsilon_U\right)|sin\,(2k_U\pi x_1)| + \frac{2}{|J_1|}\sum_{j\in J_1} h(u_j) \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - x_1 + \left(\frac{1}{2k_U} + \varepsilon_U\right)|sin\,(2k_U\pi x_1)| + \frac{2}{|J_2|}\sum_{j\in J_2} h(u_j) \\
&\text{where, } J_1, J_2 \text{ and } u_j \text{ are same as in Eq. (A.16),} \\
&h(u_j) = 2u_j^2 - cos\,(4\pi u_j) + 1, k_U \text{ is a positive integer and } \varepsilon_U > 0 \\
&\text{subjected to } 0 \le x_1 \le 1 \text{ and } -1 \le x_j \le 1, \text{ for } j = 2, \cdots, N
\end{aligned} \tag{A.21}$$

Optimal PF of Eq. (A.21) consists of $(2k_U+1)$ points having coordinates at $(i/2k_U, 1-(i/2k_U))$ for $i = 0, 1, \cdots, 2k_U$. In this thesis, $k_U = 10$ and $\varepsilon_U = 0.1$ are considered as per [138, 191].

A.4.6 UF6 problem

This 2-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 + max\,\{0, h(x_1)\} + \frac{2}{|J_1|}\left(4 + \sum_{j\in J_1} u_j^2 - 2\prod_{j\in J_1} cos\left(\frac{20\pi u_j}{\sqrt{j}}\right) + 2\right) \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - x_1 + max\,\{0, h(x_1)\} + \frac{2}{|J_2|}\left(4 + \sum_{j\in J_2} u_j^2 - 2\prod_{j\in J_2} cos\left(\frac{20\pi u_j}{\sqrt{j}}\right) + 2\right) \\
&\text{where, } J_1, J_2 \text{ and } u_j \text{ are same as in Eq. (A.16)} \\
&\text{and } h(u_j) = 2\left(\frac{1}{2k_U} + \varepsilon_U\right) sin\,(2k_U\pi u_j) \text{ with } k_U \text{ as a positive integer and } \varepsilon_U > 0 \\
&\text{subjected to } 0 \le x_1 \le 1 \text{ and } -1 \le x_j \le 1, \text{ for } j = 2, \cdots, N
\end{aligned} \tag{A.22}$$

Optimal PF of Eq. (A.22) consists of one isolated point at $(0, 1)$ and k_U disconnected parts lying on $f_2(X) = 1 - f_1(\mathbf{X})$ where $f_1(\mathbf{X}) \in \cup_{i=1}^{k_U} ((2i-1)/(2k_U), i/k_U)$. In this thesis, $k_U = 2$ and $\varepsilon_U = 0.1$ are considered as per [138, 191].

A.4.7 UF7 problem

This 2-objective problem is defined as:

$$\text{Minimize: } f_1(\mathbf{X}) = \sqrt[5]{x_1} + \frac{2}{|J_1|} \sum_{j \in J_1} u_j^2$$

$$\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt[5]{x_1} + \frac{2}{|J_2|} \sum_{j \in J_2} u_j^2$$

where, J_1, J_2 and u_j are same as in Eq. (A.16)

$$\text{subjected to } 0 \le x_1 \le 1 \text{ and } -1 \le x_j \le 1, \text{ for } j = 2, \cdots, N \tag{A.23}$$

Optimal PF of Eq. (A.23) corresponds to $f_2(\mathbf{X}) = 1 - f_1(\mathbf{X})$ with $0 \le f_1(\mathbf{X}) \le 1$.

A.4.8 UF8 problem

This 3-objective problem is defined as:

$$\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi}{2}x_1\right) cos\left(\frac{\pi}{2}x_2\right) + \frac{2}{|J_1|} \sum_{j \in J_1} u_j^2$$

$$\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi}{2}x_1\right) sin\left(\frac{\pi}{2}x_2\right) + \frac{2}{|J_2|} \sum_{j \in J_2} u_j^2$$

$$\text{Minimize: } f_3(\mathbf{X}) = sin\left(\frac{\pi}{2}x_1\right) + \frac{2}{|J_3|} \sum_{j \in J_3} u_j^2$$

where, $J_1 = \{j | 3 \le j \le N, \text{ and } j-1 \text{ is a multiple of } 3\}$,

$J_2 = \{j | 3 \le j \le N, \text{ and } j-2 \text{ is a multiple of } 3\}$,

$J_3 = \{j | 3 \le j \le N, \text{ and } j-3 \text{ is a multiple of } 3\}$

and $u_j = x_j - 2x_2 sin\left(2\pi x_1 + \dfrac{j\pi}{N}\right)$, for $j = 3, \cdots, N$,

$$\text{subjected to } 0 \le x_i \le 1 \text{ and } -2 \le x_j \le 2, \text{ for } i = 1, 2 \text{ and } j = 3, \cdots, N \tag{A.24}$$

Optimal PF of Eq. (A.24) corresponds to $\{f_1(\mathbf{X})\}^2 + \{f_2(\mathbf{X})\}^2 + \{f_3(\mathbf{X})\}^2 = 1$ with $0 \le f_i(\mathbf{X}) \le 1$ for $i = 1, 2, 3$.

A.4.9 UF9 problem

This 3-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = 0.5\left[max\left\{0,(1+\varepsilon_U)(1-4(2x_1-1)^2)\right\}+2x_1\right]x_2+\frac{2}{|J_1|}\sum_{j\in J_1}u_j^2\\
&\text{Minimize: } f_2(\mathbf{X}) = 0.5\left[max\left\{0,(1+\varepsilon_U)(1-4(2x_1-1)^2)\right\}-2x_1+2\right]x_2+\frac{2}{|J_2|}\sum_{j\in J_2}u_j^2\\
&\text{Minimize: } f_3(\mathbf{X}) = 1-x_2+\frac{2}{|J_3|}\sum_{j\in J_3}u_j^2\\
&\text{where, } J_1, J_2, J_3 \text{ and } u_j \text{ are same as in Eq. (A.24) and } \varepsilon_U>0,\\
&\text{subjected to } 0\le x_i\le 1 \text{ and } -2\le x_j\le 2, \text{ for } i=1,2 \text{ and } j=3,\cdots,N \qquad \text{(A.25)}
\end{aligned}$$

Optimal PF of Eq. (A.25) has two parts corresponding to Eq. (A.26) and Eq. (A.27). In this thesis, $\varepsilon_U = 0.1$ is considered as per [138, 191].

$$\begin{aligned}
&0\ge f_3\ge 1,\\
&0\ge f_1\ge \frac{1}{4}(1-f_3),\\
&f_2 = 1-f_1-f_3
\end{aligned}\qquad \text{(A.26)}$$

$$\begin{aligned}
&0\ge f_3\ge 1,\\
&\frac{3}{4}(1-f_3)\ge f_1\ge 1,\\
&f_2 = 1-f_1-f_3
\end{aligned}\qquad \text{(A.27)}$$

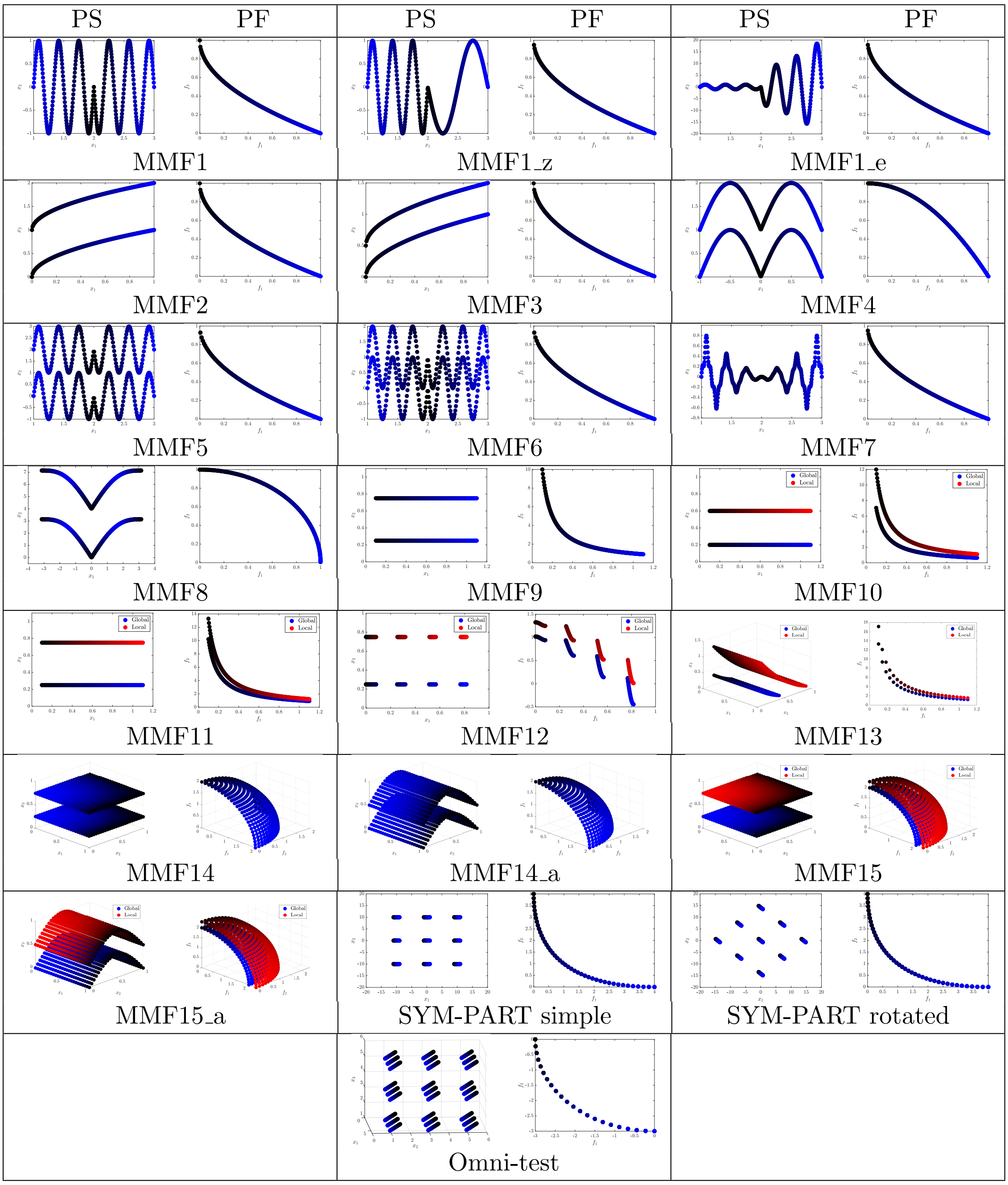
A.4.10 UF10 problem

This 3-objective problem is defined as:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi}{2}x_1\right)cos\left(\frac{\pi}{2}x_2\right)+\frac{2}{|J_1|}\sum_{j\in J_1}\left[4u_j^2-cos\left(8\pi u_j\right)+1\right]\\
&\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi}{2}x_1\right)sin\left(\frac{\pi}{2}x_2\right)+\frac{2}{|J_2|}\sum_{j\in J_2}\left[4u_j^2-cos\left(8\pi u_j\right)+1\right]\\
&\text{Minimize: } f_3(\mathbf{X}) = sin\left(\frac{\pi}{2}x_1\right)+\frac{2}{|J_3|}\sum_{j\in J_3}\left[4u_j^2-cos\left(8\pi u_j\right)+1\right]\\
&\text{where, } J_1, J_2, J_3 \text{ and } u_j \text{ are same as in Eq. (A.24)}\\
&\text{subjected to } 0\le x_i\le 1 \text{ and } -2\le x_j\le 2, \text{ for } i=1,2 \text{ and } j=3,\cdots,N \qquad \text{(A.28)}
\end{aligned}$$

Optimal PF of Eq. (A.28) corresponds to $\{f_1(\mathbf{X})\}^2+\{f_2(\mathbf{X})\}^2+\{f_3(\mathbf{X})\}^2 = 1$ with $0\le f_i(\mathbf{X})\le 1$ for $i=1,2,3$.

Table A.1: Cartesian coordinate plots of true PSs and true PFs for multi-modal multi-objective test instances from CEC 2019 session showing global surfaces in shades of blue and local surfaces in shades of red.



