Capturing Relationships in Multi-Objective Optimization

Gian Fritsche*, Andrei Strickler*, Aurora Pozo* and Roberto Santana†
*Computer Science's Department, Federal University of Paraná
Curitiba, Paraná, Brazil
Email: {gfritsche,astrickler,aurora}@inf.ufpr.br
†Department of Computer Science and Artificial Intelligence
University of the Basque Country, San Sebastin-Donostia 20080, Spain
Email: roberto.santana@ehu.es

***Abstract*—When applied to multi-objective problems (MOPs), evolutionary algorithms (EAs) can be noticeably improved by representing and exploiting information about the interactions between the components of the problem (variables and objectives). However, accurate detection of such relationships is a challenging question that involves other related issues such as finding the right metric for measuring the interaction, deciding about the timing for testing the interactions, and deciding on appropriate ways to represent the relationships found. In this paper we investigate the performance of three correlation measures (Kendall, Spearman and Pearson) in the context of multi-objective optimization using the MOEA/D-DRA algorithm. We analyze the accuracy of the measures at different stages of the evolution and for different types of relationships. Moreover, the paper proposes a meaningful way for visualizing and interpreting the captured interactions.**

***Index Terms*—Variables and objectives relationships; Correlation measures**

I. Introduction

Several real world problems are multi-objective, it means that they have to optimize more than one criteria at once, usually conflicting. A variety of algorithms have been proposed and applied to solve hard benchmark and real-world MOPs [1]. A number of works have suggested to explicitly capture the relationships between the components of the MOP (variables and objectives) as a way to improve the efficiency of the algorithms. A common approach is to investigate the correlations between the objectives of the MOP [2], [3]. Removing redundant objectives using information about the correlations can be an alternative to the challenge of many-objective problems [4]. Another kind of methods that capture and exploit some of the relationships between the MOP's components are multi-objective estimation of distribution algorithms (MOEDAs) [5]. However, most of MOEDAs, as is the case with single-objective EDAs, are focused on learning and using the dependencies between the variables of the problem. They do not explicitly represent the relationships between variables and objectives.

Recent developments on the question of exploiting the relationships between the components of the MOPs include the introduction of multidimensional Bayesian networks to represent dependencies between objectives, between variables, and between objectives and variables [6]. This is a very general framework to represent relationships. However, the method applied to learn the relationships between the components can be inaccurate, particularly for local relationships. This is so because the algorithm used for searching the best global structure is based on a greedy search method, and the global scoring metric is derived from the covariance between the components of the MOP.

Another recent proposal [7] is based on learning a mapping from the objective space to the decision space. The idea is that such a mapping could be used to create the offspring by sampling the objective space. Although very efficient, this approach has some limitations: the models map single objectives to single decision variables and direct dependencies between the variables are not captured.

Our work is also related with research on *innovization* [8], a research direction that aims to extract important design principles hidden in Pareto set approximations found by the MOEAs. Although innovization searches for a meaningful representation of the relationships that exist in the optimal solutions, this representation is not necessarily in terms of direct dependencies between variables and objectives. Furthermore, the analysis is carried out after a set of optimal solutions has been found and not from intermediate steps of the search as we do.

In this paper we focus on a required building block for any method aiming at exploiting the relationship between the components of an MOP. We evaluate the accuracy of different measures to capture the interactions between the components of the problem. This is a difficult issue that transcends the simple analysis of the data because the generation of solutions in MOEAs is dynamically determined by the search process. Therefore, the relationships between variables and objectives are influenced by the function, by the selection process, and by the replacement methods applied in MOEAs to guarantee diversity. A consequence of this is that the relationships between the components of the problem can continuously change along the evolution. Determining the right time for obtaining an accurate estimation of the interactions is itself a problem.

To address the problem, we start by defining an experimen-

978-1-5090-0016-6/15 $31.00 © 2015 IEEE
DOI 10.1109/BRACIS.2015.25

tal framework determined by a function for which the interactions between the components is known a priori. We use the multiobjective evolutionary algorithm based on decomposition with Dynamic Resource Allocation (MOEA/D-DRA) [9] to evaluate the accuracy of three measures (Kendall, Spearman and Pearson) and analyze the behavior of the interactions along evolution. The main contributions presented in the paper are the following: 1) We investigate the efficiency of different correlation measures for the task of capturing the relationships between the components of MOPs. 2) We analyze the dynamics of the relationships between the components of a MOP in the context of the MOEA/D-DRA algorithm. 3) We propose a method for visualizing and interpreting the relationships between the components of the MOP.

II. Background

Our goal is to infer how the problem components (decision variables and objectives) are related, i.e., if the value of one depends on the value of the others. This can be done by evaluating if the statistical metrics applied to the data generated by the MOEAs are able to capture these relationships. The long-term goal is to be able to use the captured relationships, improving the optimization along the evolutionary multi-objective process.

