# Identifying Correlations in Understanding and Solving Many-Objective Optimisation Problems

T. Chugh A. Gaspar-Cunha A.H. Deutz J.A. Duro
D.C. Oara A. Rahat *

**Abstract**

Optimisation problems involving multiple objectives are commonly found in real-world applications. The existence of conflicting objectives produces trade-offs where a solution can be better with respect to one objective but requires a compromise in the other objectives. In many real-world problems the relationship between objectives is unknown or uncertain, and it is common to find problems with non-conflicting objectives. Understanding these relationships has been proven useful in different ways. The search efficiency of a multi-objective optimisation algorithm can benefit if objectives that are not essential to describe the Pareto-optimal front are omitted during the search procedure. Analysts and decision makers might get a better understanding about exiting synergies between the objectives, in turn facilitating the decision-making process of identifying the best solution. One particular useful technique to capture the relationships between objective functions is to rely on correlation measures. This chapter explores the literature of finding correlations among objective functions in solving multi-objective optimisation problems. Particularly, we focus on innovization and objective reduction approaches. We explain different statistical correlation measures and also provide details of benchmark and real-world optimisation problems solved by exploiting the correlations. This chapter provides an insight in solving multi-objective optimisation problems by considering the correlation among objective functions.

## 1 Introduction

Many real-world optimisation problems involve several conflicting objectives. These problems are termed as multi-objective optimisation problems. Several methods including *multi-objective evolutionary algorithms* (MOEAs) have been proposed and tested on numerous benchmark and real-world problems. Most of these methods do not consider correlations between objective

*The author names are arranged alphabetically.

functions. In many problems, it may happen that the objectives positively or negatively correlate with each other. For instance, [11] presented a geometry optimisation problem for a diffuser using *computational fluid dynamics* (CFD) simulations where the two objectives were anti-correlated to a degree due to the complex interactions between flow considerations and unusual shapes. These correlations between objective functions can be used as additional information when solving a given multi-objective problem. For instance, if two or more objectives are strongly correlated for any arbitrary decision variable, we may keep only one representative objective when optimising as knowing one of the objectives immediately permits us to deduce the others for any solution, and thus reduce the complexity of solving the overall problem.

It has been reported in the literature [29, 21] that the search ability of several MOEAs deteriorates when the number of objectives in a multi-objective optimisation problem increases beyond three objectives. Such problems are known as *many-objective optimisation problems* (MaOPs). For these problems, a particular class of MOEAs that relies on Pareto-dominance for selection and preservation is severely affected. This is because, in higher (i.e., four or more) dimensions, it is more likely that two solutions will be mutually non-dominated. As such this measure cannot effectively discriminate between solutions, and therefore MOEAs may stagnate due to the lack of selection pressure. Furthermore, it is difficult to visualise high-dimensional objective spaces [33]. While effective visualisation techniques exist in this domain, decision making is still challenging. Another more fundamental issue is the number of solutions required to represent the non-dominated set with the same coverage, since this number increases exponentially with the number of objectives. Hence, exploiting correlations between the objectives for MaOPs is a promising avenue to explore, especially from the *objective reduction* perspective.

In addition to tackling the challenges of MaOPs mentioned above, the knowledge of correlations between the objectives and decision variables can provide important insight into the nature of the problem at hand. In this context, the *innvovization* techniques that aim to find correlations between the decision variables and/or objective functions and exploit these in search have been successfully applied to real-world optimisation problems [4].

Given the importance of correlations, it is no wonder that several approaches have been proposed to reduce the number of objectives and provide insights into the problem. However, to the best of our knowledge, there is a lack of a concise review of these techniques in the literature.

In this work, we review methods and approaches which have been used to find correlations. The main contributions of this review are as follows:

1. We focus on data mining, objective reduction, and innovization methodologies using different types of correlation measures. We expect that

practitioners and decision-makers (DMs), who wish to develop a deep understanding of the problem at hand and select a solution with this knowledge, would find this review useful.

2. We provide details of benchmark and real-world optimisation problems that have been developed and used to test different methodologies in this domain. The review of the existing problems can provide insights on how the exploitation of correlations may be beneficial.

The rest of the chapter is organized as follows. In Section 2, we describe several correlation measures that are commonly used in the fields of applied sciences and numerical optimisation. The relationships between objectives, namely conflict and harmony, and how these can be captured by correlation measures is discussed in Section 3. Existing methodologies that exploit the use of correlations in fields such as data mining, *innovization* and objective reduction, are described in Section 4. In Section 5, we provide details of existing benchmark problems, and one real-world problem, which have been designed to exploit correlations either explicitly or implicitly.

# 2 Identifying correlations from data

Correlation is a statistical relationship that measures the degree to which a pair of parameters change in relation to each other. If a correlation is positive it implies that the values of both increase (or decrease) simultaneously. In the case of a negative correlation, the direction of change in one is opposite of the other.

The methods to quantify correlations from data are broadly used in applied sciences and numerical optimisation. This is primarily because they are a useful summary of the evidence or data, and may help identify and explore structural characteristics of a problem at hand.

It is important to appreciate that although correlation coefficients are useful indicators, misinterpretations and abuses of underlying assumptions may yield meaningless results [26]. Therefore, caution should be taken to ensure that a metric is appropriate within the context of the analysis.

In the rest of this section we discuss a range of popular correlation coefficients.

## 2.1 Pearson's correlation measure

One of the most widely used statistical correlation estimation measures is Pearson's correlation [27] for continuous random variables. Several variants of the Pearson's measure exist such as Pearson's correlation distance and weighted correlation coefficient.

Considering a pair of random variables $Y$ and $Z$, in this correlation measure, there are several assumptions made about the nature of these variables and their relationships [17]:

1. The two variables are correlated and continuous.
2. The relationship is linear.
3. The joint distribution $p(Y, Z)$ is Normal.
4. A pair of samples $(y_i, z_i) \sim p(Y, Z)$ is collected through independent random sampling.

Now, if we have observed two vectors $\mathbf{y} = (y_1, \ldots, y_n)^\top$ and $\mathbf{z} = (z_1, \ldots, z_n)^\top$, we can measure the Pearson's correlation for these observations as follows:

$$\rho_p(\mathbf{y}, \mathbf{z}) = \frac{\text{cov}(\mathbf{y}, \mathbf{z})}{\sigma_y \sigma_z} = \frac{\sum_{i=1}^{n} (z_i - \bar{z})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (z_i - \bar{z})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}, \tag{1}$$

where $\text{cov}(\mathbf{y}, \mathbf{z})$ is the sample covariance for $\mathbf{y}$ and $\mathbf{z}$, $\bar{y}$ and $\bar{z}$ are means of the sampled vectors $\mathbf{y}$ and $\mathbf{z}$ respectively, and $\sigma_y$ and $\sigma_z$ are their standard deviations.

Here, $\text{cov}(\mathbf{y}, \mathbf{z}) = \sum_{i=1}^{n} (z_i - \bar{z})(y_i - \bar{y})$ is positive when, on average, the signs $(z_i - \bar{z})$ and $(y_i - \bar{y})$ agree more frequently with higher magnitude. This is further normalised by the standard deviations $\sigma_y = (\sum_{i=1}^{n} (y_i - \bar{y})^2)^{\frac{1}{2}}$ and $\sigma_z = (\sum_{i=1}^{n} (z_i - \bar{z})^2)^{\frac{1}{2}}$. In other words, this measure is the product of average distances from the mean of the samples, each normalised by the associated standard deviations. As such this measure describes how two vectors of samples are correlated with each other.

The value of Pearson's correlation measure can vary in the range $[-1, 1]$, with $-1$ representing perfect negative correlation, and 1 representing perfect positive correlation.

### 2.2 Spearman's correlation measure

To deal with ordinal variables, i.e. discrete or categorical variables with a clear order or ranking between choices, the Spearman's or Kendall's correlation measure can be used. First, we discuss how Spearman's measure [32] work. This is a ranked correlation measured developed from Pearson's correlation (see [24] for a discussion on the topic). To apply this measure, the samples are ranked (or sorted) in ascending order, first by $y_i$, and then by $z_i$. This produces two ranking vectors: $\mathbf{r}^y$ and $\mathbf{r}^z$. The correlation measure is then computed based on these vectors:

$$\rho_s(\mathbf{y}, \mathbf{z}) = \frac{\text{cov}(\mathbf{r}^y, \mathbf{r}^z)}{\sigma_{r^y}, \sigma_{r^z}}. \tag{2}$$

Essentially, by extracting the correlation coefficient from the rankings, we avoid relying on Euclidean distances (as done in Pearson's measure). This is because we cannot assign such distances between ordinal variables.

If we assume that no intra-sample ties exists: $r_i^y \neq r_j^y$ and $r_i^z \neq r_j^z$, $\forall i, j \in [1, n]$, then the measure can be computed as follows:

$$\rho_s(\mathbf{y}, \mathbf{z}) = 1 - \frac{6 \sum_{i=1}^{n} (r_i^y - r_i^z)^2}{n(n^2 - 1)}, \tag{3}$$

where $n$ is the number of observations. This is a robust measure of correlations, and works well even if a few intra-sample ties exist.