A.5 CEC 2019 Multi-Modal Multi-objective Problems [112]

As the multi-modal multi-objective evolutionary algorithms (MMMOEAs) from literature [91, 113, 120, 188] have been tested on instances from this test suite, the performance of MMMOEAs (presented in this thesis) is also analyzed on this test suite for comparison. The CEC 2019 test suite [112] consists of 22 box-constrained multi-modal multi-objective problems. As for MMMOEAs explores the decision space along with the objective space, true Pareto-optimal Sets (PSs) are illustrated with true PF in Table A.1 for this test suite.

A.5.1 MMF1 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1 - 2| \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - sin\left(6\pi\,|x_1 - 2| + \pi\right)\right)^2 \\
&\text{subjected to } 1 \le x_1 \le 3 \text{ and } -1 \le x_2 \le 1
\end{aligned} \tag{A.29}$$

Global PF of Eq. (A.29) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$. Global PS of Eq. (A.29) corresponds to $x_2 = sin\left(6\pi\,|x_1 - 2| + \pi\right)$ with $1 \le x_1 \le 3$.

A.5.2 MMF1_z problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1 - 2| \\
&\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - sin\left(2k_M\pi\,|x_1 - 2| + \pi\right)\right)^2, & \text{if } x_1 \in [1, 2) \\ 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - sin\left(2\pi\,|x_1 - 2| + \pi\right)\right)^2, & \text{if } x_1 \in [2, 3] \end{cases} \\
&\text{where } k_M > 0 \text{ controls the degree of deformation in the PS corresponding to } x_1 \in [1, 2) \\
&\text{subjected to } 1 \le x_1 \le 3 \text{ and } -1 \le x_2 \le 1
\end{aligned} \tag{A.30}$$

Global PF of Eq. (A.30) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$. Global PSs of Eq. (A.30) are at: (i) $x_2 = sin\left(2k_M\pi\,|x_1 - 2| + \pi\right)$ with $1 \le x_1 < 2$ and (ii) $x_2 = sin\left(2\pi\,|x_1 - 2| + \pi\right)$ with $2 \le x_1 \le 3$. In this thesis, $k_M = 3$ is considered as per [112].

A.5.3 MMF1_e problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\text{Minimize: } f_1(\mathbf{X}) = |x_1 - 2|$$

$$\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - sin\left(6\pi\left|x_1 - 2\right| + \pi\right)\right)^2, & \text{if } x_1 \in [1,2) \\ 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - a_M^{x_1} sin\left(6\pi\left|x_1 - 2\right| + \pi\right)\right)^2, & \text{if } x_1 \in [2,3] \end{cases}$$

where $a_M > 0$ and $a_M \neq 1$ controls the amplitude of the PS corresponding to $x_1 \in [2,3]$

subjected to $1 \leq x_1 \leq 3$ and $-a_M^3 \leq x_2 \leq a_M^3$ (A.31)

Global PF of Eq. (A.31) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. Global PSs of Eq. (A.31) are at: (i) $x_2 = sin\left(6\pi\left|x_1 - 2\right| + \pi\right)$ with $1 \leq x_1 < 2$ and (ii) $x_2 = a_M^{x_1} sin\left(2\pi\left|x_1 - 2\right| + \pi\right)$ with $2 \leq x_1 \leq 3$. In this thesis, $a_M = e$ is considered as per [112].

A.5.4 MMF2 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

Minimize: $f_1(\mathbf{X}) = x_1$

Minimize: $f_2(\mathbf{X})$

$$= \begin{cases} 1 - \sqrt{x_1} + 2\left(4\left(x_2 - \sqrt{x_2}\right)^2 - 2cos\left(\frac{20\pi(x_2 - \sqrt{x_2})}{\sqrt{2}}\right) + 2\right), & \text{if } x_2 \in [0,1] \\ 1 - \sqrt{x_1} + 2\left(4\left(x_2 - 1 - \sqrt{x_2}\right)^2 - 2cos\left(\frac{20\pi(x_2 - 1 - \sqrt{x_2})}{\sqrt{2}}\right) + 2\right), & \text{if } x_2 \in (1,2] \end{cases}$$

subjected to $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 2$ (A.32)

Global PF of Eq. (A.32) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. Global PSs of Eq. (A.32) are at: (i) $x_1 = x_2^2$ with $0 \leq x_2 \leq 1$ and (ii) $x_1 = (x_2 - 1)^2$ with $1 < x_2 \leq 2$.

A.5.5 MMF3 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

Minimize: $f_1(\mathbf{X}) = x_1$

$$
\begin{aligned}
&\text{Minimize: } f_2(\mathbf{X})\\
&= \begin{cases} 1-\sqrt{x_1}+2\left(4\left(x_2-\sqrt{x_2}\right)^2\right. \\ \left. -2cos\left(\frac{20\pi(x_2-\sqrt{x_2})}{\sqrt{2}}\right)+2\right), & \text{if } x_2\in[0,0.5], x_2\in(0.5,1)\ \&\ x_1\in(0.25,1] \\ 1-\sqrt{x_1}+2\left(4\left(x_2-0.5-\sqrt{x_2}\right)^2\right. \\ \left. -2cos\left(\frac{20\pi(x_2-0.5-\sqrt{x_2})}{\sqrt{2}}\right)+2\right), & \text{if } x_2\in[1,1.5], x_1\in[0,0.25)\ \&\ x_2\in(0.5,1) \end{cases}\\
&\text{subjected to } 0\le x_1\le 1 \text{ and } 0\le x_2\le 1.5 \qquad \text{(A.33)}
\end{aligned}
$$

Global PF of Eq. (A.33) corresponds to $f_2(\mathbf{X}) = 1-\sqrt{f_1(\mathbf{X})}$ with $0 \le f_1(\mathbf{X}) \le 1$. Global PSs of Eq. (A.33) are at: (i) $x_1 = x_2^2$ when $x_2 \in [0, 0.5]$, $x_2 \in (0.5, 1)$ & $x_1 \in (0.25, 1]$ and (ii) $x_1 = (x_2 - 0.5)^2$ when $x_2 \in [1, 1.5]$, $x_1 \in [0, 0.25)$ & $x_2 \in (0.5, 1)$.

A.5.6 MMF4 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1|\\
&\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} 1-x_1^2+2\left(x_2-sin\left(\pi\left|x_1\right|\right)\right)^2, & \text{if } x_2\in[0,1) \\ 1-x_1^2+2\left(x_2-1-sin\left(\pi\left|x_1\right|\right)\right)^2, & \text{if } x_2\in[1,2] \end{cases}\\
&\text{subjected to } -1\le x_1\le 1 \text{ and } 0\le x_2\le 2 \qquad \text{(A.34)}
\end{aligned}
$$

Global PF of Eq. (A.34) corresponds to $f_2(\mathbf{X}) = 1 - \{f_1(\mathbf{X})\}^2$ with $0 \le f_1(\mathbf{X}) \le 1$. Global PSs of Eq. (A.34) are at: (i) $x_2 = sin\left(\pi\left|x_1\right|\right)$ with $-1 \le x_1 \le 1$ and (ii) $x_2 = sin\left(\pi\left|x_1\right|\right) + 1$ with $-1 \le x_1 \le 1$.

A.5.7 MMF5 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1-2|\\
&\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} 1-\sqrt{|x_1-2|}+2\left(x_2-sin\left(6\pi\left|x_1-2\right|+\pi\right)\right)^2, & \text{if } x_2\in[-1,1] \\ 1-\sqrt{|x_1-2|}+2\left(x_2-2-sin\left(6\pi\left|x_1-2\right|+\pi\right)\right)^2, & \text{if } x_2\in(1,3] \end{cases}\\
&\text{subjected to } 1\le x_1\le 3 \text{ and } -1\le x_2\le 3 \qquad \text{(A.35)}
\end{aligned}
$$

Global PF of Eq. (A.35) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. Global PSs of Eq. (A.35) are at: (i) $x_2 = sin\left(6\pi\left|x_1 - 2\right| + \pi\right)$ with $1 \leq x_1 \leq 3$ and (ii) $x_2 = sin\left(6\pi\left|x_1 - 2\right| + \pi\right) + 2$ with $1 \leq x_1 \leq 3$.

A.5.8 MMF6 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1 - 2| \\
&\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - sin\left(6\pi\left|x_1 - 2\right| + \pi\right)\right)^2, & \text{if } x_2 \in [-1, 1] \\ 1 - \sqrt{|x_1 - 2|} + 2\left(x_2 - 1 - sin\left(6\pi\left|x_1 - 2\right| + \pi\right)\right)^2, & \text{if } x_2 \in (1, 3] \end{cases} \\
&\text{subjected to } 1 \leq x_1 \leq 3 \text{ and } -1 \leq x_2 \leq 3
\end{aligned} \tag{A.36}$$

Global PF of Eq. (A.36) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. Global PSs of Eq. (A.36) are at: (i) $x_2 = sin\left(6\pi\left|x_1 - 2\right| + \pi\right)$ with $1 \leq x_1 \leq 3$ and (ii) $x_2 = sin\left(6\pi\left|x_1 - 2\right| + \pi\right) + 1$ with $1 \leq x_1 \leq 3$.

A.5.9 MMF7 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = |x_1 - 2| \\
&\text{Minimize: } f_2(\mathbf{X}) = 1 - \sqrt{|x_1 - 2|} + \\
&\left[x_2 - \left\{0.3\left|x_1 - 2\right|^2 cos\left(24\pi\left|x_1 - 2\right| + 4\pi\right) + 0.6\left|x_1 - 2\right|\right\} sin\left(6\pi\left|x_1 - 2\right| + \pi\right)\right]^2 \\
&\text{subjected to } 1 \leq x_1 \leq 3 \text{ and } -1 \leq x_2 \leq 1.
\end{aligned} \tag{A.37}$$

Global PF of Eq. (A.37) corresponds to $f_2(\mathbf{X}) = 1 - \sqrt{f_1(\mathbf{X})}$ with $0 \leq f_1(\mathbf{X}) \leq 1$. Global PS of Eq. (A.37) corresponds to Eq. (A.38) with $1 \leq x_1 \leq 3$.

$$x_2 = \left\{0.3\left|x_1 - 2\right|^2 cos\left(24\pi\left|x_1 - 2\right| + 4\pi\right) + 0.6\left|x_1 - 2\right|\right\} sin\left(6\pi\left|x_1 - 2\right| + \pi\right) \tag{A.38}$$

A.5.10 MMF8 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = sin\,|x_1| \\
&\text{Minimize: } f_2(\mathbf{X}) = \begin{cases} \sqrt{1-(sin\,|x_1|)^2} + 2\left(x_2 - sin\,|x_1| - |x_1|\right)^2, & \text{if } x_2 \in [0,4] \\ \sqrt{1-(sin\,|x_1|)^2} + 2\left(x_2 - 4 - sin\,|x_1| - |x_1|\right)^2, & \text{if } x_2 \in (4,9] \end{cases} \\
&\text{subjected to } -\pi \le x_1 \le \pi \text{ and } 0 \le x_2 \le 9
\end{aligned} \tag{A.39}$$

Global PF of Eq. (A.39) corresponds to $\{f_1(\mathbf{X})\}^2 + \{f_2(\mathbf{X})\}^2 = 1$ with $0 \le f_i(\mathbf{X}) \le 1$ for $i = 1, 2$. Global PSs of Eq. (A.39) are at: (i) $x_2 = sin\,|x_1| + |x_1|$ with $-\pi \le x_1 \le \pi$ and (ii) $x_2 = sin\,|x_1| + |x_1| + 4$ with $-\pi \le x_1 \le \pi$.

A.5.11 MMF9 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = \frac{h(x_2)}{x_1} \\
&\text{where } h(x_i) = 2 - sin^6\left(k_M \pi x_i\right) \text{ with } k_M \text{ denoting the number of PSs} \\
&\text{subjected to } 0.1 \le x_i \le 1.1, \text{ for } i = 1, 2
\end{aligned} \tag{A.40}$$

Global PF of Eq. (A.40) corresponds to $f_2\left(\mathbf{X}\right) = h\left(\frac{1}{2k_M}\right)/f_1\left(\mathbf{X}\right)$ with $0.1 \le f_1(\mathbf{X}) \le 1.1$. The i^{th} global PS of Eq. (A.40) is at $x_2 = \frac{1}{2k_M} + \frac{i-1}{k_M}$, $x_1 \in [0.1, 1.1]$ for $i = 1, \cdots, k_M$. In this thesis, $k_M = 2$ is considered as per [112].

