Universiteit Leiden
The Netherlands

Multi-objective evolutionary algorithms for optimal scheduling
Wang, Y.

Citation
Wang, Y. (2022, January 19). *Multi-objective evolutionary algorithms for optimal scheduling*. Retrieved from https://hdl.handle.net/1887/3250350

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3250350

Note: To cite this publication please use the final published version (if applicable).

Chapter 2

Preliminaries

This chapter provides an introduction to critical basics of optimization, multi-objective optimization, multi-objective evolutionary algorithm, and the discussion of order relations for multi-objective optimization.

2.1 Optimization

An optimization problem is the problem of finding the best solution from all feasible solutions. All sorts of optimization problems arise in different disciplines from mathematics, computer science to engineering and economics, and so on.

2.1.1 Applications

Some applications listed below can give a rough impression on optimization problems. It is worthy noting that this is just a drop in the ocean, the real-world optimization problems go far beyond these disciplines and applications.

Agriculture: managing river basins to satisfy urban and agricultural consumptive demand, also in-stream environmental demands [126].

Architecture: designing a building with respect to thermal comfort, energy efficiency, and construction cost criteria [58].

Aviation: planning airport construction to minimize the cost of all the items influenced by the site layout; maximize the safety of airport operations during construction; reduce construction-related security breaches; and improve the safety of construction operations [69].

Aerospace: designing satellite orbits to minimize the spatial resolution requirement (at-nadir resolution and off-nadir resolution) and temporal resolution requirement (the repeat cycle and the revisit time) [98].

Chemistry: finding desirable molecule drug which can improve solubility, metabolic stability, cell permeability, and with reduced side effects. [125].

Engineering: designing hydraulic presser to maximize the nominal pressure rating and fully loaded power while minimizing the oil injection volume [123].

Environment: designing marine protected area networks to maximize network effectiveness, species persistence, and minimize cost of protection [47].

Investment: choosing an optimal set of assets in order to minimize the risk and maximize the profit of the investment [101].

Machine learning: assisting machine learning algorithms to optimize their hyperparameters, selecting models to minimize model complexity and maximize classification accuracy [65].

Manufacturing: making an efficient supply chain plan to minimize total losses of supply chain including production cost, hiring, firing and training cost, raw material and end product inventory holding cost, transportation and shortage cost, simultaneously, minimize the sum of the maximum amount of shortages among the customers' zones in all periods to improve customer satisfaction [78].

Medical: searching for new therapeutic drugs to maximize the potency of the drug, at the same time, minimize synthesis costs and unwanted side effects [108].

Scheduling: determining the vehicle routing to minimize the total distance traveled, the total time required, the total tour cost, and the fleet size, and maximizing the quality of the service and the profit collected [66].

All these problems have in common that a software/ search algorithm framework which can support human decision makers in solving such problems is desirable due to the large number of alternative solutions.

2.1.2 Mathematical Definition

The goal of the optimization process is to find the values of decision variables that result in a maximum or minimum of a function called the objective function. In mathematical terms, optimization problems can be formulated as:

$$\text{Minimize} \quad f(\mathbf{x}) \tag{2.1}$$
$$\text{Subject to} \quad g_i(\mathbf{x}) = 0, \qquad i = 1, \cdots, p \tag{2.2}$$
$$h_j(\mathbf{x}) \geq 0, \qquad j = 1, \cdots, q \tag{2.3}$$
$$\mathbf{x} \in \mathcal{X}. \tag{2.4}$$

Here $\mathbf{x}$ is the set of decision variables. The decision variables are the numerical quantities for which values are under our control and are to be chosen to find an optimal solution. The decision variables consist of independent variables, a vector x containing n decision variables can be represented by: $x = (x_1, x_2...., x_n)^T$. Decision variables may have continuous values which can take on any value in a specified interval or discrete values which are restricted to a specified interval of integers. This leads to continuous optimization and discrete optimization problems. In theory, continuous optimization problems tend to be easier to solve than discrete optimization problems because the information about points in a neighborhood of one decision variable can be deduced more smoothly.

The constraints, i.e., equality constraints in Eq. (2.2) and inequality constraints in Eq. (2.3), are also functions of the decision variables. Their values decide which solutions are feasible. Some of the optimization problems do not have any constraints and they are therefore called unconstrained optimization problems. Sometimes, only simple constraints on the range of the input variables are given by means of intervals. These problems are usually referred to as box-constraints problems. Constrained optimization problems can be reformulated to unconstrained optimization problems in which the constraints are replaced by a penalty term in the objective function.

The optimization process is to find the values of decision variables that result in a maximum or minimum of the objective function $f(\mathbf{x})$, i.e., Eq. (2.1). Without loss of generality, the objective function is to be minimized in this work. In other words, the objective function is a measure to compare alternative solutions. The optimization problems having a single objective function are single-objective optimization problems. But in the real-world, the optimization problems with multiple objective functions, i.e., multi-objective optimization problems, are more common.

2.2 Multi-objective Optimization

A multi-objective optimization problem is an optimization problem with more than one objective function to be minimized. That is to say, instead of one single objective function $f(\mathbf{x})$, multiple objective functions, $f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})$, are optimized simultaneously. Here, m ($m \geq 2$) indicates the number of objectives.

In multi-objective optimization, the objectives are usually conflicting with each other. Therefore, there does not typically exist a feasible solution that minimizes all objective functions simultaneously; and the trade-off among different objectives gives rise to a set of potential compromise solutions. A minimal requirement for a compromise solution is that it should be a Pareto optimal solution. Pareto optimal solutions are solutions that cannot be improved in any of the objectives without deteriorating at least one of the other objectives.