The value of Spearman's correlation measure can vary in the range $[-1, 1]$, with $-1$ representing perfect negative correlation, and 1 representing perfect positive correlation.

### 2.3 Kendall's correlation measure

Unlike the Pearson's and Spearman's correlation measures which are variance based approaches, the Kendall's correlation measure [23] is a probability based approach. Given a paired observation vectors $\mathbf{y}$ and $\mathbf{z}$ of size $n$ with the associated ranking vectors $\mathbf{r}^y$ and $\mathbf{r}^z$, we can determine how many times these rankings are concordant. We define a pair of ranks $(r_i^y, r_i^z)$ and $(r_j^y, r_j^z)$ as *concordant*, when $r_i^y < r_j^y$ then $r_i^z < r_j^z$, and when $r_i^y > r_j^y$ then $r_i^z > r_j^z$ for any $i < j$. In other words, two rankings are concordant if the sort order agrees in a pair of ranks. If these rank relationships are violated, we call them *discordant*. With the number of concordant pairs $N_c$ and the number of discordant pairs $N_d$, the Kendall's rank correlation measure is defined as:

$$\rho_k(\mathbf{y}, \mathbf{z}) = \frac{N_c - N_d}{\binom{n}{2}} = \frac{N_c - N_d}{\frac{n(n-1)}{2}}, \tag{4}$$

where the denominator represents the binomial coefficient is the number of ways to choose 2 rankings form a set of $n$ paired observations [1].

This is an estimation of the difference between the frequencies of pairs being concordant and discordant. If the frequencies were the same, we get a correlation of 0. If there are no concordant pairs, i.e. $N_c = 0$, then the correlation is perfectly negative $-1$. For perfect positive correlation, we thus require $N_d = 0$.

The original formulation does not consider ties, i.e. $r_i^y = r_j^y$ and $r_i^z = r_j^z$ for any $i < j$. Agresti, in [2], modified the Kendall correlation to account for tied ranks as:

$$\rho_k(\mathbf{y}, \mathbf{z}) = \frac{N_c - N_d}{\sqrt{(N_c + N_d + T_y)(N_c + N_d + T_z)}}, \tag{5}$$

where $T_y$ and $T_z$ are the number of tied ranks in for the observations $\mathbf{y}$ and $\mathbf{z}$, respectively.

### 2.4 Goodman & Kruskal's correlation measure

Goodman and Kruskal introduced a third rank correlation measure in [16] for problems where the number of tied ranks are small and can be ignored. The measure is defined as:

$$\rho_{gk}(\mathbf{y}, \mathbf{z}) = \frac{N_c - N_d}{N_c + N_d}. \tag{6}$$

This measure is directly considering the relative frequency difference between concordant and discordant pairs.

All the above ranking measures can be used for continuous variables as well: basically we would construct a rank vector for each observation and compute the measures accordingly.

### 2.5 Cramér's correlation measure

Cramér introduced a correlation measure for nominal variables, i.e. categorical variables with no natural order between the choices, in [9]. In this case, $Y$ and $Z$ can have multiple categories. Let, the set of possible categories for $Y$ be $K$, and for $Z$ be $L$. The number of possible categories are $\eta_y = |K|$ and $\eta_z = |L|$. The association between $\mathbf{y}$ and $\mathbf{z}$ is thus:

$$\rho_c(\mathbf{y}, \mathbf{z}) = \sqrt{\frac{\chi/n}{min(\eta_y, \eta_z) - 1}}, \tag{7}$$

where $\chi$ is the Pearson's chi-squared statistic as follows:

$$\chi^2 = \sum_{k \in K} \sum_{l \in L} \frac{(N_{kl} - \bar{N}_{kl})^2}{\bar{N}_{kl}}, \tag{8}$$

where $N_{kl}$ is the sample frequency of observing the pair $(y_i = k \in K, z_i = l \in L)$. $\bar{N}_{kl}$ is the expected frequency, i.e. $\bar{N}_{kl} = N_{k\star} N_{\star l}/n$, where $N_{k\star}$ and $N_{\star l}$ are the number of samples with $y = k$ and $z = l$ respectively.

Cramér's correlation measure reduces to $\phi$ coefficient, $\phi = \chi/\sqrt{N}$, when at least one of the variables becomes a binary variable (i.e $\eta_y = 2$ or $\eta_z = 2$). The correlation for the nominal variables is usually measured between $[0, 1]$. This is because a lack of natural order (or ordinal relationship) between categories means that there is no directional change in correlations: observations are either correlated to a degree or not.

### 2.6 Nonlinear correlation information entropy (NCIE)

The metrics discussed thus far are effective for identifying linear correlations. Even though they are often used on non-linearly correlated data, they are not as insightful. An alternative is nonlinear correlation information entropy for rankings, first proposed by Wang *et al.* in [38], and it can capture both linear and nonlinear correlations.

The main idea is based on mutual information entropy. Given two discrete random variables $Y$ and $Z$, each with domains of $K$ and $L$ possibilities respectively, it can be defined as:

$$I(Y, Z) = H(Y) + H(Z) - H(Y, Z), \tag{9}$$

where $H(Y)$ and $H(Z)$ are the information entropy measures for $Y$ and $Z$, and $H(Y, Z)$ is the joint entropy between $Y$ and $Z$. These terms are defined as follows:

$$\begin{aligned} H(Y) &= -\sum_{i=1}^{K} p_i \ln(p_i), H(Z) = -\sum_{j=1}^{L} p_j \ln(p_j), \\ H(Y, Z) &= -\sum_{i=1}^{K}\sum_{j=1}^{L} p_{ij} \ln(p_{ij}). \end{aligned} \tag{10}$$

Here, $p_i$ is the probability of observing the $i$th option for $Y$ (and $p_j$ has the same interpretation for $Z$), and $p_{ij}$ is the probability of observing both the $i$th option for $Y$ and the $j$th option for $Z$ together.

The primary issue with the above is that the span of mutual information is not guaranteed to be within the range $[0, 1]$. Therefore, Wang *et al.* proposed the following modifications [38]: let, the paired observation vectors $\mathbf{y}$ and $\mathbf{z}$ of size $n$, and their associated rankings $\mathbf{r}^y$ and $\mathbf{r}^z$. Then, they divide the ranks into $b = \lceil\sqrt{n}\rceil$ rank grids, i.e. put the first set of $b$ top ranked individuals into *first* rank, the second set of $b$ top ranked become members of *second* rank, and so on. Now, they created a $b \times b$ rank grid, and place every pair $\{y_i, z_i\}_{i\in[1,n]}$ in these grids by comparing them to the rank sequences $\mathbf{r}^y$ and $\mathbf{r}^z$. With this, the modified mutual information, also known as NCIE, is defined as follows:

$$\rho_{NCIE}(\mathbf{y}, \mathbf{z}) = I_{NCIE}(\mathbf{y}, \mathbf{z}) = 2 + \sum_{i=1}^{b}\sum_{j=1}^{b} \frac{n_{ij}}{n} \ln\left(\frac{n_{ij}}{n}\right), \tag{11}$$

where, $n_{ij}$ is the count of the number of samples in the $(i, j)$ cell of the grid.

The nonlinear correlation metric $\rho_{NCIE}$ varies between $[0, 1]$, where 0 indicates minimum correlation and 1 indicates maximum correlation. Since, $\rho_{NCIE}$ does not indicate a direction, Wang *et al.* [37] considered the product [sign $(\text{cov}(\mathbf{y}, \mathbf{z})) \times \rho_{NCIE}(\mathbf{y}, \mathbf{z})$] and thus derived a directional version of the metric.

## 3 Conflict and harmony between objectives

We have so far described correlation measures for measuring the degree by which two variables relate to one another, from a statistical point of view.

In this section other relationships between objectives that are commonly found in the multi-objective optimisation literature are described, and we investigate how these relate to the correlation measures.

Two well studied relationships between objectives are conflict and harmony [28, 15]. It is common to consider all objectives in a multi-objective optimisation problem to be in conflict, and in case they are all equally important, then multiple trade-off optimal solutions are likely to exist. On the other hand, two objectives can be in harmony, meaning that an improvement in one objective leads also to an improvement in the other. These two dependency relationships are not mutually exclusive, and could change across the Pareto front.

## 3.1 Definitions and metrics of conflict and harmony

Consider an hypothetical multi-objective optimisation problem with $M$ objective functions, denoted by $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x}))^\top$, that have to be simultaneously minimised. The set of all realisable objective vectors is $\mathcal{Z}$ and one element of this set is $\mathbf{z} = (z_1, \ldots, z_M)^\top$ where $\mathbf{z} \in \mathbb{R}^M$. Let $\mathcal{Z}_R$ be a particular region of interest in objective space with cardinality $|\mathcal{Z}_R| = N$ that may have been obtained by an MOEA such that $\mathcal{Z}_R \subseteq \mathcal{Z}$.