A.5.12 MMF10 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = \frac{h(x_2)}{x_1} \\
&\text{where } h(x_i) = 2 - \exp\left[-\left(\frac{x_i - 0.2}{0.004}\right)^2\right] - 0.8\exp\left[-\left(\frac{x_i - 0.6}{0.4}\right)^2\right]
\end{aligned}$$

$$\text{subjected to } 0.1 \leq x_i \leq 1.1, \text{ for } i = 1, 2 \tag{A.41}$$

Global PF of Eq. (A.41) corresponds to $f_2(\mathbf{X}) = h(0.2)/f_1(\mathbf{X})$ with $0.1 \leq f_1(\mathbf{X}) \leq 1.1$ and its local PF corresponds to $f_2(\mathbf{X}) = h(0.6)/f_1(\mathbf{X})$ with $0.1 \leq f_1(\mathbf{X}) \leq 1.1$. Global PS of Eq. (A.41) is at $x_2 = 0.2$, $x_1 \in [0.1, 1.1]$ and its local PS is at $x_2 = 0.6$, $x_1 \in [0.1, 1.1]$.

A.5.13 MMF11 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = \frac{h(x_2)}{x_1} \\
&\text{where } h(x_i) = 2 - \exp\left[-2\log(2)\left(\frac{x_i - 0.1}{0.8}\right)^2\right] sin^6\left(k_M \pi x_i\right) \\
&\text{with } k_M \text{ denoting the number of PSs} \\
&\text{subjected to } 0.1 \leq x_j \leq 1.1, \text{ for } j = 1, 2
\end{aligned} \tag{A.42}$$

Global PF of Eq. (A.42) corresponds to $f_2(\mathbf{X}) = h\left(\frac{1}{2k_M}\right)/f_1(\mathbf{X})$ with $0.1 \leq f_1(\mathbf{X}) \leq 1.1$ and its i^{th} local PF corresponds to $f_2(\mathbf{X}) = h\left(\frac{1}{2k_M} + \frac{i-1}{k_M}\right)/f_1(\mathbf{X})$ with $0.1 \leq f_1(\mathbf{X}) \leq 1.1$ and $i = 2, \cdots, k_M$. Global PS of Eq. (A.42) is at $x_2 = \frac{1}{2k_M}$, $x_1 \in [0.1, 1.1]$ and its i^{th} local PS is at $x_2 = \frac{1}{2k_M} + \frac{i-1}{k_M}$, $x_1 \in [0.1, 1.1]$ for $i = 2, \cdots, k_M$. In this thesis, $k_M = 2$ is considered as per [112].

A.5.14 MMF12 problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = h_1(x_2) h_2(f_1, h_1) \\
&\text{where } h_1(x_i) = 2 - \exp\left[-2\log(2)\left(\frac{x_i - 0.1}{0.8}\right)^2\right] sin^6\left(k_M \pi x_i\right) \\
&\text{and } h_2(f_1, h_1) = 1 - \left(\frac{f_1}{h_1}\right)^2 - \frac{f_1}{h_1} sin\left(2\pi a_M f_1\right) \\
&\text{with } k_M \text{ and } a_M \text{ denoting the number of PSs and discontinuous pieces in PF(PS)}
\end{aligned}$$

$$\text{subjected to } 0 \leq x_i \leq 1, \text{ for } i = 1, 2 \tag{A.43}$$

Global PF of Eq. (A.43) has discontinuous pieces at $f_2(\mathbf{X}) = h_1^{g\star} h_2(f_1, h_1^{g\star})$ and its local PF has discontinuous pieces at $f_2(\mathbf{X}) = h_1^{l\star} h_2(f_1, h_1^{l\star})$ where $h_1^{g\star}$ and $h_1^{l\star}$ are the global and local optima of $h_1(.)$, respectively. The ranges of discontinuous pieces depend on the minima of $f_2(\mathbf{X}) = h_1^{g\star} h_2(f_1, h_1^{g\star})$. Global PS of Eq. (A.43) is at $x_2 = \frac{1}{2k_M}$, $x_1 \in [0,1]$ and its i^{th} local PS is at $x_2 = \frac{1}{2k_M} + \frac{i-1}{k_M}$, $x_1 \in [0,1]$ for $i = 2, \cdots, k_M$. In this thesis, $k_M = 2$ and $a_M = 4$ are considered as per [112].

A.5.15 MMF13 problem

Defined on 3-dimensional decision vector, this 2-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = x_1 \\
&\text{Minimize: } f_2(\mathbf{X}) = \frac{h_1(u)}{x_1} \\
&\text{where } h_1(u) = 2 - \exp\left[-2\log(2)\left(\frac{u-0.1}{0.8}\right)^2\right] sin^6\left(k_M \pi u\right) \\
&\text{with } k_M \text{ denoting the number of PSs and } u = x_2 + \sqrt{x_3} \\
&\text{subjected to } 0.1 \leq x_i \leq 1.1, \text{ for } i = 1, 2, 3
\end{aligned} \tag{A.44}$$

Global PF of Eq. (A.44) corresponds to Eq. (A.45) with $i = 1$ and its i^{th} local PF corresponds to Eq. (A.45) with $i = 2, \cdots, k_M$. Global PS of Eq. (A.44) is at $x_2 + \sqrt{x_3} = \frac{1}{2k_M}$, $x_1 \in [0.1, 1.1]$ and its i^{th} local PS is at $x_2 + \sqrt{x_3} = \frac{1}{2k_M} + \frac{i-1}{k_M}$, $x_1 \in [0.1, 1.1]$ for $i = 2, \cdots, k_M$. In this thesis, $k_M = 2$ is considered as per [112].

$$f_2(\mathbf{X}) = \frac{2 - \exp\left[-2\log(2)\left(\frac{\left(\frac{1}{2k_M} + \frac{i-1}{k_M}\right) - 0.1}{0.8}\right)^2\right] sin^6\left(k_M \pi \left(\frac{1}{2k_M} + \frac{i-1}{k_M}\right)\right)}{f_1(\mathbf{X})} \tag{A.45}$$

A.5.16 MMF14 problem

Defined on N-dimensional decision vector, this M-objective problem is given by:

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots cos\left(\frac{\pi x_{M-2}}{2}\right) cos\left(\frac{\pi x_{M-1}}{2}\right)\left(1 + h(\mathbf{X}_M)\right) \\
&\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots cos\left(\frac{\pi x_{M-2}}{2}\right) sin\left(\frac{\pi x_{M-1}}{2}\right)\left(1 + h(\mathbf{X}_M)\right)
\end{aligned}$$

$$\text{Minimize: } f_3(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots sin\left(\frac{\pi x_{M-2}}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\vdots$$

$$\text{Minimize: } f_{M-1}(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) sin\left(\frac{\pi x_2}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{Minimize: } f_M(\mathbf{X}) = sin\left(\frac{\pi x_1}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{where } h(\mathbf{X}_M) = h\left(x_M, x_{M+1}, \cdots, x_{M-1+a_M}\right) = 2 - sin^2\left(k_M \pi x_{M-1+a_M}\right)$$

$$\text{with } a_M = N - (M-1) \text{ and } k_M \text{ denoting the number of PSs}$$

$$\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, 2, \cdots, N \tag{A.46}$$

Considering the global optima of $h(.)$ as $h^{\star}$, the global PF of Eq. (A.46) is located at $\sum_{j=1}^{M}\{f_j(\mathbf{X})\}^2 = (1 + h^{\star})^2$. The i^{th} global PS of Eq. (A.46) is located at $x_N = \frac{1}{2k_M} + \frac{i-1}{k_M}$, $x_j \in [0,1]$ with $i = 1, \cdots, k_M$ and $j = 1, \cdots, N-1$. In this thesis, $a_M = 1$ and $k_M = 2$ are considered as per [112].

A.5.17 MMF14_a problem

Defined on N-dimensional decision vector, this M-objective problem is given by:

$$\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots cos\left(\frac{\pi x_{M-2}}{2}\right) cos\left(\frac{\pi x_{M-1}}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots cos\left(\frac{\pi x_{M-2}}{2}\right) sin\left(\frac{\pi x_{M-1}}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{Minimize: } f_3(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right) \cdots sin\left(\frac{\pi x_{M-2}}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\vdots$$

$$\text{Minimize: } f_{M-1}(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) sin\left(\frac{\pi x_2}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{Minimize: } f_M(\mathbf{X}) = sin\left(\frac{\pi x_1}{2}\right)(1 + h(\mathbf{X}_M))$$

$$\text{where } h(\mathbf{X}_M) = h\left(x_M, x_{M+1}, \cdots, x_{M-1+a_M}\right)$$

$$= 2 - sin^2\left(k_M \pi\left(x_{M-1+a_M} - 0.5 sin\left(\pi x_{M-2+a_M}\right) + \frac{1}{2k_M}\right)\right)$$

$$\text{with } a_M = N - (M-1) \text{ and } k_M \text{ denoting the number of PSs}$$

$$\text{subjected to } 0 \leq x_j \leq 1, \text{ for } j = 1, 2, \cdots, N \tag{A.47}$$

Considering the global optima of $h(.)$ as $h^{\star}$, the global PF of Eq. (A.47) is located at $\sum_{j=1}^{M}\{f_j(\mathbf{X})\}^2 = (1 + h^{\star})^2$. The i^{th} global PS of Eq. (A.47) is located at $x_N =$

$0.5sin\left(\pi x_{N-1}\right)+\frac{i-1}{k_M}$, $x_j \in [0,1]$ with $i=1,\cdots,k_M$ and $j=1,\cdots,N-1$. In this thesis, $a_M=1$ and $k_M=2$ are considered as per [112].

A.5.18 MMF15 problem

Defined on N-dimensional decision vector, this M-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots cos\left(\frac{\pi x_{M-2}}{2}\right) cos\left(\frac{\pi x_{M-1}}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots cos\left(\frac{\pi x_{M-2}}{2}\right) sin\left(\frac{\pi x_{M-1}}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{Minimize: } f_3(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots sin\left(\frac{\pi x_{M-2}}{2}\right)(1+h(\mathbf{X}_M))\\
&\vdots\\
&\text{Minimize: } f_{M-1}(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) sin\left(\frac{\pi x_2}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{Minimize: } f_M(\mathbf{X}) = sin\left(\frac{\pi x_1}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{where } h(\mathbf{X}_M) = h\left(x_M, x_{M+1},\cdots,x_{M-1+a_M}\right)\\
&= 2-\exp\left[-2\log(2)\left(\frac{x_{M-1+a_M}-0.1}{0.8}\right)^2\right] sin^2\left(k_M\pi x_{M-1+a_M}\right)\\
&\text{with } a_M = N-(M-1) \text{ and } k_M \text{ denoting the number of PSs}\\
&\text{subjected to } 0 \le x_j \le 1, \text{ for } j=1,2,\cdots,N
\end{aligned}
\tag{A.48}
$$

Considering the global and the i^{th} local optima of $h(.)$ as $h^\star$ and $h_i^\star$, the global and the i^{th} local PF of Eq. (A.48) are located at $\sum_{j=1}^{M}\{f_j\left(\mathbf{X}\right)\}^2=(1+h^\star)^2$ and $\sum_{j=1}^{M}\{f_j\left(\mathbf{X}\right)\}^2=(1+h_i^\star)^2$, respectively. The i^{th} PS of Eq. (A.48) is located at $x_N=\frac{1}{2k_M}+\frac{i-1}{k_M}$, $x_j\in[0,1]$ with $j=1,\cdots,N-1$, $i=1$ for the global PS and $i=2,\cdots,k_M$ for the local PSs. In this thesis, $a_M=1$ and $k_M=2$ are considered as per [112].