2.2.1 Pareto Optimal and Non-dominated Solutions

The solutions are evaluated by the objective functions which represent a mapping from the *decision space* to the *objective space*. For an optimization problem, the decision space $\mathcal{X}$ comprises all candidate solutions. When the problem has m objectives, an m-dimensional Euclidean space forms its objective space in which objective function vectors coexist and where each coordinate axis corresponds to one objective. For each solution in the decision space, there is a point in the objective space. At the same time, multiple solutions in the decision space may be projected onto the same point in the objective space. A relative comparison between solutions can be achieved by the dominance relation.

Definition 2.1 (Dominance (Objective Space))**.** Given two solutions in the objective space, that is $\mathbf{y}^{(1)} \in \mathbb{R}^m$ and $\mathbf{y}^{(2)} \in \mathbb{R}^m$, solution $\mathbf{y}^{(1)}$ is said to dominate solution $\mathbf{y}^{(2)}$ if and only if $\forall\, i \in \{1, \ldots, m\} : y_i^{(1)} \leq y_i^{(2)}$ and $\exists\, j \in \{1, \ldots, m\} : y_j^{(1)} < y_j^{(2)}$, in symbols $\mathbf{y}^{(1)} \prec \mathbf{y}^{(2)}$.

Definition 2.2 (Dominance (Decision Space))**.** Given two solutions $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ in the decision space, then solution $\mathbf{x}^{(1)}$ is said to dominate $\mathbf{x}^{(2)}$ if and only if $\forall\, i \in \{1, \ldots, m\} : f_i(\mathbf{x}^{(1)}) \leq f_i(\mathbf{x}^{(2)})$ and $\exists\, j \in \{1, \ldots, m\} : f_j(\mathbf{x}^{(1)}) < f_j(\mathbf{x}^{(2)})$.

Definition 2.3 (Incomparability (objective space))**.** Solution $\mathbf{y}^{(1)}$ is said to be incomparable to solution $\mathbf{y}^{(2)}$ if and only if $\mathbf{y}^{(1)} \neq \mathbf{y}^{(2)}$, $\mathbf{y}^{(1)} \nprec \mathbf{y}^{(2)}$ and $\mathbf{y}^{(2)} \nprec \mathbf{y}^{(1)}$, in symbols $\mathbf{y}^{(1)} \| \mathbf{y}^{(2)}$.

Definition 2.4 (Indifference and incomparability (decision space))**.** Solution $\mathbf{x}^{(1)}$ is said to be indifferent to solution $\mathbf{x}^{(2)}$ ($\mathbf{x}^{(1)} \sim \mathbf{x}^{(2)}$) if and only if $\mathbf{f}(\mathbf{x}^{(1)}) = \mathbf{f}(\mathbf{x}^{(2)})$. Here $\mathbf{x}^{(1)} \sim \mathbf{x}^{(2)} \nRightarrow \mathbf{x}^{(1)} = \mathbf{x}^{(2)}$. Solution $\mathbf{x}^{(1)}$ is said to be incomparable to solution $\mathbf{x}^{(2)}$ ($\mathbf{x}^{(1)} \| \mathbf{x}^{(2)}$) if and only if $\mathbf{f}(\mathbf{x}^{(1)}) \| \mathbf{f}(\mathbf{x}^{(2)})$.

Definition 2.5 (Pareto Optimal and Non-dominated Solution)**.** In decision space, a decision vector x^* is a Pareto optimal solution if there does not exist a decision vector x ($x \neq x^*$) that dominates it, i.e., $\nexists\, x \in \mathcal{X} : \mathbf{f}(x) \prec \mathbf{f}(x^*)$. If x^* is Pareto optimal, $\mathbf{f}(x^*)$ is called a non-dominated point (solution).

The set of all Pareto optimal vectors in the decision space is referred to as the *Pareto optimal set* or efficient set; and the image of the Pareto optimal set in the objective space is referred to as the *Pareto Front*.

2.2.2 Pareto Front Geometry

Figure 2.1 shows several typical types of the Pareto fronts: convex, concave, neither convex nor concave and disconnected Pareto fronts. The Pareto front in the bottom left image consists of convex and concave parts.

The Pareto front can be represented by a function, $u : \mathbb{R}^{m-1} \to \mathbb{R}$ and m the number of objectives. A function is said to be convex if it satisfies the following equation [11]:

$$u(\theta x + (1-\theta)y) \leq \theta u(x) + (1-\theta)u(y) \tag{2.5}$$

with x, y in the domain of g and $\theta \in [0,1]$. In words, it means that the line between $(x, u(x))$ and $(y, u(y))$ is above the graph between x to y. Accordingly, a function u is concave if $-u$ is convex.