Purshouse and Fleming [28] have categorised different relationships between objectives as either dependent or independent. In the dependent case the relationships could be further categorised as either in conflict or in harmony. They have also proposed several definitions with respect to the aforementioned relationships, and we focus here in just two of them. Consider two objective vectors $\mathbf{z}^a, \mathbf{z}^b \in \mathcal{Z}_{\mathcal{R}}$ where $a, b \in \{1, \ldots, N\}$ and let $(a, b)$ denote a pair of instances such that $a \neq b$, then the definition are as follows:

**Definition 3.1** (Conflict)**.** There is evidence of conflict between objectives $i$ and $j$ if the following condition is satisfied $(z_i^a < z_i^b) \wedge (z_j^a > z_j^b)$. In case $\nexists(a, b)$ such that the condition does not hold then there is no conflict. In case $\exists(a, b)$ then there is conflict, and there is total conflict if the condition holds true $\forall(a, b)$.

**Definition 3.2** (Harmony)**.** Levels of harmony are determined by the condition $(z_i^a < z_i^b) \wedge (z_j^a < z_j^b)$. In case $\nexists(a, b)$ such that the condition does not hold then there is no harmony. In case $\exists(a, b)$ then there is harmony, and there is total harmony if the condition holds true $\forall(a, b)$.

De Freitas et al. [15] have proposed metrics to quantify the relationships of conflict and harmony. Depending how the objectives are normalised, the conflict metrics are known as direct, max-min or non-parametric, and the metric for harmony is inversely proportional to non-parametric conflict.

Let the objective vector corresponding to the $i$th objective be denoted by $\dot{\mathbf{z}}_i \in \mathbb{R}^N$, and let its maximum and minimum be denoted by $u_{\dot{\mathbf{z}}_i}$ and

$l_{\dot{\mathbf{z}}_i}$, respectively. The solutions are ranked[1] (sorted in ascending order) and let $\mathbf{r}_i \in \mathbb{N}^N$ be a vector of ranks corresponding to the $i$th objective. To determine the conflict between two objectives each objective vector needs to be normalised, and for the $i$th objective let the normalised vector be denoted by $\ddot{\mathbf{z}}_i \in \mathbb{R}^N$. The measure of conflict between objectives $i$ and $j$ is given by

$$C_{ij} = \|\ddot{\mathbf{z}}_i - \ddot{\mathbf{z}}_j\|, \tag{12}$$

where $\|\bullet\|$ is the absolute norm, and let the maximum value of the measure be denoted by $c_{\max}$. Depending on how the objective vectors are normalised we can have the following types of conflict:

1. Direct: absolute difference between two objectives, which is only suitable if both objectives have the same units (or their values lie in the same range). The normalised objective vectors are given by $\ddot{\mathbf{z}}_i = \dot{\mathbf{z}}_i - l_{\dot{\mathbf{z}}_i}$ and $c_{\max}$ is not bounded.

2. Max-min: same as in direct but objectives values are normalised in the range between 0 and 1. The normalised objective vectors are given by $\ddot{\mathbf{z}}_i = \frac{\dot{\mathbf{z}}_i - l_{\dot{\mathbf{z}}_i}}{u_{\dot{\mathbf{z}}_i} - l_{\dot{\mathbf{z}}_i}}$ and $c_{\max} = N$.

3. Non-parametric: this metric operates on ranks between solutions with respect to each objective and is equivalent to measuring the degree to which lines cross in a parallel coordinates plot. The normalised objective vectors are replaced by their ranks as given by $\ddot{\mathbf{z}}_i = \mathbf{r}_i$ and $c_{\max} = \sum_i^N |2i - N - 1|$.

The concept of harmony is not necessarily the opposite of conflict, but harmony is inversely proportional to non-parametric conflict. Complete harmony happens when all solutions between two objectives have exactly the same values. Based on non-parametric conflict, a measure of global harmony that returns values that range from 0 (lowest level of harmony) and 1 (high level of harmony) is given by

$$H_{ij} = 1 - C_{ij}/c_{\max}. \tag{13}$$

### 3.2 Comparing conflict and harmony with correlation measures

Consider Figure 1, which shows four solution sets for a 2-objective problem, and each solution set contains $N = 30$ solutions. The performance of the solutions are shown on scatter plots in Figures 1a, 1b, 1c, 1d and also on parallel coordinate plots in Figures 1e, 1f, 1g, 1h. Based on Definitions 3.1

[1]In the determination of the ranks, the ties are resolved by using the rank that causes least conflict with the other objective being compared.

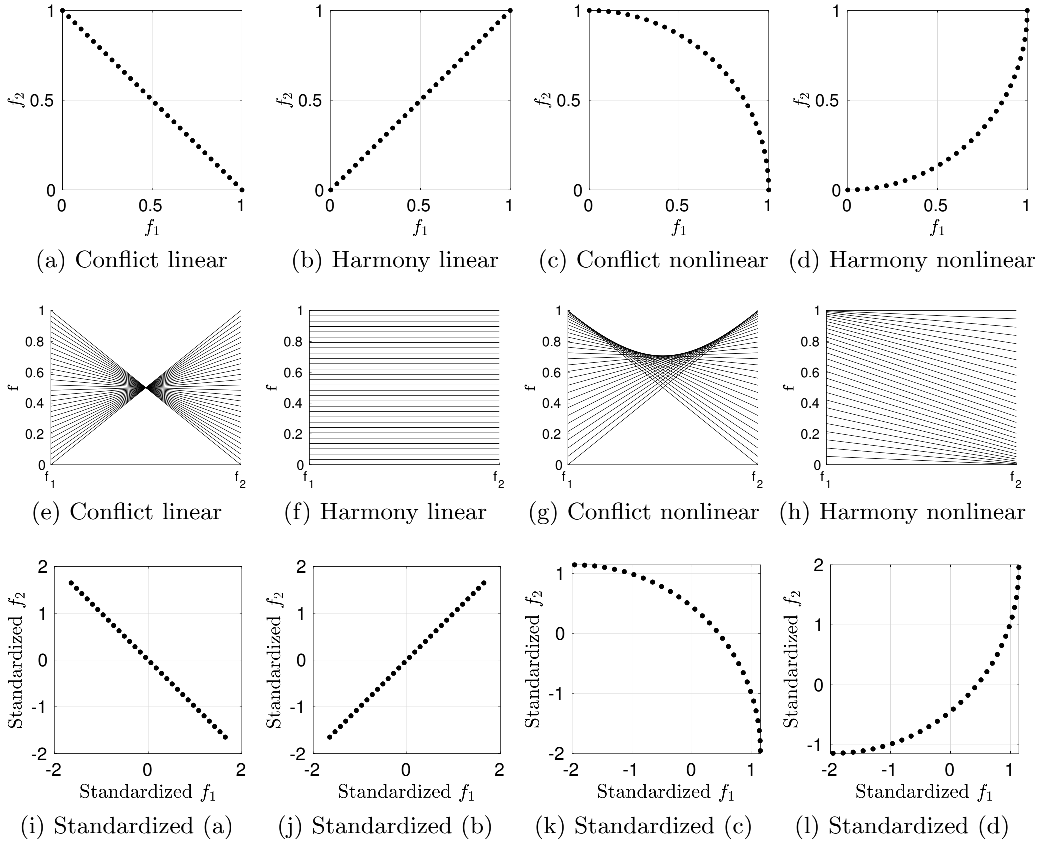
and 3.2, $f_1$ and $f_2$ are in total conflict (Figures 1a and 1c), or are in total harmony (Figures 1b and 1d). In some cases the relationship between $f_1$ and $f_2$ is linear (Figures 1a and 1b), and in other cases it is nonlinear (Figures 1c and 1d). Our intention here is to analyse the effect that these solution sets have on the Pearson and Kendall correlation measures, and as well on the non-parametric metrics of conflict (Equation 12) and harmony (Equation 13). For the Pearson correlation, the solutions are standardized following the process explained below, and for the Kendall correlation measure we simply make use of the parallel coordinate plots, in that, discordance produces crossing lines (implying conflict) whilst concordance does not (implying harmony) (see Section 2.3 for more details).

To determine the Pearson correlation, two objective vectors $\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2 \in \mathbb{R}^N$ corresponding to $f_1$ and $f_2$, respectively, first need to be standardized. This means that each vector is mean centered and needs to have unit standard deviation, and let $\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2 \in \mathbb{R}^N$ denote the standardized version of $\dot{\mathbf{z}}_1$ and $\dot{\mathbf{z}}_2$, respectively. The Pearson correlation coefficient can be determined by $\frac{1}{N-1}\bar{\mathbf{z}}_1^\top \bar{\mathbf{z}}_2$, which is equivalent to Equation 1. Consider the following observations:

1. Conflict linear: in Figure 1i the points are symmetrical with respect to the origin, and when one objective is negative the other is positive, and vice-versa. This means that the Pearson correlation will be negative, and the coefficient is $\rho_p(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = -1$. In Figure 1e all lines are crossing, implying that the Kendall correlation coefficient is $\rho_k(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = -1$. The non-parametric measure of conflict gives $C_{1,2}/c_{\max} = 1$ and the measure of harmony gives $H_{1,2} = 0$. The correlation measures indicate that $f_1$ and $f_2$ are negatively correlated with maximum correlation strength, and the non-parametric measures indicate conflict.