A.5.19 MMF15_a problem

Defined on N-dimensional decision vector, this M-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots cos\left(\frac{\pi x_{M-2}}{2}\right) cos\left(\frac{\pi x_{M-1}}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{Minimize: } f_2(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots cos\left(\frac{\pi x_{M-2}}{2}\right) sin\left(\frac{\pi x_{M-1}}{2}\right)(1+h(\mathbf{X}_M))\\
&\text{Minimize: } f_3(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) cos\left(\frac{\pi x_2}{2}\right)\cdots sin\left(\frac{\pi x_{M-2}}{2}\right)(1+h(\mathbf{X}_M))\\
&\vdots
\end{aligned}
$$

$$\text{Minimize: } f_{M-1}(\mathbf{X}) = cos\left(\frac{\pi x_1}{2}\right) sin\left(\frac{\pi x_2}{2}\right)(1 + h(\mathbf{X}_M))$$
$$\text{Minimize: } f_M(\mathbf{X}) = sin\left(\frac{\pi x_1}{2}\right)(1 + h(\mathbf{X}_M))$$
$$\text{where } h(\mathbf{X}_M) = h\left(x_M, x_{M+1}, \cdots, x_{M-1+a_M}\right)$$
$$= 2 - \exp\left[-2\log(2)\left(\frac{u-0.1}{0.8}\right)^2\right] sin^2\left(k_M \pi u\right)$$
$$\text{with } u = \left(x_{M-1+a_M} - 0.5 sin\left(\pi x_{M-2+a_M}\right) + \frac{1}{2k_M}\right),$$
$$a_M = N - (M-1) \text{ and } k_M \text{ denoting the number of PSs}$$
$$\text{subjected to } 0 \le x_j \le 1, \text{ for } j = 1, 2, \cdots, N \tag{A.49}$$

Considering the global and the i^{th} local optima of $h(.)$ as $h^\star$ and $h_i^\star$, the global and the i^{th} local PF of Eq. (A.49) are located at $\sum_{j=1}^{M}\{f_j(\mathbf{X})\}^2 = (1+h^\star)^2$ and $\sum_{j=1}^{M}\{f_j(\mathbf{X})\}^2 = (1+h_i^\star)^2$, respectively. The i^{th} PS of Eq. (A.49) is located at $x_N = 0.5 sin\left(\pi x_{N-1}\right) + \frac{i-1}{k_M}$, $x_j \in [0,1]$ with $j = 1, \cdots, N-1$, $i = 1$ for the global PS and $i = 2, \cdots, k_M$ for the local PSs. In this thesis, $a_M = 1$ and $k_M = 2$ are considered as per [112].

A.5.20 SYM-PART simple problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$\text{Minimize: } f_1(\mathbf{X}) = (J_1 + a_M)^2 + J_2^2$$
$$\text{Minimize: } f_2(\mathbf{X}) = (J_1 - a_M)^2 + J_2^2$$
$$\text{where } J_1 = x_1 - \{sgn\left(u_1\right) min\left(\left|u_1\right|, 1\right)\}\left(2a_M + c_M\right)$$
$$\text{and } J_2 = x_2 - \{sgn\left(u_2\right) min\left(\left|u_2\right|, 1\right)\} b_M$$
$$\text{with } u_1 = sgn(x_1)\left\lceil \frac{|x_1| - \left(a_M + \frac{c_M}{2}\right)}{2a_M + c_M} \right\rceil \text{ and } u_2 = sgn(x_2)\left\lceil \frac{|x_2| - \frac{b_M}{2}}{b_M} \right\rceil$$
$$\text{subjected to } -20 \le x_j \le 20, \text{ for } j = 1, 2 \tag{A.50}$$

Global PF of Eq. (A.50) corresponds to $f_1(\mathbf{X}) = 4a_M^2 u_3^2$ and $f_2(\mathbf{X}) = 4a_M^2(1-u_3)^2$ with $u_3 \in [0,1]$. Global PSs of Eq. (A.50) are at $x_1 = J_1$, $x_2 = 0$. In this thesis, $a_M = 1$, $b_M = 10$ and $c_M = 8$ are considered as per [112].

A.5.21 SYM-PART rotated problem

Defined on 2-dimensional decision vector, this 2-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = (J_1 + a_M)^2 + J_2^2 \\
&\text{Minimize: } f_2(\mathbf{X}) = (J_1 - a_M)^2 + J_2^2 \\
&\text{where } J_1 = x_1 - \{sgn\,(u_1)\,min\,(|u_1|\,,1)\}\,(2a_M + c_M) \\
&\text{with } u_1 = sgn(u_2)\left\lceil \frac{|u_2| - \left(a_M + \frac{c_M}{2}\right)}{2a_M + c_M} \right\rceil, u_2 = (cos\ \omega)x_1 - (sin\ \omega)x_2 \\
&\text{and } J_2 = x_2 - \{sgn\,(u_3)\,min\,(|u_3|\,,1)\}\,b_M \\
&\text{with } u_3 = sgn(u_4)\left\lceil \frac{|u_4| - \frac{b_M}{2}}{b_M} \right\rceil, u_4 = (sin\ \omega)x_1 + (cos\ \omega)x_2 \\
&\text{subjected to } -20 \le x_j \le 20, \text{ for } j = 1, 2
\end{aligned}
\tag{A.51}
$$

Global PF of Eq. (A.51) corresponds to $f_1(\mathbf{X}) = 4a_M^2 u_5^2$ and $f_2(\mathbf{X}) = 4a_M^2(1 - u_5)^2$ with $u_5 \in [0, 1]$. Global PSs of Eq. (A.51) are at $x_1 = J_1$, $x_2 = 0$. In this thesis, $a_M = 1$, $b_M = 10$, $c_M = 8$ and $\omega = \pi/4$ are considered as per [112].

A.5.22 Omni-test problem

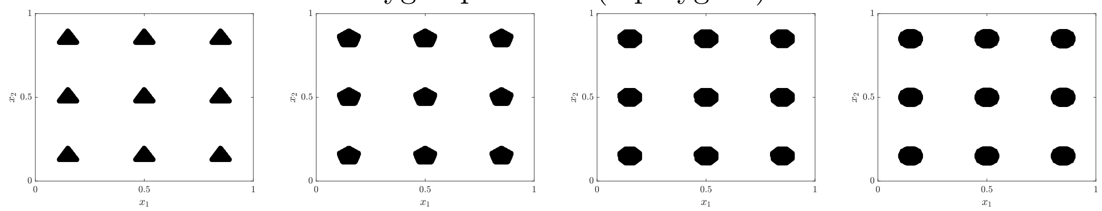
Defined on N-dimensional decision vector, this 2-objective problem is given by:

$$
\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = \sum_{j=1}^{N} sin\,(\pi x_j) \\
&\text{Minimize: } f_2(\mathbf{X}) = \sum_{j=1}^{N} cos\,(\pi x_j) \\
&\text{subjected to } 0 \le x_j \le 6, \text{ for } j = 1, \cdots, N
\end{aligned}
\tag{A.52}
$$

Global PF of Eq. (A.52) corresponds to $f_2(\mathbf{X}) = -\sqrt{N^2 - \{f_1(\mathbf{X})\}^2}$ with $-N \le f_1(\mathbf{X}) \le 0$. Global PSs of Eq. (A.52) are at $x_j \in [2k_M + 1, 2k_M + 3/2]$ with $j = 1, \cdots, N$ and k_M takes integer values. In this thesis, $N = 3$ is considered as per [112].

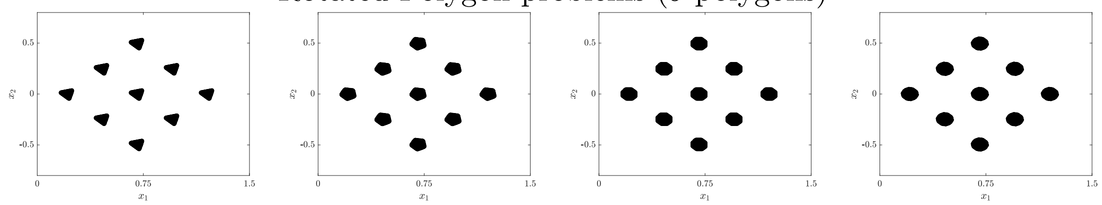
Table A.2: Cartesian coordinate plots of the 2-dimensional PSs of M-objective polygon and rotated polygon problems.

Polygon problems (9 polygons)



$M = 3$ | $M = 5$ | $M = 8$ | $M = 10$

Rotated Polygon problems (9 polygons)



$M = 3$ | $M = 5$ | $M = 8$ | $M = 10$

A.6 Multi-Modal Many-objective Polygon Problems [76]

The performance of MMMOEAs are also analyzed using M-objective polygon and rotated (by 45 degrees) polygon problems [76, 170], as these problems have variable number of objectives. For these problems, the number of objectives (M) is equal to the number of vertices of the polygons and i^{th} objective to be minimized is given by the Euclidean distance of a solution to its nearest i^{th} vertex over any of the given number of polygons. Hence, a 3-objective polygon problem searches for triangles, a 5-objective polygon problem searches for pentagons and so on, as shown in Table A.2. For example, a 3-objective 9-polygon problem is defined by Eq. (A.53) where $\mathbf{V}_{1,i}$, $\mathbf{V}_{2,i}$ and $\mathbf{V}_{3,i}$ are the three-vertices of the i^{th} polygon (as shown in Fig. A.5) and the function $D_E(.)$ evaluates the Euclidean distance between two vectors.

$$\begin{aligned}
&\text{Minimize: } f_1(\mathbf{X}) = min\{D_E(\mathbf{X}, \mathbf{V}_{1,1}), D_E(\mathbf{X}, \mathbf{V}_{1,2}), \cdots, D_E(\mathbf{X}, \mathbf{V}_{1,9})\}\\
&\text{Minimize: } f_2(\mathbf{X}) = min\{D_E(\mathbf{X}, \mathbf{V}_{2,1}), D_E(\mathbf{X}, \mathbf{V}_{2,2}), \cdots, D_E(\mathbf{X}, \mathbf{V}_{2,9})\}\\
&\text{Minimize: } f_3(\mathbf{X}) = min\{D_E(\mathbf{X}, \mathbf{V}_{3,1}), D_E(\mathbf{X}, \mathbf{V}_{3,2}), \cdots, D_E(\mathbf{X}, \mathbf{V}_{3,9})\}
\end{aligned} \tag{A.53}$$

The specifications of the eight problem instances, considered in this thesis, are as follows:

- Dimension of decision space: $N = 2$
- Bounds of decision space ($\mathcal{D}$): $\mathbf{X}^L = [-1, -1]$ and $\mathbf{X}^U = [2, 2]$

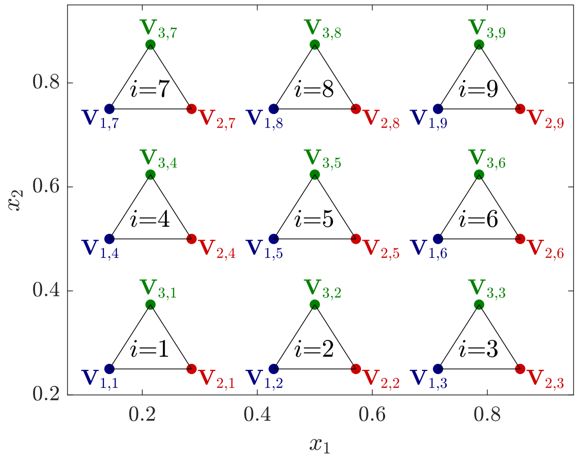


Figure A.5: Decision space of a 3-objective 9-polygon problem.

- Dimension of objective space: $M \in \{3, 5, 8, 10\}$
- Number of Subsets in PS: #Sets = number of polygons = 9
- Vertices of the polygons are obtained from `https://sites.google.com/view/nimmopt/`.

Appendix B

Visualizing an M-objective Pareto-Front using Polar Plots

B.1 Steps for Visualization

The steps to visualize an M-dimensional Pareto-Front (PF: $\mathcal{A}_{\mathbf{F}}$) using the polar coordinate plot [68] are as follows:

1. The M-dimensional objective space is partitioned into n_{dir} sub-spaces using Das and Dennis' approach [40] of reference vectors generation (Section 3.2.1, Python implementation in `http://worksupplements.droppages.com/refvecgen`). For example in Fig. B.1, the set of reference vectors $\mathcal{W} = [\mathbf{W}_1, \cdots, \mathbf{W}_{n_{dir}}]$ is formed with $p_1 = 4$ for $M = 2$ and with $p_1 = 2$ (boundary layer) and $p_2 = 1$ (inside layer) for $M = 3$.