Multi-objective optimization problems with more than three objectives are called many-objective optimization problems [41] and they form a special and important case of multi-objective optimization problems. An increase in the number of objectives causes a large portion of solutions to become non-dominated. This leads to the difficulty in searching for Pareto optimal solutions, meanwhile, a huge number of solutions may be needed to estimate the entire Pareto front. Many-objective optimization gives rise to a new set of challenges [3, 61]. The need for tackling many-objective

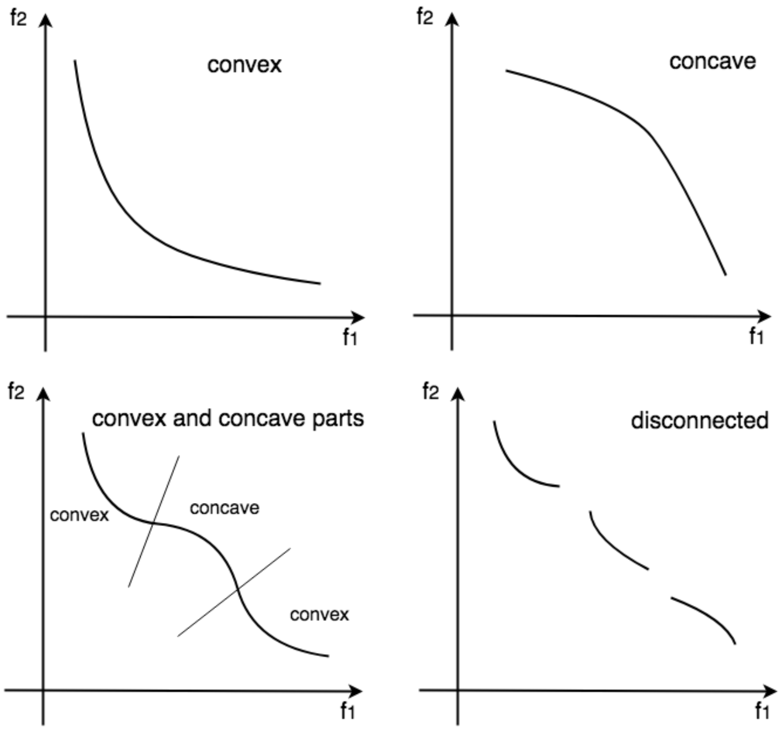

Figure 2.1: Examples of Pareto fronts.

problems became evident recently because it would allow us to solve more complex real world problems.

2.3 Multi-objective Evolutionary Algorithm

EA has been successfully adapted to dealing with multi-objective optimization and these specialised algorithms are called multi-objective evolutionary algorithm (MOEA) or, sometimes, also evolutionary multi-objective optimization algorithm (EMOA). The optimization mechanism of MOEA is very similar to EA, such as population-based search and information exchange among solutions (individuals). One special characteristic of MOEA is the use of the dominance relationship to assign the fitness to each solution in the population. In detail, at each iteration, the objective values are calculated for each individual and then used to determine the relationship of dominance in the population in order to choose a potentially better solution for the creation of the offspring population. At the same time, the ability to maintain diversity within a population of individuals is another key component of MOEA.

2.3.1 Classification

Classical *Pareto dominance-based MOEAs*, such as NSGA-II [29], use Pareto dominance as a first ranking criterion and use a second ranking criterion to maintain and increase diversity. Pareto dominance-based MOEAs have been a mainstream class for a long time in the field of evolutionary multi-objective optimization (EMO). They are very efficient on multi-objective optimization problems with two or three objectives. However, their performance degrades significantly on many-objective optimization problems due to their ineffectiveness in distinguishing the quality of solutions when the number of objectives becomes large.

As the performance assessment of MOEAs reached a mature stage, performance measures (indicators) on the quality of Pareto front approximations were adopted to search for solutions. These indicators capture both convergence and diversity in a single value. Additionally for Pareto compliant indicators, it can be shown that they obtain their maximum in a diversified set of solutions on the Pareto front. In general, *indicator-based MOEAs* (IBEA) [139], such as SMS-EMOA [9] and R2-EMOA [104], have strong theoretical support. However, the commonly used performance indicators lead to a convergence in distribution with a high density on the boundary of the Pareto front, as well as on knee regions [9].

Decomposition is a search paradigm that was originally applied by EMO two decades ago [53] and recently regained prominence from the MOEA/D framework [135] and NSGA-III [26]. *Decomposition-based MOEAs* transform the original multi-objective problem into simpler, single-objective subproblems by means of scalarizations with different weights or reference vectors, therefore they can converge to a well defined, diverse set. However, the central issue in decomposition-based methods is how to select a set of weighting vectors that can provide a well distributed set of Pareto optimal points, given that the location and shape of the Pareto front are unknown a priori. Moreover, the number of weights required to sample a Pareto front with a sufficient resolution suffers an exponential growth from the objective space dimension [51].

2.3.2 Quality Measures

The goal of solving a multi-objective optimization problem is to approximate or compute all or a representative set of Pareto optimal solutions. The quality of the approximation sets is evaluated to compare different algorithms when solving multi-objective optimization problems. The major characteristics for evaluating fronts include convergence and diversity, and the diversity consists of two aspects: distribution and spread. A good multi-objective optimization algorithm is required to generate solutions that are close to the Pareto front, well distributed and spread widely over the entire Pareto front at minimum computational cost.

Among numerous metrics, we choose the following ones to evaluate the quality of the obtained Pareto front approximation, which are also performance metrics commonly used by the evolutionary multi-objective evolutionary community.

Hypervolume Indicator

The hypervolume (HV) indicator [141], previously also known as S metric [137] or Lebesgue measure [72], is one of the most popular indicators for multi-objective optimization. It has been proven that the maximization of this performance measure is equivalent to finding the Pareto front [45] provided it is a finite set. In other cases, it leads to a well distributed approximation of the Pareto optimal set if the number of objectives is small (say ≤ 4). The HV indicator is an unary metric which evaluates one approximation set, and it measures the volume of the objective space jointly dominated by the Pareto front approximation, relative to a reference point $r \in \mathbb{R}^m$.