2. Harmony linear: in Figure 1j the points are symmetrical with respect to the origin, and both objectives are either positive or negative. This means that the Pearson correlation will be positive, and the coefficient is $\rho_p(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = 1$. In Figure 1f all lines are not crossing, implying that the Kendall correlation coefficient is $\rho_k(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = 1$. The non-parametric measure of conflict gives $C_{1,2}/c_{\max} = 0$ and the measure of harmony gives $H_{1,2} = 1$. The correlation measures indicate that $f_1$ and $f_2$ are positively correlated with maximum correlation strength, and the non-parametric measures indicate harmony.

3. Conflict nonlinear: in Figure 1k the points are not symmetrical with respect to the origin, and the obtained correlation coefficient is $\rho_p(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = -0.9129$. In Figure 1g all lines are crossing, implying that the Kendall correlation coefficient is $\rho_k(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = -1$. The non-parametric measure of conflict gives $C_{1,2}/c_{\max} = 0$ and the measure of harmony gives $H_{1,2} = 1$. Although both correlation measures indicate that $f_1$ and




(a) Conflict linear (b) Harmony linear (c) Conflict nonlinear (d) Harmony nonlinear

(e) Conflict linear (f) Harmony linear (g) Conflict nonlinear (h) Harmony nonlinear

(i) Standardized (a) (j) Standardized (b) (k) Standardized (c) (l) Standardized (d)

Figure 1: Solution sets for a two-objective problem showing cases of total conflict and total harmony.

$f_2$ are negatively correlated, the correlation strength obtained by the Pearson correlation measure is lower than the Kendall measure. The non-parametric measures indicate conflict.

4. Harmony nonlinear: in Figure 1l the points are not symmetrical with respect to the origin, and the obtained correlation coefficient is $\rho_p(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = 0.9129$. In Figure 1h all lines are not crossing, implying that the Kendall correlation coefficient is $\rho_k(\dot{\mathbf{z}}_1, \dot{\mathbf{z}}_2) = 1$. The non-parametric measure of conflict gives $C_{1,2}/c_{\max} = 1$ and the measure of harmony gives $H_{1,2} = 0$. Although both correlation measures indicate that $f_1$ and $f_2$ are positively correlated, the correlation strength obtained by the Pearson correlation measure is lower than the Kendall measure. The non-parametric measures indicate harmony.

As shown by the above cases, both Pearson and Kendall correlation measures are able to provide some indication about the relationship type between two pairs of objectives. We have considered cases where the pairwise relationships between objectives is either in total conflict or in total harmony. However, if the relationship is not linear then the Pearson correlation may not be the most appropriate statistical measure, and the Kendall correlation might be more suitable. We have also demonstrated for the same cases, that the Kendall correlation measure is equivalent to the non-parametric measures by De Freitas et al. [15], since both approaches are sensitive to the existence of crossing lines in a parallel coordinate plot.

# 4 Exploiting correlations

Several techniques and approaches have been proposed and used to exploit the correlations between objectives and/or decision variables. These methods are generally used for data mining, innovization and objective reduction.

## 4.1 Data mining

Bandaru et al. [5] detailed several data mining methods that can be used to extract knowledge from the given data in solving multi-objective optimisation problems. These methods are summarised in Table 1. These include: descriptive statistics, visual data mining, and machine learning. Moreover, the same authors have presented some examples of using such techniques on multi-objective data sets.

In another study, Chiba et al. [7] used three data mining techniques: self-organizing maps, functional analysis of variance and rough set theory and applied them to solve an aerodynamic shape design optimisation problem with multiple objectives. The problem had four objectives and 71 decision variables related to the wing shape of a two-stage-to-orbit reusable vehicle. The study showed that the self-organizing maps were able to provide the correlation of different variables on the objective functions. On the other hand, analysis of variance and rough set theory were able to tailored down the number of decision variables.

## 4.2 Innovization

Solving real-world multi-objective optimisation problems not only aims to find an approximation of Pareto optimal solutions, but also to decide in favour of a single solution to be used in practice. However, in some cases, e.g. dynamic changes in the importance of objectives during the search can pose a challenge in decision-making. In such a scenario the objectives to be optimised can be changed based on their relevance during the search process. Therefore, it is important to exploit the existing correlations between objectives and/or decision variables. The knowledge of such correlations can be further used in the optimisation algorithm when solving a given problem.

Table 1: Methods for data mining.

| Descriptive Statistics | | | |
|---|---|---|---|
| Central Tendency | Variability | Distribution type | Correlation |
| Mean | Standard Deviation | Skewness | Pearson |
| Median | Quartiles | Kurtosis | Spearman |
| | | | Kendall |
| | | | Cramér |

| Visual Data Mining | | |
|---|---|---|
| Graphical | Clustering | Manifold Learning |
| Heat maps | Biclustering | Sammon Mapping |
| Pareto Race | Hierarchical | Isomaps |
| Prosection Method | k-Means Clustering | Self-Organizing Maps |
| Level Diagrams | Kernel based or Density based | Proper Orthogonal decomposition |

| Machine Learning | | |
|---|---|---|
| Supervised | Unsupervised | Beyond Learning |
| Decision Trees | Rough Set Theory | Randomization |
| Neural Networks | Biclustering | Ensemble Learning |
| Radial Basis Function | Pattern Mining | Multi-instance learning |
| Bayesian Networks | Automated Inovization | Semisupervised Learning |

In real-world optimisation problems, a DM or an expert is usually involved in formulating the problem and also to learn and gain insights. However, understanding a high-dimensional problem (both in objective and decision spaces) can result in high cognitive load. Moreover, high-dimensional problems raise the challenge of visualisation of solutions.

Addressing these challenges, Bandaru and Deb [4] proposed a framework for automatic innovization. The main idea behind this methodology is to use sophisticated data analysis methods, based on data-mining and/or machine learning techniques, to identify important relationships between decision variables and objectives. We illustrate a schematic of the innovization methodology in Figure 2. The problem under study is described by some mathematical models that can be implemented in a computer; it can be a simple equation. An optimisation algorithm, in the present case an MOEA, is then able to determine an approximation to the Pareto-optimal front (POF) by taking into consideration the relationships between the decision variables and the objectives. At the end, the DM, using data-mining and/or machine learning techniques, is able to find correlations between all parameters involved in the system.

The efficacy of the framework is based on the ability to describe a particular relationship (or a design rule) with a product of $N$ basis functions. Here, each basis function combines any scalar function of the decision variables, objectives and constraints. For instance, a relationship for a vector $\mathbf{z}$ (objective and/or decision vector) can be expressed as:

$$\text{constant} = \prod_{j=1}^{N} \phi(\mathbf{z})_j^{a_{ij} b_{ij}}, \tag{14}$$

where $a_{ij}$ is a Boolean indicator that reflects the presence ($a_{ij} = 1$) or absence ($a_{ij} = 0$) of the $j$th basis function, $b_{ij}$ represents the powers of contributions of the $j$th basis function towards the $i$th design rule, $\phi()_j$ represents the $j$th basis function. The following equation is an example of such relationship, where, for the problem under study, one of such relationships can be defined as:

$$\text{constant} = x_1^{2.1} x_2^{-1} f_1^{0.5} \tag{15}$$

With this, we can gather data on the estimated POF following several runs of an MOEA. Then it is straightforward to analyse this data and identify significant relationships between parameters of interest. This insight into various relationships between parameters may then be exploited based on statistical ranking.

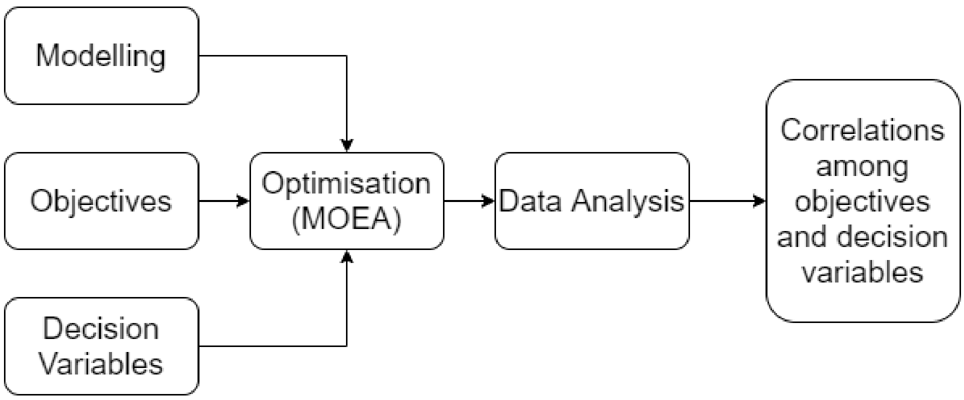



Figure 2: A schematic of an innovization method in solving a multi-objective optimisation problem

### 4.3 Objective reduction

Objective reduction approaches have been exploited to counter the limitations of existing MOEAs in dealing with MaOPs. These approaches attempt to eliminate objectives that are not essential to describe the POF. If the number of objectives is reduced to three or less, it is often the case that MOEAs are well suited for solving such type of problems. In case the number of objectives remains higher than three, it is still expected for any reduction to improve the search efficiency of an optimiser, to reduce the number of solutions required to provide a fair coverage across the POF, and to ease the decision-making process (i.e. to reduce the cognitive load associated with the process of selecting a solution).