2. In the polar plot, n_{dir} uniformly spread directions are chosen to correspond to the n_{dir} sub-spaces ($\mathcal{S}_1$ to $\mathcal{S}_{n_{dir}}$) from the objective space. However, the correspondence of a sub-space with a direction is randomly fixed. For example, both the transformations of the objective space to the polar coordinates are equivalent in Fig. B.1. As a thumb-rule, the i^{th} sub-space ($\mathcal{S}_i$) is assumed to be transformed into the direction at an angle (θ_i^{rad}) as follows:

$$\theta_i^{rad} = \frac{2\pi\,(i-1)}{n_{dir}}. \tag{B.1}$$

3. The transformation of an objective vector $\mathbf{F}$ from the objective space to the polar coordinates is also dictated by the shape of the PF.

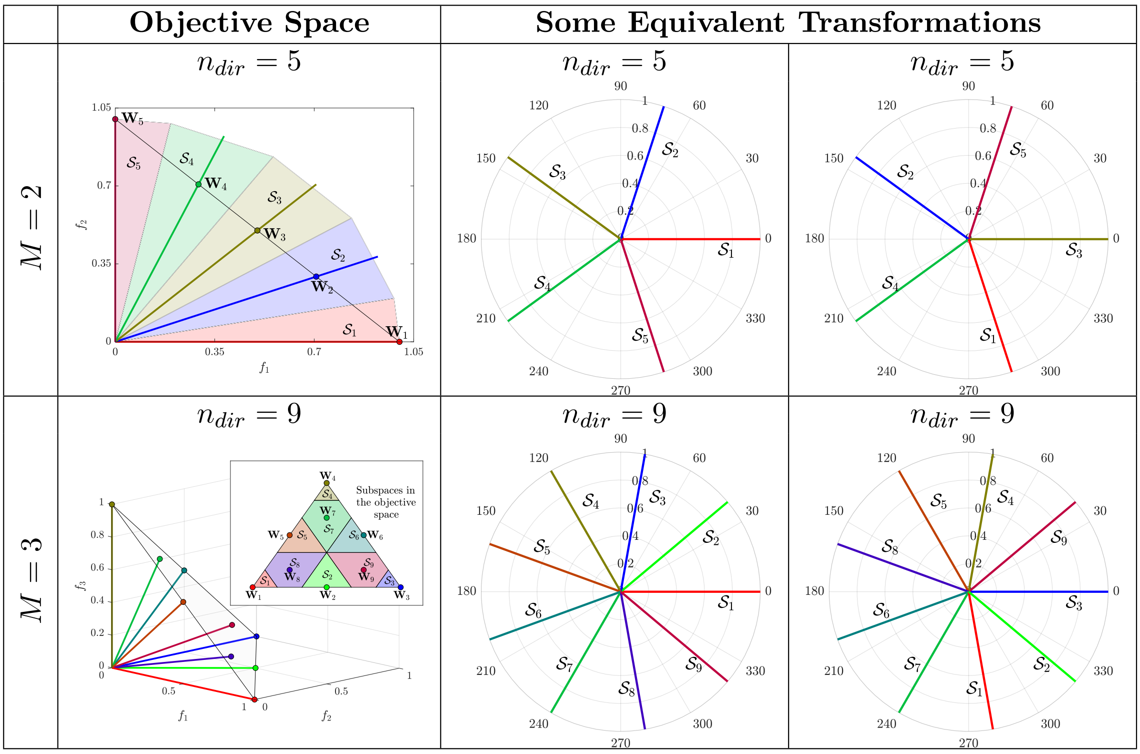


Figure B.1: Multiple equivalent transformations to map sub-spaces from the objective space into the polar coordinates.

- When the PF has a concave shape (Eq. (4.8) with $\delta_c > 1$), the radius $(\rho^{rad}(\mathbf{F}))$ of the objective vector $\mathbf{F}$ in the polar plot satisfies the following equation:

$$\rho^{rad}(\mathbf{F}) = \sqrt{\sum_{i=1}^{M} \left(f_i - f_i^{ide}\right)^2}. \tag{B.2}$$

- When the PF has a convex shape (Eq. (4.8) with $\delta_c < 1$), the radius $(\rho^{rad}(\mathbf{F}))$ of the objective vector $\mathbf{F}$ in the polar plot satisfies the following equation:

$$\rho^{rad}(\mathbf{F}) = \sqrt{\sum_{i=1}^{M} \left\{\rho^{rad}(\mathbf{F}) - \left(f_i - f_i^{ide}\right)\right\}^2}. \tag{B.3}$$

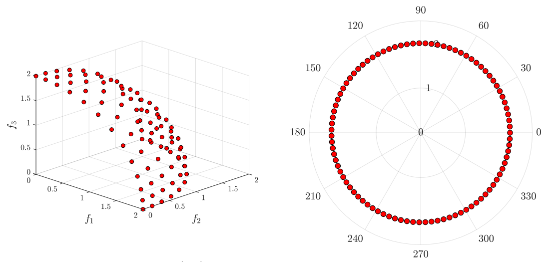
- When the PF has a linear shape (Eq. (4.8) with $\delta_c = 1$), the radius $(\rho^{rad}(\mathbf{F}))$ of the objective vector $\mathbf{F}$ in the polar plot satisfies the following equation:

$$\rho^{rad}(\mathbf{F}) = \sum_{i=1}^{M} \left(f_i - f_i^{ide}\right). \tag{B.4}$$

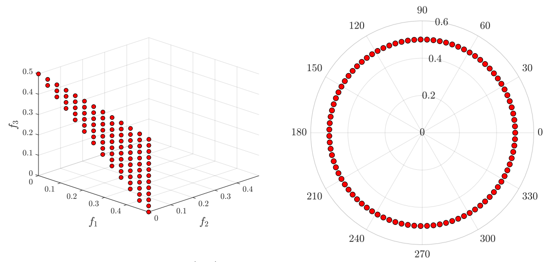
- It can be seen that for any regular PF, $\rho^{rad}(\mathbf{F})$ is constant $\forall \mathbf{F} \in \mathcal{A}_{\mathbf{F}}$. Ideally, for an irregular PF, different parts are locally regular and hence, a mixture of

$\rho^{rad}(\mathbf{F})$ using Eqs. (B.2) to (B.4) should be used. However, to keep things simple, in this thesis, $\rho^{rad}(\mathbf{F})$ is obtained by Eq. (B.2) for irregular PF.

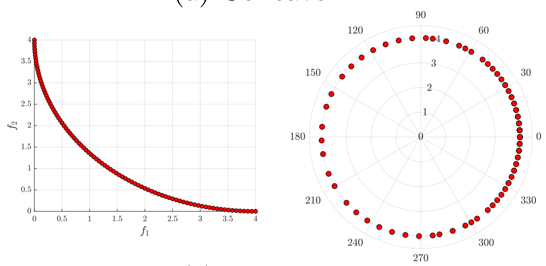
4. Assuming an objective vector $\mathbf{F}$ is associated with sub-space $\mathcal{S}_i$ (using Eq. (3.2)), its polar coordinates are computed as $\left(\rho^{rad}(\mathbf{F}), \theta_i^{rad}\right)$. For illustrating the mapping, the Cartesian-coordinate plots and the corresponding polar coordinate plots are presented in Fig. B.2a for a 3-objective concave PF (e.g., MMF14), in Fig. B.2b for a 3-objective linear PF (e.g., DTLZ1), in Fig. B.2c for a 2-objective convex PF (e.g., SYM-PART simple), and in Fig. B.2d for a 3-objective irregular PF (e.g., DTLZ7).



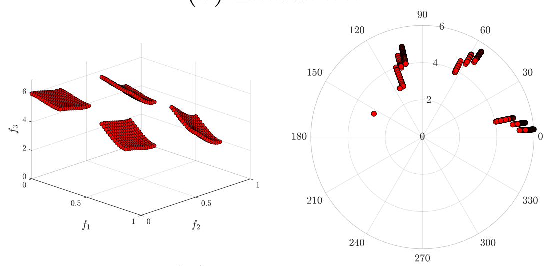
(a) Concave PF



(b) Linear PF



(c) Convex PF



(d) Irregular PF

Figure B.2: Objective space mapping between the Cartesian coordinate plots and the polar coordinate plots for different shapes of the PF with $n_{dir} = 91$.

B.2 Knowledge Retained and Lost by Polar Coordinate Plots

The following insights can be obtained from the polar coordinate plot:

- *Convergence*: For a regular PF, $\rho^{rad}(\mathbf{F}) = \rho^{rad}_{\mathcal{A}_\mathbf{F}}$ is constant $\forall \mathbf{F} \in \mathcal{A}_\mathbf{F}$. For two approximations of the PF ($\mathcal{A}_{\mathbf{F},1}$ and $\mathcal{A}_{\mathbf{F},2}$), if $\rho^{rad}_{\mathcal{A}_{\mathbf{F},1}} < \rho^{rad}_{\mathcal{A}_{\mathbf{F},2}}$, then $\mathcal{A}_{\mathbf{F},1}$ has better convergence than $\mathcal{A}_{\mathbf{F},2}$.

- *Shape*: For a regular PF, as $\rho^{rad}(\mathbf{F})$ will be constant using one of the rules (Eqs. (B.2) to (B.4)), the shape of the PF can also be understood using the rule yielding a constant $\rho^{rad}(\mathbf{F})$.

- *Diversity*: The diversity information can be obtained from the number of solutions associated with the sub-spaces. Thus, a well-diverse PF will have an equal number of solutions on each of the directions in the polar coordinate plot. Although the polar plot is uniformly spread for both concave PF and linear PF, it is not so for the convex PF in Fig. B.2c. It should be noted that this ambiguity is not from an information loss due to the transformation. Rather, it is due to the convexity of the PF, which associates more solutions to the sub-spaces near the objective axes.

- *Scalability*: The approach is applicable for any number of objectives and can efficiently reflect a large number of solutions. Another important aspect of this polar plot visualization is the easy comparison of more than one PF using the same plot.

Due to the advantages mentioned above, the polar plot visualization technique is used in this thesis. However, it suffers from the following disadvantages:

- For a proper representation of an irregular PF, obtaining $\rho^{rad}(\mathbf{F})$ locally is cumbersome. The polar coordinate plot is easier to obtain for the regular PF.

- As shown in Fig. B.1, there can be multiple possible transformations due to the random ordering of the sub-spaces. Hence, the spatial relation between the sub-spaces is not preserved.

- The solution distribution within each sub-space cannot be observed through this visualization method as all the objective vectors within a sub-space are plotted along the same direction in the polar coordinate plot.

- If the niche count is not uniform across all the sub-spaces, it cannot be ascertained which particular sub-space has poor diversity. Hence, it becomes difficult to understand the working of the algorithm and the characteristics of the problem.

Bibliography

[1] A. A. Alyafi, V.-B. Nguyen, Y. Laurillau, P. Reignier, S. Ploix, G. Calvary, J. Coutaz, M. Pal, and J.-P. Guilbaud. From usable to incentive-building energy management systems. *Modeling and Using Context*, 2(Issue 1):1–30, 2018.

[2] A. A. Alyafi, M. Pal, S. Ploix, P. Reignier, and S. Bandyopadhyay. Differential explanations for energy management in buildings. In *2017 Computing Conference*, pages 507–516, July 2017.

[3] M. Amayri, A. Arora, S. Ploix, S. Bandhyopadyay, Q.-D. Ngo, and V. R. Badarla. Estimating occupancy in heterogeneous sensor environment. *Energy and Buildings*, 129:46–58, 2016.

[4] M. Amayri, S. Ploix, and S. Bandyopadhyay. Estimating occupancy in an office setting. In *Sustainable Human–Building Ecosystems*, pages 72–80, 2015.

[5] L. M. Antonio and C. A. C. Coello. Indicator-based cooperative coevolution for multi-objective optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 991–998, July 2016.