Definition 2.6 (Hypervolume Indicator)**.**

$$HV(Y, r) = \lambda_m(\bigcup_{y \in Y} [y, r]) \tag{2.6}$$

here λ_m denotes the Lebesgue measure on $\mathbb{R}^m$, with m being the number of objective functions.

The HV indicator considers both convergence and diversity. The HV indicator, and its variations, are the only known unary indicator to be strictly monotonic [138], i.e., if an Pareto front approximation A strictly dominates another Pareto front approximation B, $HV(A, r) > HV(B, r)$. Therefore, the HV indicator is said to be Pareto compliant. The major disadvantage of the HV indicator is calculating hypervolume exactly is NP-hard and exponential in the number of objectives [8]. For a small constant number of objectives, however, there exists fast computation algorithms.

To evaluate the fitness value of each solution in the Pareto front approximation, the hypervolume contribution can be used. The hypervolume contribution of a point $y \in Y$ is defined as the difference between the hypervolume indicator of Y and the

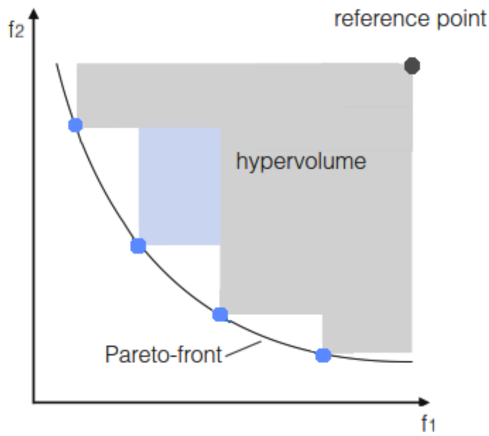

Figure 2.2: Illustration of the hypervolume indicator and hypervolume contribution for a bi-objective problem.

hypervolume indicator of $Y \setminus \{y\}$. Figure 2.2 shows the hypervolume indicator and hypervolume contribution for a bi-objective problem. The size of the blue part in the dominated region is the hypervolume contribution of one solution.

Inverted generational distance (IGD)

IGD [17] has been widely considered as a reliable performance indicator. It is complementary to generational distance (GD). Both IGD and GD use the true Pareto front as a reference set; if the true Pareto front is unknown, the reference set is usually a combination of the non-dominated points of several approximate fronts.

They are given by the following formulas:

Definition 2.7 (IGD Metric)**.**

$$IGD(Y, P) = \frac{1}{|P|}(\sum_{i=1}^{|P|} d(r_i, Y)^2)^{\frac{1}{2}} \tag{2.7}$$

where $|P|$ is the number of points in the reference front P and Y is the obtained Pareto front approximation; $d(r_i, Y)$ denotes the minimum Euclidean distance between a point in the reference front and the solutions in the Pareto front approximation Y.

Definition 2.8 (GD Metric)**.**

$$GD(Y,P) = \frac{1}{|Y|}(\sum_{i=1}^{|Y|} d(y_i,P)^2)^{\frac{1}{2}} \tag{2.8}$$

where $|Y|$ is the number of solutions in the Pareto set approximation Y and P is the reference front; $d(y_i, P)$ denotes the minimum Euclidean distance between solution y_i and the points in the reference front P.

It can be seen that $IGD(Y,P) = GD(P,Y)$, but there is significant difference between IGD and GD. IGD uses the reference front as reference and calculates the distance of each point from the reference front to the Pareto front approximation, which means no part of the reference front (or "true" Pareto front) can be missed. If sufficient members of the reference front are known, IGD could measure both the diversity and the convergence of a Pareto front approximation. The smaller the value of this metric, the closer the obtained front is to the true Pareto front. IGD is efficient to compute in low dimensions of the objective space, but it requires the knowledge of the Pareto front.

Beyond the metrics introduced here, the interested reader is referred to [89] and [4] for a general overview and introduction to performance metrics in multi-objective optimization.

2.3.3 Dominance Relations

The concept of Pareto dominance is of fundamental importance to multi-objective optimization. We use this section to discuss further the Pareto dominance and introduce other dominance relations. Let us first review the basic concept of binary relations and some general properties that binary relations can potentially have (see also [34]).

Definition 2.9 (Binary relation)**.** A binary relation R over a set X is defined as a set of pairs of elements of X, that is, a subset of $X \times X = \{(x,y)|\ x,y \in X\}$. The statement $(x,y) \in R$ reads "x is R-related to y" and is denoted by xRy.

Definition 2.10 (Properties of Binary Relations)**.** Given a set X, a binary relation R is said to be

- reflexive, if and only if $\forall x \in X : (x,x) \in R$.
- irreflexive, if and only if $\forall x \in X, (x,x) \notin R$.

- symmetric, if and only if $\forall x \in X, \forall y \in X : (x, y) \in R \Leftrightarrow (y, x) \in R$.
- asymmetric, if and only if $\forall x \in X, \forall y \in X : (x, y) \in R \Rightarrow (y, x) \notin R$.
- anti-symmetric, if and only if $\forall x \in X, \forall y \in X : (x, y) \in R \wedge (y, x) \in R \Rightarrow x = y$.
- transitive, if and only if $\forall x \in X, \forall y \in X, \forall z \in X : (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$.

Based on these properties, different types of orders then can be defined.

Definition 2.11 (Pre-order, Partial Order, Strict Partial Order)**.** A binary relation R is

- pre-order (aka quasi-order), if and only if it is transitive and reflexive.
- partial order, if and only if it is an antisymmetric pre-order.
- strict partial order, if and only if it is irreflexive and transitive.