Existing approaches operate on the objective vectors of the non-dominated solutions that describe the POF, and the aim is to identify the smallest set of conflicting objectives that generates the same POF as the original problem formulation. When two objectives are not in conflict, these two objectives share the same optimal solution, implying that there are no other trade-off solutions where a gain in one objective leads to a sacrifice in the other. Therefore, from each set of non-conflicting objectives all objectives but one can be omitted without affecting the POF. The smallest set of conflicting objectives is often termed as *essential*, while all objectives that can be omitted without affecting the POF are termed as *redundant*. To identify the essential objective set, existing objective reduction approaches rely on the analysis of the Pareto-dominance structure, or on the interpretation of the correlations between objectives.

In this section we review existing objective reduction approaches, covering those that rely on dominance (Section 4.3.1) and correlation structure preservation (Section 4.3.2). We focus on how conflicting objectives are identified and which optimisation problems have these approaches been applied to. Also, for those approaches that preserve the correlation structure, we also discuss which correlation measures are employed, and the criteria used to indicates whether two objectives are correlated or not.

### 4.3.1 Approaches that preserve the dominance structure

Brockhoff and Zitzler [6] have proposed an objective reduction approach known as Dominance Relation Preservation (DRP) and as the name suggests, it relies on the concept of dominance. Objectives are considered to be redundant, as long as the dominance structure is preserved in case they are omitted. The concept of dominance is often used by multi-objective optimisation algorithms where two solutions are compared on whether one dominates the other. For an arbitrary multi-objective problem with $M$ objectives consider the subset $\mathcal{F}' \subseteq \mathcal{F}$ where the set $\mathcal{F} = \{f_1, \ldots, f_M\}$ includes all objectives. Let $\mathbf{z}^a, \mathbf{z}^b \in \mathbb{R}^M$ be two objective vectors, $\mathbf{z}^a$ is said to dominate $\mathbf{z}^b$ if the following two conditions are satisfied[2] (assuming minimisation):

1. $\mathbf{z}^a$ is no worse than $\mathbf{z}^b$ in all objectives (i.e. $z_i^a \leq z_i^b \; \forall i = 1, \ldots, M$); and

[2]This definition of dominance relationship between two solutions is sometimes referred to as weak dominance relation.

2. $\mathbf{z}^a$ is strictly better than $\mathbf{z}^b$ in at least one objective (i.e. $\exists i \in \{1, \ldots, M\}$ s.t. $z_i^a < z_i^b$).

If any of the two conditions are violated, then it is said that $\mathbf{z}^a$ does not dominate $\mathbf{z}^b$, and if either solution does not dominate the other then they are said to be non-dominated. The effect of omitting objectives in the dominance relation can be studied by considering only those objectives in the subset $\mathcal{F}'$. Two objective subsets $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{F}$ are said to be non-conflicting if the same solutions remain dominated (or non-dominated), otherwise they are said to be in conflict. By applying this relation to all possible objective subsets, it is possible to determine the smallest objective subset that preserves the dominance structure of the original objective set. The authors have also proposed an error measure and introduced the concepts of $\delta$-Minimum Objective Subset ($\delta$-MOSS) and Minimum Objective Subset of Size $k$ with minimim error ($k$-EMOSS), in that:

1. $\delta$-MOSS corresponds to a situation where a DM alows for a $\delta$ error, and wants to know the smallest set of conflicting objectives, such that the error incurred is less than or equal to $\delta$;

2. $k$-EMOSS corresponds to a situation where a DM specifies a fraction of the original number of objectives to be retained, and wishes to know which objective set incurs the minimal error.

The $\delta$ error is zero when the dominance relations between $\mathcal{F}$ and $\mathcal{F}'$ are identical, and in case it is desirable to further reduce the number of objectives, then it indicates by how much the objective values have to be adjusted by an additive term $\delta$ such that the corresponding dominance relations are identical. DRP has been demonstrated by the authors on several benchmark problems, including DTLZ2, DTLZ5, DTLZ7 ([13]) (see Section 5.2.1 for a description), and a multi-objective variant of the knapsack problem. It has also been applied to a real-world radar waveform problem ([19]) (more details in Section 5.2.2).

Singh et al. [31] have proposed the Pareto Corner Search Evolutionary Algorithm (PCSEA) to search for the corners of the POF, to which they apply an objective reduction approach. The motivation is that generating a representative set of the whole POF is difficult if the number of objectives is too high, and that the boundaries of the POF (which are less difficult to find) should aid in the accurate estimation of the true dimensionality of the POF. The criterion used by this approach relies on the interpretation of the ratio between the number of non-dominated solutions found in a reduced objective set, and the number of non-dominated solutions found in entire objective set. However, this requires the user to specify a threshold that is used to indicate whether or not an objective is redundant, and the order on which the objectives are omitted could lead to different essential objective sets. It is also required to conduct a search over the problem model for the corner solutions (in this case by using a multi-objective evolutionary algorithm) with a solution comparison criterion that promotes the corner solutions. This differs from the other objective reduction approaches that apply directly to the non-dominated solutions generated by an optimizer. PCSEA has been demonstrated by the authors on several benchmark problems, in particular on an extended version of the scalable DTLZ5 benchmark problem known as DTLZ5($I$,$M$) ([12]), where the Pareto front is $I$-dimensional ($I < M$) (more details in Section 5.2.1). It has also been applied to DTLZ2, two variants of WFG3 ([18]), and two engineering design problems, a storm drainage system problem ([25]) and the above radar waveform problem.

Yuan et al. [39] have proposed two objective reduction approaches based on dominance relation preservation. The first relies on a bi-objective formulation where the number of objectives and the error incurred due to the reduction in the number of objectives, are posed as objectives. The second is based on the work by [31] since it makes use of the ratio between the number of non-dominated solutions found in a reduced objective set, and the number found when entire objective set is used. These approaches have been validated on DTLZ2, DTLZ5($I$,$M$) and WFG3.

### 4.3.2 Approaches that preserve the correlation structure

The objective reduction approach proposed by Jaimes et al. [22] relies on the Pearson correlation (Section 2.1) to identify conflicting objectives. The authors have proposed a modified correlation coefficient given by $1 - \rho_p$ which takes values in the interval $[0, 2]$. Thus, a value of 2 indicates that the objectives are negatively correlated, and a value of 0 indicates that the objectives positively correlated. The authors have not defined a criterion (possibly by using a threshold) that would indicate whether or not two objectives are said to be correlated or not, based on the correlation coefficient. Instead, the proposed approach quantifies the error incurred in case an objective is to be omitted. The error quantification approach relies on a clustering technique, where the objectives are grouped together into neighbourhoods of fixed sizes based on their distance, measured by the correlation coefficient. When an objective is omitted, the error corresponds to the biggest distance in that neighbourhood. This approach has been demonstrated on DTLZ5($I$,$M$).

Saxena et al. [30] proposed a machine learning based approach that makes use of principal components analysis (PCA) and maximum variance unfolding (MVU) techniques, leading to a linear and nonlinear algorithm. The linear and nonlinear algorithms are know as L-PCA and NL-MVU-PCA, respectively. The premise is that the essential objective set can be identified by transforming the high-dimensional data such that the correlation structure is preserved, in a process where redundancies (or dependencies) are minimised. PCA removes the second-order dependencies (variance and covariance) by performing eigenvalue decomposition on a correlation matrix determined with respect to the objective vectors, revealing the principal components. MVU removes higher order dependencies by learning a kernel matrix to which PCA can be applied to, as opposed to the correlation matrix. The process is analogous to unfolding a high-dimensional data manifold by maximizing the Euclidean distance between points, while locally preserving distances and angles between nearby points. The correlation between objectives is interpreted as follows. The first step is to construct a correlation matrix by using the Pearson correlation coefficient applied to all pairwise objective vectors. Then:

1. Along each principal component, the objectives are selected as follows. The objective with the highest contribution by magnitude, and all objectives with opposite sign are picked. If all objectives have the same sign, then the two objectives with the top two contributions by magnitude are picked. All objectives not picked at this stage are interpreted as being non-conflicting.