[6] E. Asadi, M. G. Da Silva, C. H. Antunes, and L. Dias. Multi-objective optimization for building retrofit strategies: A model and an application. *Energy and Buildings*, 44:81–87, 2012.

[7] E. Asadi, M. G. da Silva, C. H. Antunes, L. Dias, and L. Glicksman. Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application. *Energy and Buildings*, 81:444–456, 2014.

[8] J. Bader and E. Zitzler. A hypervolume-based optimizer for high-dimensional objective spaces. In D. Jones, M. Tamiz, and J. Ries, editors, *New Developments in*

Multiple Objective and Goal Programming, pages 35–54, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[9] J. Bader and E. Zitzler. HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary computation*, 19(1):45–76, 2011.

[10] S. Bandyopadhyay and A. Mukherjee. An algorithm for many-objective optimization with reduced objective computations: A study in differential evolution. *IEEE Transactions on Evolutionary Computation*, 19(3):400–413, 2015.

[11] S. Bandyopadhyay, S. K. Pal, and B. Aruna. Multiobjective GAs, quantitative indices, and pattern classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(5):2088–2099, 2004.

[12] N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669, 2007.

[13] J. Blank, K. Deb, Y. Dhebar, S. Bandaru, and H. Seada. Generating well-spaced points on a unit simplex for evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, pages 1–12, 2020.

[14] P. A. N. Bosman and D. Thierens. The balance between proximity and diversity in multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 7(2):174–188, 2003.

[15] J. Brest, S. Greiner, B. Boskovic, M. Mernik, and V. Zumer. Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. *IEEE Transactions on Evolutionary Computation*, 10(6):646–657, Dec 2006.

[16] D. Brockhoff, T. Wagner, and H. Trautmann. On the properties of the R2 indicator. In *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pages 465–472. ACM, 2012.

[17] D. Brockhoff and E. Zitzler. Are all objectives necessary? on dimensionality reduction in evolutionary multiobjective optimization. In *Parallel Problem Solving from Nature-PPSN IX*, pages 533–542. Springer, 2006.

[18] D. Brockhoff and E. Zitzler. Objective reduction in evolutionary multiobjective optimization: theory and applications. *Evolutionary Computation*, 17(2):135–166, 2009.

[19] P. Buser. Über eine ungleichung von Cheeger (On an inequality of Cheeger). *Mathematische Zeitschrift (in German)*, 158(3):245–252, Oct 1978.

[20] L. Castillo, R. Enríquez, M. Jiménez, and M. Heras. Dynamic integrated method based on regression and averages, applied to estimate the thermal parameters of a room in an occupied office building in madrid. *Energy and Buildings*, 81:337–362, 2014.

[21] K. P. Chan and T. Ray. An evolutionary algorithm to maintain diversity in the parametric and the objective space. In *International Conference on Computational Robotics and Autonomous Systems (CIRAS), Centre for Intelligent Control, National University of Singapore*, 2005.

[22] S. Chand and M. Wagner. Evolutionary many-objective optimization: a quick-start guide. *Surveys in Operations Research and Management Science*, 20(2):35–42, 2015.

[23] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In R. C. Gunning, editor, *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. Princeton Univ. Press, Princeton, NJ, 1970.

[24] H. Chen, R. Cheng, W. Pedrycz, and Y. Jin. Solving many-objective optimization problems via multistage evolutionary search. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.

[25] H. Chen, Y. Tian, W. Pedrycz, G. Wu, R. Wang, and L. Wang. Hyperplane assisted evolutionary algorithm for many-objective optimization problems. *IEEE Transactions on Cybernetics*, 50(7):3367–3380, 2020.

[26] H. Chen, G. Wu, W. Pedrycz, P. N. Suganthan, L. Xing, and X. Zhu. An adaptive resource allocation strategy for objective space partition-based multiobjective optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.

[27] H. Chen, X. Zhu, W. Pedrycz, S. Yin, G. Wu, and H. Yan. Pea: Parallel evolutionary algorithm by separating convergence and diversity for large-scale multi-objective op-

timization. In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pages 223–232. IEEE, 2018.

[28] R. Cheng, Y. Jin, M. Olhofer, and B. Sendhoff. A reference vector guided evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(5):773–791, 2016.

[29] T. Chugh, Y. Jin, K. Miettinen, J. Hakanen, and K. Sindhya. A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 22(1):129–142, Feb 2018.

[30] C. A. C. Coello. An introduction to multi-objective particle swarm optimizers. In A. Gaspar-Cunha, R. Takahashi, G. Schaefer, and L. Costa, editors, *Soft Computing in Industrial Applications*, pages 3–12, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[31] C. A. C. Coello. An introduction to evolutionary multi-objective optimization with some applications in pattern recognition. In E. Bayro-Corrochano and E. Hancock, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 1–13, Cham, 2014. Springer International Publishing.

[32] C. A. C. Coello. Recent results and open problems in evolutionary multiobjective optimization. In C. Martín-Vide, R. Neruda, and M. A. Vega-Rodríguez, editors, *Theory and Practice of Natural Computing*, pages 3–21, Cham, 2017. Springer International Publishing.

[33] C. A. C. Coello. Constraint-handling techniques used with evolutionary algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '19, page 485–506, New York, NY, USA, 2019. ACM.

[34] C. A. C. Coello, S. G. Brambila, J. F. Gamboa, M. G. C. Tapia, and R. H. Gómez. Evolutionary multiobjective optimization: open research areas and some challenges lying ahead. *Complex & Intelligent Systems*, 6(2):221–236, 2020.

[35] C. A. C. Coello, G. B. Lamont, and D. A. Van Veldhuizen. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007.

[36] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga. Handling multiple objectives with particle swarm optimization. *IEEE Transactions on evolutionary computation*, 8(3):256–279, 2004.

[37] M. B. Cohen, S. Elder, C. Musco, C. Musco, and M. Persu. Dimensionality reduction for k-means clustering and low rank approximation. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 163–172, New York, NY, USA, 2015. ACM.

[38] D. W. Corne, N. R. Jerram, J. D. Knowles, and M. J. Oates. PESA-II: Region-based selection in evolutionary multiobjective optimization. In *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*, pages 283–290. Morgan Kaufmann Publishers Inc., 2001.

[39] B. Cui, C. Fan, J. Munk, N. Mao, F. Xiao, J. Dong, and T. Kuruganti. A hybrid building thermal modeling approach for predicting temperatures in typical, detached, two-story houses. *Applied Energy*, 236:101 – 116, 2019.

[40] I. Das and J. E. Dennis. Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, 8(3):631–657, 1998.

[41] S. Das, S. S. Mullick, and P. N. Suganthan. Recent advances in differential evolution – an updated survey. *Swarm and Evolutionary Computation*, 27:1–30, 2016.

[42] S. Dasgupta, S. Das, A. Biswas, and A. Abraham. On stability and convergence of the population-dynamics in differential evolution. *AI Communications*, 22(1):1–20, 2009.

[43] K. Deb. *Multi-objective optimization using evolutionary algorithms*, volume 16. John Wiley & Sons, 2001.

[44] K. Deb and R. B. Agrawal. Simulated binary crossover for continuous search space. *Complex systems*, 9(2):115–148, 1995.

[45] K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with

box constraints. *IEEE Transactions on Evolutionary Computation*, 18(4):577–601, 2014.

[46] K. Deb, M. Mohan, and S. Mishra. Evaluating the ε-domination based multi-objective evolutionary algorithm for a quick computation of pareto-optimal solutions. *Evolutionary computation*, 13(4):501–525, 2005.

[47] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[48] K. Deb and D. Saxena. Searching for pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. In *Proceedings of the World Congress on Computational Intelligence (WCCI-2006)*, pages 3352–3360, 2006.

[49] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA)*, pages 825–830. Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA), 2002.

[50] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable test problems for evolutionary multiobjective optimization. In A. Abraham, L. Jain, and R. Goldberg, editors, *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications*, pages 105–145, London, 2005. Springer London.

[51] K. Deb and S. Tiwari. Omni-optimizer: A procedure for single and multi-objective optimization. In C. A. C. Coello, A. H. Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization*, pages 47–61, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[52] J. Demšar. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7:1–30, 2006.

[53] S. D'Oca and T. Hong. A data-mining approach to discover patterns of window opening and closing behavior in offices. *Building and Environment*, 82:726–739, 2014.

[54] N. Drechsler, R. Drechsler, and B. Becker. Multi-objective optimisation based on relation favour. In *Evolutionary Multi-criterion Optimization*, pages 154–166. Springer, 2001.

[55] K. Elsayed and C. Lacor. CFD modeling and multi-objective optimization of cyclone geometry using desirability function, artificial neural networks and genetic algorithms. *Applied Mathematical Modelling*, 37(8):5680–5704, 2013.

[56] Q. Fan and X. Yan. Solving multimodal multiobjective problems through zoning search. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pages 1–12, 2019.

[57] Y. Fan and X. Xia. A multi-objective optimization model for building envelope retrofit planning. *Energy Procedia*, 75:1299–1304, 2015.

[58] M. Farina, K. Deb, and P. Amato. Dynamic multiobjective optimization problems: test cases, approximations, and applications. *IEEE Transactions on Evolutionary Computation*, 8(5):425–442, Oct 2004.

[59] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975.

[60] C. M. Fonseca and P. J. Fleming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary computation*, 3(1):1–16, 1995.

[61] R. H. Gómez and C. A. C. Coello. Improved metaheuristic based on the R2 indicator for many-objective optimization. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 679–686. ACM, 2015.

[62] A. P. Guerreiro and C. M. Fonseca. Hypervolume sharpe-ratio indicator: Formalization and first theoretical results. In J. Handl, E. Hart, P. R. Lewis, M. López-Ibáñez, G. Ochoa, and B. Paechter, editors, *Parallel Problem Solving from Nature – PPSN XIV*, pages 814–823, Cham, 2016. Springer International Publishing.

[63] D. Hadka and P. Reed. Borg: An auto-adaptive many-objective evolutionary computing framework. *Evolutionary computation*, 21(2):231–259, 2013.

[64] F. Haldi and D. Robinson. The impact of occupants' behaviour on building energy demand. *Journal of Building Performance Simulation*, 4(4):323–338, 2011.

[65] N. Hamada and K. Goto. Data-driven analysis of pareto set topology. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '18, pages 657–664, New York, NY, USA, 2018. ACM.

[66] T. Hanne. On the convergence of multiobjective evolutionary algorithms. *European Journal of Operational Research*, 117(3):553 – 564, 1999.

[67] Z. He and G. G. Yen. An improved visualization approach in many-objective optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 1618–1625. IEEE, 2016.

[68] Z. He and G. G. Yen. Visualization and performance metric in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(3):386–402, June 2016.

[69] Z. He, G. G. Yen, and J. Zhang. Fuzzy-based pareto optimality for many-objective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 18(2):269–285, 2014.

[70] M. Helbig and A. P. Engelbrecht. Performance measures for dynamic multi-objective optimisation algorithms. *Information Sciences*, 250:61 – 81, 2013.

[71] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pages 65–70, 1979.

[72] T. Hong, S. C. Taylor-Lange, S. D'Oca, D. Yan, and S. P. Corgnati. Advances in research and applications of energy-related occupant behavior in buildings. *Energy and Buildings*, 116:694–702, 2016.

[73] S. Huband, L. Barone, L. While, and P. Hingston. A scalable multi-objective test problem toolkit. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 280–295. Springer, 2005.

[74] S. Huband, P. Hingston, L. Barone, and L. While. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506, 2006.

[75] E. J. Hughes. MSOPS-II: A general-purpose many-objective optimiser. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pages 3944–3951. IEEE, 2007.

[76] H. Ishibuchi, N. Akedo, and Y. Nojima. A many-objective test problem for visually examining diversity maintenance behavior in a decision space. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11, pages 649–656, New York, NY, USA, 2011. ACM.

[77] H. Ishibuchi, R. Imada, N. Masuyama, and Y. Nojima. Comparison of hypervolume, IGD and IGD+ from the viewpoint of optimal distributions of solutions. In K. Deb, E. Goodman, C. A. Coello Coello, K. Klamroth, K. Miettinen, S. Mostaghim, and P. Reed, editors, *Evolutionary Multi-Criterion Optimization*, pages 332–345, Cham, 2019. Springer International Publishing.