Note that a strict partial order is necessarily asymmetric (and therefore also anti-symmetric). Next, the definition of Pareto order is given in the objective space $\mathbb{R}^m$ and it can be viewed as a cone order from a geometrical perspective, as will be shown later on.

Recall Definition 2.1 introduced the concept of dominance in the objective space or Pareto dominance. In this section, to distinguish it from other orders, we denote it with $\prec_{Pareto}$. The Pareto order $\prec_{Pareto}$ is a strict partial order defined in the objective space, i.e., the m-dimensional Euclidean space $\mathbb{R}^m$ with the objective function values being the coordinate axes. It allows a comparison between (some) pairs of feasible solutions in the objective space. Moreover, it is a transitive relation and as in the more general case of a pre-order, minimal elements and maximal elements are defined. In comparison to the more general pre-order, a partial order relation also is constrained by the anti-symmetry axiom, which implicates that indifference falls together with incomparability. The concept of Pareto dominance implies that, for a solution to dominate another one, it should not be worse in any objective and must be strictly better in at least one objective. The Pareto order is a special case of a cone order, which are a family of partial orders defined on vector spaces.

Definition 2.12 (Non-trivial Cone)**.** A set $\mathcal{C} \subset \mathbb{R}^m$ with $\emptyset \neq \mathcal{C} \neq \mathbb{R}^m$ is called a non-trivial cone, if and only if $\forall \alpha \in \mathbb{R}, \alpha > 0, \forall c \in \mathcal{C} : \alpha c \in \mathcal{C}$.

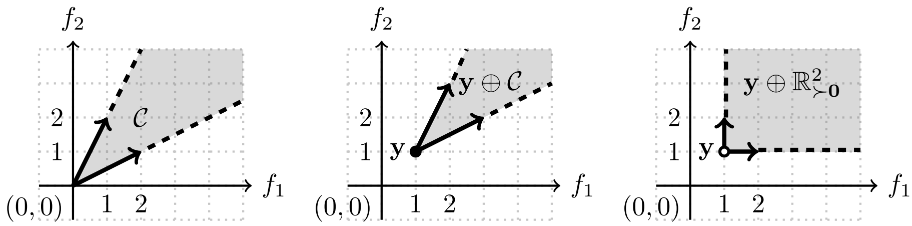

Figure 2.3: Example of a polyhedral cone $\mathcal{C}$ generated by $(2,1)$ and $(1,2)$ (left), Minkowski sum of a singleton $\{\mathbf{y}\} = \{(1,1)\}$ and $\mathcal{C}$ (middle), and Minkowski sum of $\{\mathbf{y}\}$ and the cone $\mathbb{R}^2_{\succ \mathbf{0}}$. The cone $\mathbb{R}^2_{\succ \mathbf{0}}$ is equal to the non-negative quadrant minus $\{(0,0)\}$. (cf. [36])

Definition 2.13 (Minkowski Sum)**.** The Minkowski sum (aka algebraic sum) of two sets $A \in \mathbb{R}^m$ and $B \in \mathbb{R}^m$ is defined as $A \oplus B := \{a + b \mid a \in A \wedge b \in B\}$. Moreover we define $\alpha A = \{\alpha a \mid a \in A\}$.

Figure 2.3 gives an illustration and examples of Minkowski sums, also refer to [36].

Definition 2.14 (Binary Relation Associated to Cone)**.** Given a cone $\mathcal{C}$, the binary relation associated to this cone, notation $R_{\mathcal{C}}$, is defined as follows: $\forall \mathbf{x} \in \mathbb{R}^m, \forall \mathbf{y} \in \mathbb{R}^m : (\mathbf{x}, \mathbf{y}) \in R_{\mathcal{C}}$ if and only if $\mathbf{y} \in \{\mathbf{x}\} \oplus \mathcal{C}$.

For any cone $\mathcal{C}$, the associated binary relation $R_{\mathcal{C}}$ is translation invariant (i.e., if $\forall \mathbf{u} \in \mathbb{R}^m : (\mathbf{x}, \mathbf{y}) \in R_{\mathcal{C}} \Rightarrow (\mathbf{x} + \mathbf{u}, \mathbf{y} + \mathbf{u}) \in R_{\mathcal{C}}$) and multiplication invariant by any positive real (i.e., $\forall \alpha > 0 : (\mathbf{x}, \mathbf{y}) \in R_{\mathcal{C}} \Rightarrow (\alpha \mathbf{x}, \alpha \mathbf{y}) \in R_{\mathcal{C}}$). At the same time, given a binary relation R which is translation invariant and multiplication invariant by any positive real, the set $\mathcal{C}_R := \{\mathbf{y} - \mathbf{x} \mid (\mathbf{x}, \mathbf{y}) \in R\}$ is a cone. The above two operations are inverses of each other, i.e., starting from a cone $\mathcal{C}$, a binary relation $R_{\mathcal{C}}$ which is translation invariant and multiplication invariant by any positive real can be associated to it; starting from a binary relation R which is translation invariant and multiplication invariant by any positive real, a cone $\mathcal{C}_R$ can be obtained. It can be seen there is a natural one to one correspondence between cones and binary relations on $\mathbb{R}^m$ which are translation invariant and multiplication invariant by positive reals (see also [80]).