2. The non-conflicting objectives identified in the previous step are omitted from the correlation matrix. Based on this reduced correlation matrix, a subset of correlated objectives is identified for each objective. This is to say that each objective is individually compared with the others, and any two objectives are said to be correlated if:

    (a) their correlation strength is equal to or higher than a correlation threshold, and;

    (b) if their correlation signs with respect to the other objectives (including the non-conflicting ones identified in the previous step) are the same.

    A closed form expression exists to determine the correlation threshold. This is based on the interpretation of the redundancy in a problem, in that, a problem is said to be highly redundant if the first principal component accounts for the majority of the variance. In contrast, a problem is said to have low redundancy if the variance is equally distributed amongst the principal components.

Based on the information provided by PCA, the authors have also derived an error measure that can indicate how much error is incurred if the redundant objectives are to be omitted. This error measure is different from the DRP approach since it quantifies the variance that is left unaccounted if an objective is omitted. Both L-PCA and NL-MVU-PCA have been validated on several benchmark problems, including DTLZ1 to DTLZ4, DTLZ5($I$,$M$), and WFG3. The error measure has been operationalised to provide the $\delta$-MOSS and $k$-EMOSS analysis in [14], and also to generated a preference ranking between the objectives. The authors have demonstrated their approach on the above radar waveform problem.

Wang and Yao [37] have proposed to use a correlation measure based on the information theory concept of mutual information, known as NCIE (Section 2.6). This quantifies the nonlinear correlation between two variables, providing a score between 0 and 1, where 0 indicates no correlation and 1 is total correlation. To indicate whether the correlation is positive or negative, the NCIE is combined with the covariance. The score of this modified NCIE takes values between -1 and 1, where -1 indicates total negative correlation, and 1 is total positive correlation. This approach starts by constructing a correlation matrix, where the modified NCIE is applied to each possible pair of objective vectors. The next step is to select the most conflicting objective, which is the objective with the largest absolute sum of its negative NCIEs to other objectives. Other objectives are interpreted as being not essential, if they are positively correlated with the most conflicting objective, and their strength of correlation is higher than a pre-defined threshold. This approach has been validated on DTLZ5($I$,$M$) and WFG3.

Yuan et al. [39], besides the two dominance relation preservation approaches mentioned above, they have also proposed another approach that relies on the preservation of the correlation structure. This approach is inspired by [22] method, in that, instead of using the Pearson's correlation coefficient, the Kendall's rank correlation (Section 2.3) is used instead. The authors claim that this measure is sensitive to both linear and nonlinear relationships. This approach has been validated on DTLZ2, DTLZ5($I$,$M$) and WFG3.

Following the above review, we summarise in Table 2 the correlation measures used by the objective reduction approaches, and the criteria used to identify the essential objective set are summarised in Table 3.

Table 2: Correlation measures used by objective reduction approaches

| Reference | Correlation measures |
|---|---|
| Jaimes et al. [22] | Pearson (Section 2.1) |
| Saxena et al. [30] | Pearson (Section 2.1) |
| Wang and Yao [37] | NCIE (Section 2.6) |
| Yuan et al. [39] | Kendall (Section 2.3) |

# 5 Benchmarking and case studies

In the literature, only few studies exist in developing multi-objective optimisation benchmark problems with correlated objectives. These problems can be broadly divided into two groups: one where we can *explicitly* impose

Table 3: Criteria used by objective reduction approaches to identify the essential objective set

| Reference | Criterion |
|---|---|
| Brockhoff and Zitzler [6] | The smallest objective subset that preserves the dominance structure. |
| Jaimes et al. [22] | The smallest objective subset that incurs in the lowest error tolerable by the user. The error corresponds to the highest distance between objectives inside a neighbourhood as determined by the correlation coefficient. |
| Singh et al. [31] | Ratio between the number of non-dominated solutions found in a reduced set and the number found in the original set, combined with a user-defined threshold. |
| Saxena et al. [30] | Interpretation of the principal components (first reduction), and interpretation of the correlation matrix combined with a data-derived threshold (second reduction). |
| Wang and Yao [37] | Interpretation of the correlation matrix, involving a user-defined threshold. |
| Yuan et al. [39] | Two dominance structure based criteria that involve the interpretation of the dominance structure and the ratio between solutions as defined in [31]. Also, a criterion based on the correlation between objectives that relies on the interpretation of Kendall's rank correlation. |

a desired level of correlation, and other where the control of correlation is *implicit*. We here review articles which focused on using and/or developing such problems.

## 5.1 Explicit correlation

### 5.1.1 One-Max Problem

In [8], a continuous version of mapped One-Max problem with two objectives for binary search spaces was introduced. The problem provided the flexibility to control the correlation among objectives. The problem was defined as:

$$f = (f_1, f_2) = (n_1(\mathbf{x}), n_2(\mathbf{y})), \tag{16}$$

where $n_1(\mathbf{x})$ is the sum of all decision variables values in decision vector $\mathbf{x}$, $n_2(\mathbf{y})$ the sum of all variable values in $\mathbf{y}$ and $\mathbf{y}$ is the mapped version of $\mathbf{x}$ i.e. $y_i = |x_i - map_i|, i = 1, \ldots, n$, ($n$ is number of decision variables). The mapped value $map_i \in [0, 1]$ is independent for each element in $\mathbf{x}$ and was set by flipping a coin biased by the degree of correlation $corr \in [-1, 1]$. The $corr$ parameter can be defined to introduce a desired correlation among objectives. For instance, $corr = 0$ means no correlation, $corr = -1$ negative correlation and $corr = 1$ means a positive correlation.

### 5.1.2 Multi-objective NK-Landscape Combinatorial Problems

The NK landscape methodology is a performance modelling approach for a general class of systems; a concise introduction on the topic can be found in [10].

The modelling paradigm considers a class of systems that has $N$ components, each with a number of discrete possibilities (usually restricted to two: 0 or 1) and a dependence on $K$ other components (where $0 \leq K \leq N - 1$). A complete state vector $\mathbf{x} = (x_1, \ldots, x_N)^\top$ consists of individual states. The NK landscape is essentially a function $f : \mathbf{x} \in \mathbb{N}^N \rightarrow \mathbb{R}$ that maps the state vector into a performance measure (or commonly referred to as fitness). This can be defined as:

$$f(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} y_i(\mathbf{x}), \tag{17}$$

where $y_i(\cdot)$ is a function accounting for the contribution of the $i$th component of $\mathbf{x}$ to the performance, and $x_i$ and $K$ other components of $\mathbf{x}$ are active for $y_i(\mathbf{x})$. Here, $K$ may be used to control the inter-dependencies between components, and usually a higher $K$ results in more complex and ragged performance landscapes. This construct allows the design and analysis of successful algorithms for this class of problems in a controlled environment, and it is particularly important as many real-world problems can be modelled with this paradigm.

It is relatively straightforward to extend this concept to MaOPs. A canonical multi-objective NK (MNK) landscape problem with $M$ objectives can be defined as follows:

$$\max_{\mathbf{x}} \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_M(\mathbf{x}))^\top, \tag{18}$$

where the $j$th objective function $f_j$ is an NK landscape performance function with a potentially distinct $K_j$. Here, the complexity is controlled with $\mathbf{K} = (K_1, \ldots, K_M)^\top$. This permits design and analysis of multi-objective algorithms. Note that the component functions may now be denoted as $y_{ij}(\cdot)$

for the $i$th component and the $j$th objective function. Interested readers should refer to the work of [3] for many key insights into the workings of successful algorithms in this domain.

In this paper, we are interested in MaOPs with correlated objectives. Extending the ideas of MNK problems, [34] have proposed CMNK and $\rho$MNK test problems where the correlation between objectives may be precisely controlled.

In CMNK, the central concept is to define a correlation structure between the component functions $y_{ij}$ for all $i \in [1, N]$ and $j \in [1, M]$. In particular, the pairwise correlation is defined as:

$$C_{de} = \text{corr}(y_{id}, y_{ie}) \in [-1, 1], \tag{19}$$

for all $i \in [1, N]$. By definition, if $d = e$ then $C_{de} = 1$. Naturally there is a restriction of symmetry: $C_{de} = C_{ed}$. This means the correlation is as per a $M \times M$ symmetric matrix $C$. They require $C$ to be positive definite. Note that the standard MNK problem can be easily derived from CMNK by setting $C_{de} = C_{ed} = 0$ when $d \neq e$.

In addition, [34] propose a restricted CMNK problem known as $\rho$MNK problems. This is a special case when $C_{de} = C_{ed} = \rho$ given that $d \neq e$. To maintain the positive definiteness of the matrix, they suggest that $\rho > \frac{-1}{M-1}$. Therefore, it may not be possible to investigate all positive and negative correlations in the range of $[-1, 1]$ beyond two objectives.