[78] H. Ishibuchi, H. Masuda, and Y. Nojima. A study on performance evaluation ability of a modified inverted generational distance indicator. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 695–702. ACM, 2015.

[79] H. Ishibuchi, T. Matsumoto, N. Masuyama, and Y. Nojima. Effects of dominance resistant solutions on the performance of evolutionary multi-objective and many-objective algorithms. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, GECCO '20, pages 507–515, New York, NY, USA, 2020. ACM.

[80] H. Ishibuchi and T. Murata. A multi-objective genetic local search algorithm and its application to flowshop scheduling. *IEEE transactions on systems, man, and cybernetics, part C (applications and reviews)*, 28(3):392–403, 1998.

[81] H. Ishibuchi, L. M. Pang, and K. Shang. Population size specification for fair comparison of multi-objective evolutionary algorithms. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1095–1102. IEEE, 2020.

[82] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima. Evolutionary many-objective optimization by NSGA-II and MOEA/D with large populations. In *2009 IEEE International Conference on Systems, Man and Cybernetics*, pages 1758–1763. IEEE, 2009.

[83] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima. How to compare many-objective algorithms under different settings of population and archive sizes. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 1149–1156. IEEE, 2016.

[84] H. Ishibuchi, Y. Setoguchi, H. Masuda, and Y. Nojima. Performance of decomposition-based many-objective algorithms strongly depends on pareto front shapes. *IEEE Transactions on Evolutionary Computation*, 21(2):169–190, 2017.

[85] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. Evolutionary many-objective optimization: A short review. In *Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on*, pages 2419–2426, 2008.

[86] A. L. Jaimes, C. A. C. Coello, and J. E. U. Barrientos. Online objective reduction to deal with many-objective problems. In *Evolutionary Multi-criterion Optimization*, pages 423–437. Springer, 2009.

[87] A. L. Jaimes, C. A. C. Coello, and D. Chakraborty. Objective reduction using a feature selection technique. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 673–680. ACM, 2008.

[88] A. L. Jaimes, A. Oyama, and K. Fujii. Space trajectory design: Analysis of a real-world many-objective optimization problem. In *2013 IEEE Congress on Evolutionary Computation (CEC)*, pages 2809–2816. IEEE, 2013.

[89] H. Jain and K. Deb. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach. *IEEE Transactions on Evolutionary Computation*, 18(4):602–622, 2014.

[90] J. Jamal, R. Montemanni, D. Huber, M. Derboni, and A. E. Rizzoli. A multi-modal and multi-objective journey planner for integrating carpooling and public transport. *Journal of Traffic and Logistics Engineering*, 5(2):68–72, December 2017.

[91] M. Javadi, H. Zille, and S. Mostaghim. Modified crowding distance and mutation for multimodal multi-objective optimization. In *Proceedings of the Genetic and Evo-*

lutionary Computation Conference Companion, GECCO '19, pages 211–212, New York, NY, USA, 2019. ACM.

[92] S. Jiang, Z. Cai, J. Zhang, and Y.-S. Ong. Multiobjective optimization by decomposition with pareto-adaptive weight vectors. In *Natural Computation (ICNC), 2011 Seventh International Conference on*, volume 3, pages 1260–1264. IEEE, 2011.

[93] S. Jiang, Y.-S. Ong, J. Zhang, and L. Feng. Consistencies and contradictions of performance metrics in multiobjective optimization. *IEEE Transactions on Cybernetics*, 44(12):2391–2404, Dec 2014.

[94] Y. Jin. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, 1(2):61 – 70, 2011.

[95] Y. Jin. Scalable model based evolutionary multi-objective optimization. In *2015 7th International Joint Conference on Computational Intelligence (IJCCI)*, volume 3, pages 7–7. IEEE, 2015.

[96] Y. Jin, S. Oh, and M. Jeon. Incremental approximation of nonlinear constraint functions for evolutionary constrained optimization. In *IEEE Congress on Evolutionary Computation*, pages 1–8. IEEE, 2010.

[97] Y. Jin, T. Okabe, and B. Sendhoff. Solving three-objective optimization problems using evolutionary dynamic weighted aggregation: Results and analysis. In E. Cantú-Paz, J. A. Foster, K. Deb, L. D. Davis, R. Roy, U.-M. O'Reilly, H.-G. Beyer, R. Standish, G. Kendall, S. Wilson, M. Harman, J. Wegener, D. Dasgupta, M. A. Potter, A. C. Schultz, K. A. Dowsland, N. Jonoska, and J. Miller, editors, *Genetic and Evolutionary Computation — GECCO 2003*, pages 636–637, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[98] Y. Jin, C. Yang, J. Ding, and T. Chai. Reference point based prediction for evolutionary dynamic multiobjective optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 3769–3776. IEEE, 2016.

[99] R. Katuwal and P. N. Suganthan. Dropout and dropconnect based ensemble of random vector functional link neural network. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1772–1778, Nov 2018.

[100] L. Kaufman and P. J. Rousseeuw. Partitioning around medoids (program pam). *Finding groups in data: an introduction to cluster analysis*, pages 68–125, 1990.

[101] W. Khan and Q. Zhang. MOEA/D-DRA with two crossover operators. In *Computational Intelligence (UKCI), 2010 UK Workshop on*, pages 1–6. IEEE, 2010.

[102] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[103] F. Kudo, T. Yoshikawa, and T. Furuhashi. A study on analysis of design variables in pareto solutions for conceptual design optimization problem of hybrid rocket engine. In *2011 IEEE Congress of Evolutionary Computation (CEC)*, pages 2558–2562, June 2011.

[104] J. Langevin, J. Wen, and P. L. Gurian. Modeling thermal comfort holistically: Bayesian estimation of thermal sensation, acceptability, and preference distributions for office building occupants. *Building and Environment*, 69:206–226, 2013.

[105] Y. Laurillau, V.-B. Nguyen, J. Coutaz, G. Calvary, N. Mandran, F. Camara, and R. Balzarini. The top-slider for multi-criteria decision making by non-specialists. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, pages 642–653. ACM, 2018.

[106] B. Li, J. Li, K. Tang, and X. Yao. Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys*, 48(1):13:1–13:35, Sept. 2015.

[107] H. Li, K. Deb, Q. Zhang, and P. N. Suganthan. Challenging novel many and multi-objective bound constrained benchmark problems. In *Technical Report.* Technical Report, updated 11 Jan, 2018.

[108] H. Li, K. Deb, Q. Zhang, P. N. Suganthan, and L. Chen. Comparison between MOEA/D and NSGA-III on a set of novel many and multi-objective benchmark problems with challenging difficulties. *Swarm and Evolutionary Computation*, 46:104–117, 2019.

[109] K. Li, K. Deb, Q. Zhang, and S. Kwong. An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE Transactions on Evolutionary Computation*, 19(5):694–716, 2015.

[110] L.-M. Li, K.-D. Lu, G.-Q. Zeng, L. Wu, and M.-R. Chen. A novel real-coded population-based extremal optimization algorithm with polynomial mutation: A non-parametric statistical study on continuous optimization problems. *Neurocomputing*, 174:577 – 587, 2016.

[111] Y. Li, Y. Zhou, Z. Zhan, and J. Zhang. A primary theoretical study on decomposition-based multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 20(4):563–576, Aug 2016.

[112] 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. *Technical Report, Computational Intelligence Laboratory, Zhengzhou University*, 2019.

[113] 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)*, pages 2454–2461. IEEE, 2016.

[114] P. Limbourg and D. E. S. Aponte. An optimization algorithm for imprecise multi-objective problem functions. In *2005 IEEE Congress on Evolutionary Computation*, volume 1, pages 459–466. IEEE, 2005.

[115] H.-L. Liu, L. Chen, K. Deb, and E. D. Goodman. Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms. *IEEE Transactions on Evolutionary Computation*, 21(3):408–425, 2017.

[116] H.-L. Liu, L. Chen, Q. Zhang, and K. Deb. An evolutionary many-objective optimisation algorithm with adaptive region decomposition. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 4763–4769, July 2016.

[117] H.-L. Liu, F. Gu, and Q. Zhang. Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems. *IEEE Transactions on Evolutionary Computation*, 18(3):450–455, 2014.

[118] Y. Liu, G. G. Yen, and D. Gong. A multi-modal multi-objective evolutionary algorithm using two-archive and recombination strategies. *IEEE Transactions on Evolutionary Computation*, 23(4):660–674, Aug 2019.

[119] J. Luo, Y. Yang, X. Li, Q. Liu, M. Chen, and K. Gao. A decomposition-based multi-objective evolutionary algorithm with quality indicator. *Swarm and Evolutionary Computation*, 39:339 – 355, 2018.

[120] K. Maity, R. Sengupta, and S. Saha. MM-NAEMO : Multimodal neighborhood-sensitive archived evolutionary many-objective optimization algorithm. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 286–294, June 2019.

[121] S. Z. Martínez and C. A. C. Coello. A multi-objective particle swarm optimizer based on decomposition. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11, pages 69–76, New York, NY, USA, 2011. ACM.

[122] Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947.

[123] Y. Mei, K. Tang, and X. Yao. Decomposition-based memetic algorithm for multiobjective capacitated arc routing problem. *IEEE Transactions on Evolutionary Computation*, 15(2):151–165, 2011.

[124] A. Menchaca-Mendez and C. A. C. Coello. GDE-MOEA: a new moea based on the generational distance indicator and ε-dominance. In *2015 IEEE Congress on Evolutionary Computation (CEC)*, pages 947–955. IEEE, 2015.

[125] S. Mishra and C. A. C. Coello. An approach for non-domination level update problem in steady-state evolutionary algorithms with parallelism. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 1006–1013. IEEE, 2019.

[126] M. W. Mkaouer, M. Kessentini, S. Bechikh, K. Deb, and M. Ó Cinnéide. High dimensional search-based software engineering: finding tradeoffs among 15 objectives for automating software refactoring using NSGA-III. In *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, pages 1263–1270. ACM, 2014.

[127] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, and C. A. C. Coello. A survey of multiobjective evolutionary algorithms for data mining: Part I. *IEEE Transactions on Evolutionary Computation*, 18(1):4–19, Feb 2014.

[128] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, and C. A. C. Coello. Survey of multiobjective evolutionary algorithms for data mining: Part II. *IEEE Transactions on Evolutionary Computation*, 18(1):20–35, Feb 2014.

[129] Y. Narahari. *Game theory and mechanism design*, volume 4. World Scientific, 2014.

[130] W. O'Brien, K. Kapsis, and A. K. Athienitis. Manually-operated window shade patterns in office buildings: A critical review. *Building and Environment*, 60:319–338, 2013.

[131] S. Otake, T. Yoshikawa, and T. Furuhashi. Basic study on aggregation of objective functions in many-objective optimization problems. In *World Automation Congress (WAC), 2010*, pages 1–6. IEEE, 2010.

[132] M. Pal, A. A. Alyafi, S. Bandyopadhyay, S. Ploix, and P. Reignier. Enhancing comfort of occupants in energy buildings. In S. Kar, U. Maulik, and X. Li, editors, *Operations Research and Optimization*, pages 133–144, Singapore, 2018. Springer Singapore.

[133] M. Pal, A. A. Alyafi, S. Ploix, P. Reignier, and S. Bandyopadhyay. Unmasking the causal relationships latent in the interplay between occupant's actions and indoor ambience: A building energy management outlook. *Applied Energy*, 238:1452 – 1470, 2019.

[134] M. Pal and S. Bandyopadhyay. Reliability of convergence metric and hypervolume indicator for many-objective optimization. In *2016 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*, pages 511–515. IEEE, Jan 2016.

[135] M. Pal and S. Bandyopadhyay. Consensus of subjective preferences of multiple occupants for building energy management. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1815–1822, Nov 2018.

[136] M. Pal and S. Bandyopadhyay. Exploration of many-objective feature selection for recognition of motor imagery tasks. In P. K. Sa, M. N. Sahoo, M. Murugappan, Y. Wu, and B. Majhi, editors, *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*, pages 331–337, Singapore, 2018. Springer Singapore.

[137] M. Pal and S. Bandyopadhyay. Differential evolution for multi-modal multi-objective problems. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '19, pages 1399–1406, New York, NY, USA, 2019. ACM.