We restrict our attention to relations which are translation invariant and positive multiplication invariant to get this bijection between cones and relations. Note if a translation invariant and positive multiplication invariant relation R is such that $\emptyset \neq R \neq \mathbb{R}^m \times \mathbb{R}^m$, the associated cone $\mathcal{C}_R$ is non-trivial. Relations associated to non-trivial cones are non-empty and not equal to $\mathbb{R}^m \times \mathbb{R}^m$ as well.

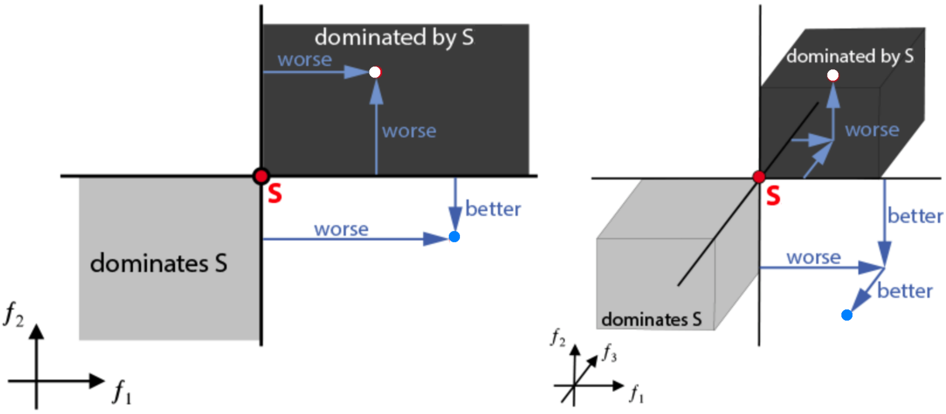

Figure 2.4: Pareto dominance relation and cone.

The reason to view the Pareto dominance relation or Pareto order as derived from a cone is that it gives the opportunity to study this order more geometrically. Figure 2.4 illustrates geometrically the Pareto order and Pareto cone in two and three dimensional spaces. In the objective space, each solution can be located in the vector space based on its objective values. One solution S dominates another solution if all objective values of S are better than the corresponding objective values of another solution, or the objective values of S are equal to but at least one better than the corresponding objective values of another solution. In other words, if a solution is located in the dark gray cone area of S (including the boundaries), it is dominated by S. Similarly, solutions in the light gray cone area of S (including the boundaries) dominate S; solutions in other areas, for examples, the blue points, are incomparable to S.

It can be seen that, in two dimensional space, the cone which associates the Pareto order is the positive quadrant and the angle between two edges of the cone is 90°. Similarly, the Pareto cone in three dimensional space is the positive octant. In the following, the detailed definitions are given.

Definition 2.15. Let m be a natural number bigger or equal to 1, the non-negative orthant of $\mathbb{R}^m$, denoted by $\mathbb{R}^m_{\geq 0}$, is the set of all elements in $\mathbb{R}^m$ whose coordinates are non-negative. Furthermore, the zero-dominated orthant, denoted by $\mathbb{R}^m_{\succ 0}$, is the set $\mathbb{R}^m_{\geq 0} \setminus \{0\}$ with 0 denoting the m dimensional vector $(0, \ldots, 0)$. Analogously, the non-positive orthant of $\mathbb{R}^m$, denoted by $\mathbb{R}_{\leq 0}$, is the set of elements in $\mathbb{R}^m$ whose coordinates are non-positive. Furthermore, the set of elements in $\mathbb{R}^m$ which dominate the zero vector 0, denoted by $\mathbb{R}^m_{\prec 0}$, is the set $\mathbb{R}^m_{\leq 0} \setminus \{0\}$. The set of positive reals is denoted by $\mathbb{R}_{>0}$ and the set of non-negative reals is denoted by $\mathbb{R}_{\geq 0}$.

The Pareto order can be associated to cone $\mathbb{R}^m_{\succ 0}$, i.e., the Pareto order $\prec_{Pareto}$ on $\mathbb{R}^m$ is given by the cone order with cone $\mathbb{R}^m_{\succ 0}$. This cone is also referred to as the Pareto cone.

Definition 2.16 (Pointed cone and convex cone)**.** A cone $\mathcal{C}$ is pointed, if and only if $\mathcal{C} \cap -\mathcal{C} \subseteq \{\mathbf{0}\}$ where $-\mathcal{C} = \{-c \mid c \in \mathcal{C}\}$ and $\mathcal{C}$ is convex if and only if $\forall c_1 \in \mathcal{C}, c_2 \in \mathcal{C}, \forall \alpha$ such that $0 \leq \alpha \leq 1 : \alpha c_1 + (1-\alpha)c_2 \in \mathcal{C}$.

As $\mathbb{R}^m_{\succ 0}$ is a pointed [1] and convex cone and $0 \notin \mathbb{R}^m_{\succ \mathbf{0}}$, the associated binary relation is irreflexive, antisymmetric and transitive, therefore strict partial order.

The following concepts are useful in order to compare order relations.

Definition 2.17 (Order Extension)**.** An order relation R' on the set X is said to extend an order relation R on the set X if $R' \supseteq R$. In other words, for all $x, x' \in X$: xRx' implies $xR'x'$.

Definition 2.18 (Minimal Element)**.** A minimal element $x \in X$ in a (strictly) partially ordered set (X, R) is an element for which there does not exist an $x' \in X$ with $x'Rx$ and $x' \neq x$. (In case, the order R is a strict partial order, $x'Rx$ implies $x' \neq x$).