It should be noted that the correlation structure is defined on the component functions $y_{ij}$ rather than between the objectives. However, they provided a proof that shows that the expected correlation between a pair of objectives is the same as the correlation defined in the relevant cell of the correlation matrix $C$: $\mathbb{E}[\text{corr}(f_d(\mathbf{x}), f_e(\mathbf{x}))] = C_{de}$. Thus, it is possible to precisely control the correlations between the objectives in this class of test problems.

Clearly, the correlation structure here is global, as in the correlation between a pair of objectives is always constant. This suite of test problems is therefore a restrictive set, but certainly allows us to explore different characteristics of the interactions between solvers and landscapes.

For a detailed treatment of the nature of these problems, and successful algorithms, please refer to [34, 35, 36].

## 5.2 Implicit correlation

### 5.2.1 DTLZ and WFG problems

DTLZ and WFG problems have been widely used in the literature to test several multiobjective optimisation algorithms. They have also been used to validate objective reduction approaches (see Section 4.3), especially DTLZ2, DTLZ5, DTLZ5($I$,$M$) and WFG3. These problems are scalable in terms of the number of objectives and also decision variables.

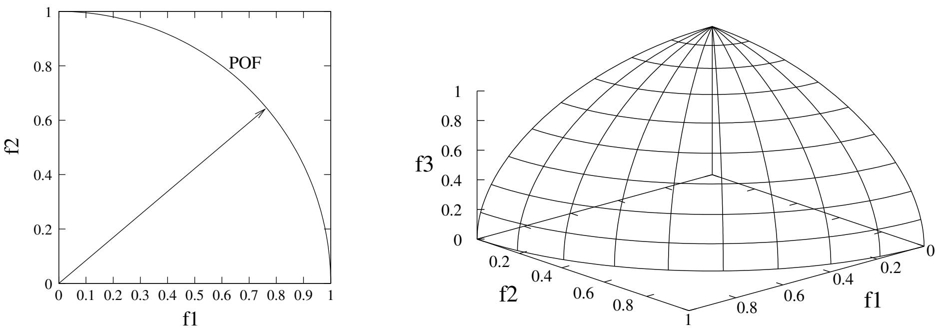
Consider the formulation for DTLZ2 in Equation 20 where the total number of decision variables is given by $N = M+k-1$, and $k$ is a parameter that can be adjusted by the user to control the number of decision variables. For POF all decision variables in the $g$ function (i.e. $x_M$ to $x_{M+k-1}$) are fixed to 0.5, implying that $g = 0$, and the other decision variables (i.e. $x_1$ to $x_{M-1}$) can take any value within their bounds. The latter set of variables that are unconstrained in the POF are referred to as free variables.

$$
\begin{aligned}
\text{Min. } & f_1 = (1+g)\Pi_{i=1}^{M-1}\cos(\theta_i)\\
\text{Min. } & f_{j=2:M-1} = (1+g)\Pi_{i=1}^{M-j}\cos(\theta_i)\sin(\theta_{M-j+1})\\
\text{Min. } & f_M = (1+g)\sin(\theta_i)\\
\text{s.t. } & 0 \le x_i \le 1, \text{ for } i = 1,\dots,M,\\
& \theta_i = \frac{\pi}{2}x_i, \text{ for } i = 1,\dots,M-1,\\
\text{and } & g = \sum_{i=M}^{M+k-1}(x_i-0.5)^2.
\end{aligned}
\tag{20}
$$

The number of free variables is given by $M-1$, implying that for $M = 2$ there is one free variable ($\theta_1$), for $M = 3$ there are two free variables ($\theta_1$ and $\theta_2$), and so on. For $M = 2$ this implies that the POF is a one-dimensional curve since there is only one degree of freedom provided by $\theta_1$, as pointed out by the arrow in Figure 3a. For $M = 3$ the POF is a two-dimensional surface where two degrees of freedom are provided by $\theta_1$ and $\theta_2$, one samples across the horizontal plane and the other across the vertical one as depicted in Figure 3b. For higher dimensions the POF is a hypersurface. Hence, two conflicting objectives implies one free variable and a one-dimensional POF; three conflicting objectives implies two free variables and a two-dimensional POF; and this generalises to $m \le M$ conflicting objectives which implies $m-1$ free variables and a $(m-1)$-dimensional POF. In DTLZ2 $m = M$ and all objectives are equally conflicting, implying that they are not positively correlated. Other DTLZ problems, such as DTLZ1, DTLZ3 and DTLZ4 also share the same property as pointed out in [30].

In DTLZ5 and DTLZ6, the problem can be reduced to only two conflicting objectives: $f_{M-1}$ and $f_M$. The rest of the objectives are positively correlated (i.e. no conflict) with $f_{M-1}$. Similarly, the POF for WFG3 degenerated into a linear hyperplane such that the first $M-1$ objectives are perfectly correlated, while the last objective is in conflict with all other objectives. Although WFG3 has been proposed as a degenerate test problem by [18], the Pareto front is actually not degenerated as argued by [20].

DTLZ5($I$,$M$) is an extended version of DTLZ5. First described in [12], this test problem has been used as a benchmark in the objective reduction



(a) DTLZ2 with 2 objectives (b) DTLZ2 with 3 objectives

Figure 3: Representation of the POF for the DTLZ2 problem.

literature given its ability to control the dimensionality of the POF by the parameter $I$, such that $I \leq M$. This is achieved by restricting the number of free variables and for this consider the problem formulation in Equation 21. For POF the variables in the $g$ function are fixed to 0.5, implying that $g = 0$ and $\theta_{i=I:M-1} = \frac{\pi}{4}$. The number of free variables is $I - 1$ implying that dimensionality of the POF is $I - 1$. As a result, the first $M - I + 1$ objectives are perfectly correlated, while the rest are in conflict with every other objective in the problem.

$$
\begin{aligned}
\text{Min. } & f_1 = (1+g)\Pi_{i=1}^{M-1}\cos(\theta_i) \\
\text{Min. } & f_{j=2:M-1} = (1+g)\Pi_{i=1}^{M-j}\cos(\theta_i)\sin(\theta_{M-j+1}) \\
\text{Min. } & f_M = (1+g)\sin(\theta_1) \\
\text{s.t. } & 0 \leq x_i \leq 1, \text{ for } i = 1, \ldots, M \\
& g = \sum_{i=M}^{M+k-1} (x_i - 0.5)^2, \\
& \theta_{i=1:I-1} = \frac{\pi}{2}x_i, \ \theta_{i=I:M-1} = \frac{\pi}{4(1+g)}(1+2gx_i), \\
& c_i = \sum_{j=0}^{I-2} f_{M-j}^2 + 2^{p_i} f_i^2 \gg 1, \text{ for } i = 1, \ldots, M-I+1, \\
& p_1 = M-1 \text{ and } p_{i=2:M-I+1} = (M-I+2)-2.
\end{aligned}
\tag{21}
$$

### 5.2.2 Radar waveform problem

The radar waveform problem, first described in [19], deals with the design a of waveform for a *Pulsed Doppler Radar*. This type of radars are typically used to equip an aircraft, and the primary aim is to locate and track other aircrafts during an air-to-air role. For this, the radar needs to determine the range and velocity, and to take into account that the target aircraft may travel at very high velocity (Mach 5 possible) and its location could be more than 100 nautical miles away. The optimisation problem has a total of 9 objective functions, the first 8 are to be minimised, and the last one is to be maximised. The physical meaning of each objective is as follows:

1. Median range/velocity extent of target before schedule is not decodable ($f_1/f_2$);
2. Median range/velocity extent of target before schedule has blind regions ($f_3/f_4$);
3. Minimum range/velocity extent of target before schedule is not decodable ($f_5/f_6$);
4. Minimum range/velocity extent of target before schedule has blind regions ($f_7/f_8$);
5. Time required to transmit total waveform ($f_9$).

This real-world problem has been used as a benchmark problem, in particular to validate objective reduction approaches. This is due to the fact that the author of this problem in [19] has revealed the expected correlations between objectives, and also which objectives should be in conflict. The following objective pairs that measure the range $f_1$ & $f_3$ and $f_5$ & $f_7$ are expected to be correlated. The same can be said about the following objective pairs that measure the velocity, that is, $f_2$ & $f_4$ and $f_6$ & $f_8$. The objectives that measure range and velocity are expected to be in conflict.

# 6 Summary

In this paper we have conducted a review focusing on approaches that employ knowledge extraction techniques (such as correlation measures) as a way to facilitate the process of solving a multi-objective optimisation problem (MOP). These approaches can be found in fields such as data mining, objective reduction and innovization. The knowledge is extracted from solutions that have been generated by an MOEA due to their population-approach, since multiple optimal solutions can be found in a single optimisation run. The rationale for these approaches lies in the fact that developing a model for an optimisation problem requires a lot of domain expertise, but knowledge about the problem (e.g. relationships between the objectives and decision variables) may not be available or simply cannot be treated as trivial, specially when dealing with models that are very complex. As these approaches have demonstrated in this review, the knowledge extracted from an MOP can be used to:

1. Reduce the number of objectives which can counter the limitations of Pareto-based MOEAs in generating a good approximation of the POF, and facilitate the decision-making process;

2. Determine a rank between the decision variables based on some criterion which can be used to reduce the dimensionality of the decision space. This also can be used to facilitate the search and decision-making processes.