[138] M. Pal and S. Bandyopadhyay. ESOEA: Ensemble of single objective evolutionary algorithms for many-objective optimization. *Swarm and Evolutionary Computation*, 50:100511, 2019.

[139] M. Pal and S. Bandyopadhyay. Multi-modality of occupants' actions for multi-objective building energy management. In S. Bhattacharyya, P. Dutta, and K. Datta, editors, *Intelligence Enabled Research: DoSIER 2020*, pages 11–19, Singapore, 2021. Springer Singapore.

[140] M. Pal and S. Bandyopadhyay. Decomposition in decision and objective space for multi-modal multi-objective optimization. *Swarm and Evolutionary Computation*, 2021 (accepted, in press).

[141] M. Pal, S. Saha, and S. Bandyopadhyay. Clustering based online automatic objective reduction to aid many-objective optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 1131–1138. IEEE, July 2016.

[142] M. Pal, S. Saha, and S. Bandyopadhyay. DECOR: Differential evolution using clustering based objective reduction for many-objective optimization. *Information Sciences*, 423:200 – 218, 2018.

[143] M. Pal, R. Sengupta, S. Bandyopadhyay, A. A. Alyafi, S. Ploix, P. Reignier, and S. Saha. Analysis of optimizers to regulate occupant's actions for building energy management. In *2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR)*, pages 1–6, Dec 2017.

[144] L. Pan, C. He, Y. Tian, H. Wang, X. Zhang, and Y. Jin. A classification-based surrogate-assisted evolutionary algorithm for expensive many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 23(1):74–88, Feb 2019.

[145] S. Papadopoulos and E. Azar. Optimizing HVAC operation in commercial buildings: a genetic algorithm multi-objective optimization framework. In *Proceedings of the 2016 Winter Simulation Conference*, pages 1725–1735. IEEE Press, 2016.

[146] R. P. Parouha and K. N. Das. A memory based differential evolution algorithm for unconstrained optimization. *Applied Soft Computing*, 38:501 – 517, 2016.

[147] A. L. Pisello and F. Asdrubali. Human-based energy retrofits in residential buildings: A cost-effective alternative to traditional physical strategies. *Applied Energy*, 133:224–235, 2014.

[148] Y. Qi, X. Ma, F. Liu, L. Jiao, J. Sun, and J. Wu. MOEA/D with adaptive weight adjustment. *Evolutionary computation*, 22(2):231–264, 2014.

[149] A. K. Qin and P. N. Suganthan. Self-adaptive differential evolution algorithm for numerical optimization. In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, volume 2, pages 1785–1791. IEEE, 2005.

[150] Z. Qingfu and L. Hui. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731, 2007.

[151] X. Qiu, J. X. Xu, K. C. Tan, and H. A. Abbass. Adaptive cross-generation differential evolution operators for multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 20(2):232–244, 2016.

[152] F. D. Ramalho, P. Y. Ekel, W. Pedrycz, J. G. P. Júnior, and G. L. Soares. Multicriteria decision making under conditions of uncertainty in application to multiobjective allocation of resources. *Information Fusion*, 49:249–261, 2019.

[153] T. Robič and B. Filipič. DEMO: Differential evolution for multiobjective optimization. In C. A. C. Coello, A. H. Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization*, pages 520–533, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[154] M. Rong, D. Gong, W. Pedrycz, and L. Wang. A multimodel prediction method for dynamic multiobjective evolutionary optimization. *IEEE Transactions on Evolutionary Computation*, 24(2):290–304, 2019.

[155] G. Rudolph and A. Agapie. Convergence properties of some multi-objective evolutionary algorithms. In *Evolutionary Computation, 2000. Proceedings of the 2000 Congress on*, volume 2, pages 1010–1016. IEEE, 2000.

[156] S. A. Sadeghi and P. Karava. Stochastic model predictive control of mixed-mode buildings based on probabilistic interactions of occupants with window blinds. In *Third International High Performance Buildings Conference, Purdue, July 14-17*, pages 1–10. Herrick Laboratories Purdue University, 2014.

[157] D. K. Saxena, J. A. Duro, A. Tiwari, K. Deb, and Z. Qingfu. Objective reduction in many-objective optimization: Linear and nonlinear algorithms. *IEEE Transactions on Evolutionary Computation*, 17(1):77–99, 2013.

[158] L. Scanu. *Towards archetypes of self-tuned models for connected buildings.* PhD thesis, Université Grenoble Alpes, 2017.

[159] L. Scanu, S. Ploix, P. Bernaud, and E. Wurtz. Model tuning approach for energy management of office and apartment settings. In *Proceedings of the 15th IBPSA Conference (Building Simulation 2017)*, pages 858–867, San Francisco, CA, USA, 2017. International Building Performance Simulation Association (IBPSA).

[160] R. Sengupta, M. Pal, S. Saha, and S. Bandyopadhyay. NAEMO: Neighborhood-sensitive archived evolutionary many-objective optimization algorithm. *Swarm and Evolutionary Computation*, 46:201 – 218, 2019.

[161] R. Sengupta, M. Pal, S. Saha, and S. Bandyopadhyay. Population dynamics indicators for evolutionary many-objective optimization. In C. R. Panigrahi, A. K. Pujari, S. Misra, B. Pati, and K.-C. Li, editors, *Progress in Advanced Computing and Intelligent Engineering*, pages 261–271, Singapore, 2019. Springer Singapore.

[162] H. Sharma, J. C. Bansal, and K. V. Arya. Fitness based differential evolution. *Memetic Computing*, 4(4):303–316, Dec 2012.

[163] H. K. Singh and K. Deb. Investigating the equivalence between PBI and AASF scalarization for multi-objective optimization. *Swarm and Evolutionary Computation*, 53:100630, 2020.

[164] R. Storn and K. Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, Dec 1997.

[165] P. Suganthan, S. Das, S. Mukherjee, and S. Chatterjee. Adaptation methods in differential evolution: A review. In *20th International Conference on Soft Computing MENDEL*, volume 2014, pages 131–140, 2014.

[166] P. N. Suganthan. Shape indexing using self-organizing maps. *IEEE Transactions on Neural Networks*, 13(4):835–840, 2002.

[167] P. N. Suganthan. Structural pattern recognition using genetic algorithms. *Pattern Recognition*, 35(9):1883–1893, 2002.

[168] P. N. Suganthan. Differential evolution: Recent advances. In *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation*, GECCO '13 Companion, pages 845–876, New York, NY, USA, 2013. ACM.

[169] R. Tanabe and H. Ishibuchi. A decomposition-based evolutionary algorithm for multi-modal multi-objective optimization. In A. Auger, C. M. Fonseca, N. Lourenço, P. Machado, L. Paquete, and D. Whitley, editors, *Parallel Problem Solving from Nature – PPSN XV*, pages 249–261, Cham, 2018. Springer International Publishing.

[170] R. Tanabe and H. Ishibuchi. A niching indicator-based multi-modal many-objective optimizer. *Swarm and Evolutionary Computation*, 49:134 – 146, 2019.

[171] R. Tanabe and H. Ishibuchi. A review of evolutionary multi-modal multi-objective optimization. *IEEE Transactions on Evolutionary Computation*, pages 1–9, 2019.

[172] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman. How to grow a mind: Statistics, structure, and abstraction. *science*, 331(6022):1279–1285, 2011.

[173] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin. An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility. *IEEE Transactions on Evolutionary Computation*, 22(4):609–622, 2017.

[174] Y. Tian, X. Zhang, R. Cheng, and Y. Jin. A multi-objective evolutionary algorithm based on an enhanced inverted generational distance metric. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 5222–5229, July 2016.

[175] C. Touati, E. Altman, and J. Galtier. Generalized Nash Bargaining Solution for bandwidth allocation. *Computer Networks*, 50(17):3242 – 3263, 2006.

[176] T. Tušar and B. Filipič. Differential evolution versus genetic algorithms in multiobjective optimization. In S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, and T. Murata, editors, *Evolutionary Multi-Criterion Optimization*, pages 257–271, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[177] M. A. ul Haq, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, F. Hussin, and D. M. Said. A review on lighting control technologies in commercial buildings, their performance and affecting factors. *Renewable and Sustainable Energy Reviews*, 33:268–279, 2014.

[178] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, Dec 2007.

[179] M. Wagner and F. Neumann. A fast approximation-guided evolutionary multi-objective algorithm. In *Proceedings of the 15th annual conference on Genetic and evolutionary computation*, pages 687–694. ACM, 2013.

[180] W. Wang, J. Chen, G. Huang, and Y. Lu. Energy efficient HVAC control for an IPS-enabled large space in commercial buildings through dynamic spatial occupancy distribution. *Applied Energy*, 207:305 – 323, 2017. Transformative Innovations for a Sustainable Future – Part II.

[181] R. Ward, R. Choudhary, Y. Heo, and A. Rysanek. Exploring the impact of different parameterisations of occupant-related internal loads in building energy simulation. *Energy and Buildings*, 123:92–105, 2016.

[182] G. Wu, R. Mallipeddi, and P. N. Suganthan. Ensemble strategies for population-based optimization algorithms–a survey. *Swarm and Evolutionary Computation*, 44:695–711, 2019.

[183] L. Yang, H. Yan, and J. C. Lam. Thermal comfort and building energy consumption implications – a review. *Applied Energy*, 115:164–173, 2014.

[184] S. Yang, M. Li, X. Liu, and J. Zheng. A grid-based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 17(5):721–736, 2013.

[185] Z. Yang, K. Tang, and X. Yao. Self-adaptive differential evolution with neighborhood search. In *Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on*, pages 1110–1116. IEEE, 2008.

[186] G. Yu, Y. Jin, and M. Olhofer. An a priori knee identification multi-objective evolutionary algorithm based on α-dominance. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '19, pages 241–242, New York, NY, USA, 2019. ACM.

[187] Y. Yuan, X. Hua, W. Bo, and Y. Xin. A new dominance relation-based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(1):16–37, 2016.

[188] C. Yue, B. Qu, and J. Liang. A multiobjective particle swarm optimizer using ring topology for solving multimodal multiobjective problems. *IEEE Transactions on Evolutionary Computation*, 22(5):805–817, Oct 2018.

[189] C. T. Yue, J. J. Liang, B. Y. Qu, K. J. Yu, and H. Song. Multimodal multiobjective optimization in feature selection. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 302–309, June 2019.

[190] H. Zhang, X. Zhang, X. Z. Gao, and S. Song. Self-organizing multiobjective optimization based on decomposition with neighborhood ensemble. *Neurocomputing*, 173:1868–1884, 2016.

[191] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, and S. Tiwari. Multiobjective optimization test instances for the CEC 2009 special session and competition. *University of Essex, Colchester, UK and Nanyang technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report*, 264, 2008.

[192] X. Zhang, Y. Tian, and Y. Jin. A knee point-driven evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 19(6):761–776, 2015.

[193] Z. Zhang, X. Wang, and J. Lu. Multi-objective immune genetic algorithm solving nonlinear interval-valued programming. *Engineering Applications of Artificial Intelligence*, 67:235–245, 2018.

[194] J. Zhao, B. Lasternas, K. P. Lam, R. Yun, and V. Loftness. Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. *Energy and Buildings*, 82:341–355, 2014.

[195] S.-Z. Zhao, P. N. Suganthan, and Q. Zhang. Decomposition-based multiobjective evolutionary algorithm with an ensemble of neighborhood sizes. *IEEE Transactions on Evolutionary Computation*, 16(3):442–446, 2012.

[196] X. Zhao, J. Venkateswaran, and Y.-J. Son. Modeling human operator decision-making in manufacturing systems using BDI agent paradigm. In *IIE Annual Conference. Proceedings*, pages 1–6. Institute of Industrial and Systems Engineers (IISE), 2005.

[197] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang. Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1):32–49, 2011.

[198] A. Zhou, Q. Zhang, and Y. Jin. Approximating the set of pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation*, 13(5):1167–1189, 2009.

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

[200] E. Zitzler and S. Künzli. Indicator-based selection in multiobjective search. In *International Conference on Parallel Problem Solving from Nature*, pages 832–842. Springer, 2004.

[201] E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. *Evolutionary Methods for Design, Optimization, and Control*, pages 95–100, 2002.