Let (X, R) and (X, R') denote two strict partially ordered sets. If R' is an order extension of R, this implies:

1. The set of incomparable pairs in (X, R') is a subset of the set of incomparable pairs in (X, R).
2. The set of minimal elements of (X, R) is a superset of minimal elements of (X, R').

The first statement is true because if a pair of elements is incomparable in R', such a pair will also be incomparable in the smaller relation R. The second statement is also clear: as $R \subseteq R'$ a minimal element with respect to R' is also a minimal element with respect to R. In other words, if an element is non-dominated in R' it cannot be dominated in a smaller relation.

In the context of many objective optimization, extensions of the Pareto dominance order play an important role, since they on the one hand preserve the important and

[1] Different definitions of pointed cone are given in literature, here we use the definition by Matthias Ehrgott [34].

somewhat essential solutions, because one would not like to consider a Pareto dominated solution to be optimal; on the other hand, by extending Pareto dominance relation, the number of minimal (non-dominated) elements and the number of incomparable solutions can be reduced. Thereby, if wisely chosen, they can provide a (partial) remedy to the curse of dimensionality that occurs if the number of objective functions (m) increases.

The study of cone-based dominance goes back to the early work of Yu [130] and Miettinen relates it in her book to proper Pareto dominance and bounded trade-off [77]. It has also been related to equity preferences in the work of Shukla [96]. Next, we introduce several recently proposed alternative dominance relations which utilizing the extensions of the Pareto dominance order. They are also called relaxed forms or loose versions of Pareto dominance. It can be shown that they are also special cases of cone orders.

α-dominance

Ikeda et al. proposed α-dominance [60] to deal with dominance resistant solutions, i.e., solutions that are extremely inferior to other solutions in at least one objective, but hardly dominated in the other objectives. The idea behind α-dominance is that a small detriment in one or perhaps several of the objectives is permitted if an attractive improvement in the other objective(s) is achieved.

The α-dominance uses linear trade-off functions to define the tolerance of dominance. The approach is to define the following m functions on $\mathbb{R}^m \times \mathbb{R}^m$ with codomain $\mathbb{R}$ using m^2 a-priori given real numbers α_{ij} $(i, j \in \{1, \cdots, m\})$ as follows:

$$g_i(\mathbf{y}, \mathbf{y}') := \sum_{j=1}^{m} \alpha_{ij}(\mathbf{y}'_j - \mathbf{y}_j).$$

In [60], $\alpha_{ij} \geq 0$ and $\alpha_{ii} = 1$. For each such m-tuple of such functions, a strict partial order on $\mathbb{R}^m$ can be defined, denoted by $\overset{\alpha}{\prec}$ as follows:

$$\mathbf{y} \overset{\alpha}{\prec} \mathbf{y}' :\Leftrightarrow \forall i \in \{1, \cdots, m\} : g_i(\mathbf{y}, \mathbf{y}') \geq 0 \text{ and } \exists k \in \{1, \cdots, m\} : g_k(\mathbf{y}, \mathbf{y}') > 0.$$

ϵ-dominance

Laumanns et al. proposed the concept of ϵ-dominance [71].

Definition 2.19 (ϵ-dominance)**.** Let $\mathbf{y}$, $\mathbf{y}' \in \mathbb{R}^m$ and $\epsilon \in \mathbb{R}$, with $\epsilon > 0$, then $\mathbf{y}$ is said to ϵ-dominate $\mathbf{y}'$ (denoted by $\mathbf{y} \prec_\epsilon \mathbf{y}'$) if and only if $\forall i \in \{1, ..., m\} : y_i - \epsilon \leq y'_i$.

Optimizing based on ϵ-dominance creates, for objective functions with bounded ranges, a finite set of non-dominated solutions that will be placed distant from each other. If the decision maker wants to maintain a set of maximally T non-dominated solutions and let K denote a constant such that $0 \leq f_i \leq K, \forall i \in \{1, \ldots, m\}$, then ϵ can be adjusted to $\epsilon = (K/T)^{1/(m-1)}$. An alleged disadvantage of ϵ-dominance is that certain regions of the Pareto front with steep trade-off are underrepresented.

Batista et al. [6] proposed cone ϵ-dominance to improve the way solutions distribute as compared to ϵ-dominance. For this they propose to use a polyhedral cone given by a $\mathrm{R}^{m \times m}$ generator matrix and exemplify their approach in two dimensions, whereas some concepts were introduced also in higher dimensions.

Control Dominance Area of Solutions (CDAS)

Sato et al. proposed an approach to control the dominance area of solutions (CDAS) [90]. In CDAS, the objective values are modified and the i-th objective value of x after modification is defined as: $\hat{f}_i(x) = \frac{r \cdot \sin(w_i + S_i \cdot \pi)}{\sin(S_i \cdot \pi)}$, where r is the norm of $f(x)$, w_i is the declination angle between $f(x)$ and the coordinate axis. The degree of expansion or contraction of the dominance area of solutions can be controlled by the user-defined parameter S, i.e., $\hat{f}_i(x) > f_i(x)$ when $S_i < 0.5$; in case of $S_i = 0.5$, $f_i(x)$ does not change; and when $S_i > 0.5$, $\hat{f}_i(x) < f_i(x)$. Depending on increasing or decreasing the parameter S, the dominance area of solutions expands or increases. Only in case all the $S_i \leq 1$, CDAS is an extension of the Pareto relation.