In this review we have first described the correlation measures that are broadly used in the fields of applied sciences and numerical optimisation. We have shown that these are useful for indicating if two objectives are either in conflict or in harmony. However, the presence of nonlinearity can affect the accuracy of some of these measures (e.g. the Pearson correlation), but others have shown to be more robust (e.g. the Kendall correlation).

Some correlation measures have been used in data mining alongside other methods (e.g. central tendency and variance statistics, or even machine learning approaches such as rough set theory). The same principles are adopted in innovization where the focus is to provide a better understanding of the problem to a designer or a practitioner, revealing information that can be useful to the task of finding the most desirable solution by the DM. In objective reduction many approaches rely on correlation measures to determine which objectives can be eliminated without affecting the POF, while others rely on the interpretation of the Pareto-dominance structure. In particular, the criterion used by [30] and [37] can be used to indicate whether or not two objectives are correlated by comparing their correlation strength against some threshold. Other information that can be of interest to a DM is how much error one incurs if one or more objectives are to be omitted, as in [6, 22, 30]. Thereupon the information provided by an error measure can be used to conduct the $\delta$-MOSS and $k$-EMOSS analysis, and even to derive a preference ranking between the objectives.

Besides the above approaches, this review has covered benchmarking and case studies focusing on correlated objectives. Depending on how the correlations are perceived, the problems have been categorised as either as explicit or implicit. In the explicit case, it is possible to specify a desirable degree of correlation between the objectives. This could be by tweaking a parameter (e.g. One-Max problem) or by defining a correlation structure (e.g. CMNK and $\rho$MNK). In the implicit case, the correlations between objectives cannot be prescribed (say by a user), but it is known which objectives are supposed to be correlated. The latter type of problems have been extensively used in the literature to validate objective reduction algorithms.

# References


[1] H. Abdi. The kendall rank correlation coefficient. *Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, CA*, pages 508–510, 2007.

[2] A. Agresti. *Analysis of ordinal categorical data*, volume 656. John Wiley & Sons, 2010.

[3] H. E. Aguirre and K. Tanaka. Working principles, behavior, and performance of moeas on mnk-landscapes. *European Journal of Operational Research*, 181(3):1670 – 1690, 2007.

[4] S. Bandaru and K. Deb. Automated innovization for simultaneous discovery of multiple rules in bi-objective problems. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 1–15. Springer, 2011.

[5] S. Bandaru, A. H. Ng, and K. Deb. Data mining methods for knowledge discovery in multi-objective optimization: Part a-survey. *Expert Systems with Applications*, 70:139–159, 2017.

[6] D. Brockhoff and E. Zitzler. Improving hypervolume-based multiobjective evolutionary algorithms by using objective reduction methods. *2007 IEEE Congr. Evol. Comput. CEC 2007*, pages 2086–2093, 2007.

[7] K. Chiba, S. Jeong, Shinkyu Obayashi, and K. Nakahashi. Knowledge discovery in aerodynamic design space for flyback–booster wing using data mining. pages 1–18, 2006.

[8] T. Chugh, R. Allmendinger, V. Ojalehto, and K. Miettinen. Surrogate-assisted evolutionary biobjective optimization for objectives with non-uniform latencies. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, pages 609–616, 2018.

[9] H. Cramér. *Mathematical methods of statistics*, volume 43. Princeton university press, 1999.

[10] F. A. Csaszar. A note on how nk landscapes work. *Journal of Organization Design*, 7(1):15, 2018.

[11] S. Daniels, A. Rahat, G. Tabor, J. Fieldsend, and R. Everson. Automated shape optimisation of a plane asymmetric diffuser using combined computational fluid dynamic simulations and multi-objective bayesian methodology. *International Journal of Computational Fluid Dynamics*, pages 1–16, 2019.

[12] K. Deb and D. Saxena. Searching for pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. In *IEEE Congress on Evolutionary Computation (CEC)*, pages 3353–3360. IEEE, 07 2006.

[13] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. *Scalable Test Problems for Evolutionary Multiobjective Optimization*, pages 105–145. Springer London, London, 2005.

[14] J. A. Duro, D. K. Saxena, K. Deb, and Q. Zhang. Machine learning based decision support for many-objective optimization problems. *Neurocomputing*, 146:30 – 47, 2014. Bridging Machine learning and Evolutionary Computation (BMLEC) Computational Collective Intelligence.

[15] A. R. R. D. Freitas, P. J. Fleming, and F. G. Guimarães. Aggregation trees for visualization and dimension reduction in many-objective optimization. *Information Sciences*, 298:288–314, 2015.

[16] L. A. Goodman and W. H. Kruskal. Measures of association for cross classifications. *Journal of the American Statistical Association*, 146(268):732–764, 1954.

[17] L. L. Havlicek and N. L. Peterson. Robustness of the pearson correlation against violations of assumptions. *Perceptual and Motor Skills*, 43(3_suppl):1319–1334, 1976.

[18] S. Huband, L. Barone, L. While, and P. Hingston. A scalable multi-objective test problem toolkit. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization*, pages 280–295, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[19] E. J. Hughes. Radar waveform optimisation as a many-objective application benchmark. *Lect. Notes Comput. Sci.*, 4403 LNCS:700–714, 2007.

[20] H. Ishibuchi, H. Masuda, , and Y. Nojima. Pareto fronts of many-objective degenerate test problems. *IEEE Trans. on Evolutionary Computation*, 2016.

[21] H. Ishibuchi, N. Tsukamoto, and Y. Nojima. Evolutionary many-objective optimization: A short review. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, pages 2419–2426, June 2008.

[22] A. L. Jaimes, C. A. Coello Coello, and D. Chakraborty. Objective reduction using a feature selection technique. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation, GECCO '08*, pages 673–680, 2008.

[23] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.

[24] M. G. Kendall. *Rank correlation methods.* Griffin, 1948.

[25] K. Musselman and J. Talavage. A tradeoff cut approach to multiple objective optimization. *Operations Research*, 28(6):1424–1435, 1980.

[26] A. J. Onwuegbuzie and L. G. Daniel. Uses and misuses of the correlation coefficient. 1999.

[27] K. Pearson. Notes on regression and inheritance in the case of two parents proceedings of the royal society of london, 58, 240-242, 1895.

[28] R. C. Purshouse and P. J. Fleming. Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimisation. In C. M. Fonseca, P. J. Fleming, E. Zitzler, L. Thiele, and K. Deb, editors, *Evolutionary Multi-Criterion Optimization*, pages 16–30, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[29] R. C. Purshouse and P. J. Fleming. On the Evolutionary Optimization of Many Conflicting Objectives. *IEEE Transactions on Evolutionary Computation*, 11(6):770–784, Dec 2007.

[30] D. K. Saxena, J. A. Duro, A. Tiwari, K. Deb, and Q. Zhang. Objective Reduction in Many-Objective Optimization: Linear and Nonlinear Algorithms. *IEEE Transactions on Evolutionary Computation*, 17(1):77–99, 2013.

[31] H. K. Singh, A. Isaacs, and T. Ray. A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, 15(4):539–556, 2011.

[32] C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 1904.

[33] A. K. A. Talukder and K. Deb. Paletteviz: A visualization method for functional understanding of high-dimensional pareto-optimal data-sets to aid multi-criteria decision making. *IEEE Computational Intelligence Magazine*, 15(2):36–48, 2020.

[34] S. Verel, A. Liefooghe, L. Jourdan, and C. Dhaenens. Analyzing the effect of objective correlation on the efficient set of mnk-landscapes. In *International Conference on Learning and Intelligent Optimization*, pages 116–130. Springer, 2011.

[35] S. Verel, A. Liefooghe, L. Jourdan, and C. Dhaenens. Pareto local optima of multiobjective nk-landscapes with correlated objectives. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, pages 226–237. Springer, 2011.

[36] S. Verel, A. Liefooghe, L. Jourdan, and C. Dhaenens. On the structure of multiobjective combinatorial search space: Mnk-landscapes with correlated objectives. *European Journal of Operational Research*, 227(2):331 – 342, 2013.

[37] H. Wang and X. Yao. Objective reduction based on nonlinear correlation information entropy. *Soft Computing*, 20(6):2393–2407, 2016.

[38] Q. Wang, Y. Shen, and J. Q. Zhang. A nonlinear correlation measure for multivariable data set. *Physica D: Nonlinear Phenomena*, 200(3):287 – 295, 2005.

[39] Y. Yuan, Y. S. Ong, A. Gupta, and H. Xu. Objective Reduction in Many-Objective Optimization: Evolutionary Multiobjective Approaches and Comprehensive Analysis. *IEEE Trans. Evol. Comput.*, 22(2):189–210, 2018.