Angle dominance

Liu et al. defined angle dominance [75]. For each point $\mathbf{y}$, a point dependent "cone" is constructed as follows. For each i ($i = 1, \cdots, m$), a point $P^{(i)} := (0, \cdots, 0, p_i, 0, \cdots 0)^\top \in \mathbb{R}^m$ is introduced, all coordinates of $P^{(i)}$ are zero, except the i-th coordinate. The i-th coordinate is derived from the worst point and a parameter $k > 0$. The worst point w is the point for which its i-th coordinate is equal to $w_i := \sup\{f_i(x) \,|\, x \in \mathcal{X}\}$, where f_i is the i-th objective ($i = 1, \cdots, m$), and $\mathcal{X}$ is the search space. Using the parameter k one defines $P^{(i)} := (0, \cdots, 0, p_i = kw_i, 0, \cdots, 0)^\top$. The second ingredient used is the ideal point (or if needed the utopian point). Denote the ideal point by $\mathbf{z}^{\text{ideal}}$. Then to a point $\mathbf{y}$, it associates m angles: $(\alpha_1, \cdots, \alpha_m)$. The cosine of α_i is equal to

$$\cos(\alpha_i) = \frac{(P^{(i)} - \mathbf{y}) \cdot (P^{(i)} - \mathbf{z}^{\text{ideal}})}{|P^{(i)} - \mathbf{y}| \, |P^{(i)} - \mathbf{z}^{\text{ideal}}|},$$

where in the numerator the inner product is used. The angle dominance relation is defined as follows.

Definition 2.20 (Angle Dominance)**.** $\mathbf{y} \prec_{\text{angle}} \mathbf{y}' :\Longleftrightarrow \forall i \in \{1, \cdots, m\} : \alpha_i \leq \alpha_i'$ and $\exists i \in \{1, \cdots, m\} : \alpha_i < \alpha_i'$, where α_i are the m angles associated to $\mathbf{y}$ and α_i' are m angles associated to $\mathbf{y}'$.

The authors show that given the premise that the parameter k is greater than 1, the angle dominance is irreflexive, asymmetric and transitive. Therefore, the angle dominance defines a strict partial order.

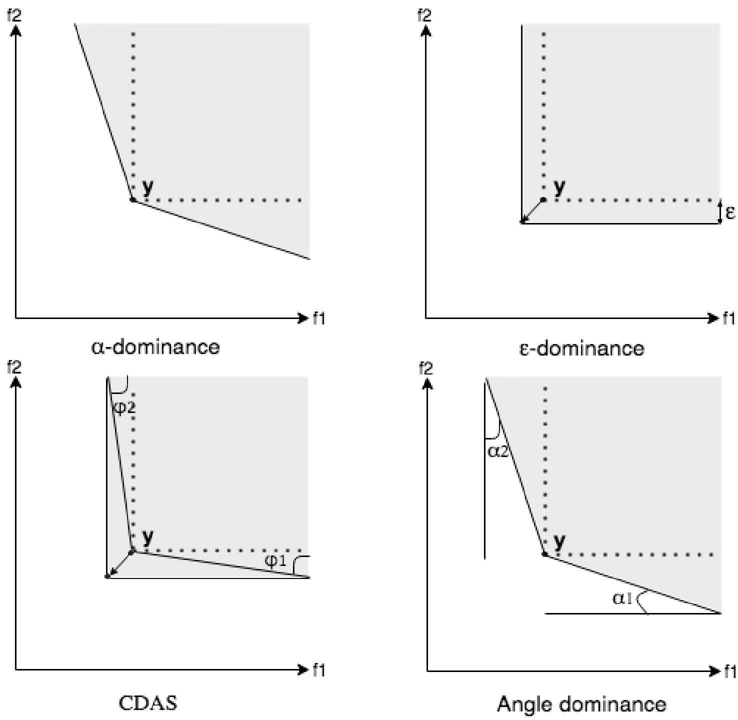

Figure 2.5: Dominance relations in a two-dimensional objective space.

Figure 2.5 illustrates these dominance relations in a two-dimensional objective space. The dotted lines indicate the space where solutions are Pareto dominated by the point $\mathbf{y}$. The gray areas show the dominated spaces by $\mathbf{y}$ based on α-dominance, ϵ-dominance, CDAS and angel dominance respectively. For the α-dominance (at the top left corner), the α-dominated area by $\mathbf{y}$ can be seen as expanding the angle of

Pareto cone and the degree of expansion is controlled by α. For ϵ-dominance (at the top right corner), the ϵ-dominated area by $\mathbf{y}$ can be seen as shifting the point $\mathbf{y}$ towards a position which Pareto dominates $\mathbf{y}$, therefore, the dominated area by $\mathbf{y}$ is expanded and the degree of expansion is controlled by ϵ. The dominated space by CDAS (at the bottom left corner) can also be seen as expanding the dominance area of $\mathbf{y}$ by translating Pareto cone. The degree of translation is decided by the parameter S because S controls φ ($\varphi_i = S_i \cdot \pi$). Angle dominance area is also obtained by expanding the angle of Pareto cone; the degree of expansion is decided by the parameter k and the worst point which controls α.

These dominance relations extend the Pareto dominance. Despite the difference in algorithm and controlling mechanisms, their eventual implementation is to allow a solution to dominate a larger space. At the same time, shifting the point to its dominating space which dominates it or opening the edges of its dominated space which is dominated by it, they are convertible. For instance, the dominated space of a point by CDAS can also be seen as opening the edges of Pareto order cone with a angle of $\pi/2 - S \cdot \pi$ because the opening angles are the same for each point. More details are available in [33]. In the next chapter, we use the geometrical interpretations directly to extend the Pareto dominance relation.