The analysis starts from some known representation of the relationships between the components of the MOP (variables and objectives). This problem is optimized using a MOEA and the samples generated by the algorithm during the search process are stored for later analysis. In this work we use the Multiobjective Evolutionary Algorithm based on Decomposition (MOEA/D-DRA) [9]. It is a state-of-the-art evolutionary algorithm for multiobjective optimization using the decomposition idea, with Dynamic Resource Allocation (DRA). This algorithm was winner of the CEC 2009 MOEA contest. Any other MOEA could be used.

In the next step, the correlation measures are computed using these samples and an *approximate* structure of the relationships between the problem components is recovered. In the final step, the quality of this approximate structure is evaluated using different criteria. In the following, we will explain correlation measures used and the MOP selected.

A. Correlation measures

The correlation between variables and objectives is extracted by statistical metrics. There are three correlation metrics that we used in this work: Kendall, Spearman and Pearson.

The Pearson's correlation coefficient between two variables (ρ) is a measure of linear correlation which is defined as the covariance of the two variables divided by the product of their standard deviations.

$$\rho = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{1}$$

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked variables

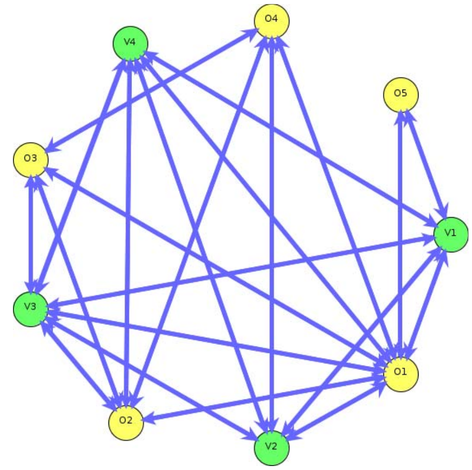

Fig. 1. A priori known graph of interactions for function WFG1. Nodes corresponding to variables $v_1, \ldots, v_4$ are shown in green. Nodes corresponding to the objectives $o_1, \ldots, o_5$ are shown in yellow.

[10]. For a sample of size n, the n raw scores X_i, Y_i are converted to ranks x_i, y_i, and r is computed from:

$$s = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} \tag{2}$$

where $d_i = x_i - y_i$, is the difference between ranks.

Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be a set of observations of the joint random variables X and Y respectively, such that all the values of x_i and y_i are unique. Any pair of observations (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree: that is, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant, if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant.

The Kendall τ coefficient is defined as [10]:

$$\tau = \frac{n_c - n_d}{\frac{1}{2} n(n-1)} \tag{3}$$

where n_c is the number of concordant pairs and n_d is the number of discordant pairs.

B. Function benchmark and a priori defined relationships

As a benchmark to compare the performance of the correlation measures we have chosen the function WFG1 of the Walking Fish Group (WFG) [11]. The problems included in the WFG set encompass a diverse set of properties that can be found in real-world MOPs and, thus, raise substantial obstacles for any multiobjective optimization algorithm. Function WFG1 comprises 5 objectives and 16 decision variables. The first $k = 4$ variables determine the solution's position and the last $l = 12$ variables serve as a distance parameter. A simplified definition of the function [6], [11] is shown in Equations (4)-(8). In these equations, the functions $h_1(.), \ldots, h_5(.)$ are shape functions, $a = g_1(x_5, \ldots, x_{16})$, and functions $g_1(.)$ and $g_2(.)$ represent a composition of transformations of the input variables. Function WFG1 was previously used in [6] to investigate the accuracy of an EDA based on Multi-Bayesian networks.

$$f_1(x) = a + 2 \times h_1(g_2(x_1), g_2(x_2), g_2(x_3), g_2(x_4)) \quad (4)$$
$$f_2(x) = a + 4 \times h_2(g_2(x_1), g_2(x_2), g_2(x_3), g_2(x_4)) \quad (5)$$
$$f_3(x) = a + 6 \times h_3(g_2(x_1), g_2(x_2), g_2(x_3)) \quad (6)$$
$$f_4(x) = a + 8 \times h_4(g_2(x_1), g_2(x_2)) \quad (7)$$
$$f_5(x) = a + 10 \times h_5(g_2(x_1)) \quad (8)$$

In real-world problems the relationships between variables and objectives can be complex. For instance, the influence of a single variable on a single objective can be linear, quadratic, or has other impact, more difficult to handle analytically. However when multiple variables influence an objective the relationships can be even more intricate with complex pattern of interactions. Similarly, the objectives can be related because they depend on similar variables or describe correlated measures.

To simplify our analysis, we define the relationships between the problem components in general terms, without analytically defining the precise nature of the relationship. Although relationships defined in this way are less specific, they are also more general in the sense that they can represent different types of underlying dynamics of interaction between the problem components. We defined the relationships between the problem components as follows:

- A pair of variables is considered to interact if they are used together in at least one objective in a non-additive way.
- Two objectives interact if they have at least one variable in common.
- An objective is related with a variable if the variable is used by the corresponding objective function.

These are essentially the same definition criteria used in [6]. However, we emphasize that the definition of the relationships can be a very problem-related question and in other MOPs, other possible definitions of the relationships could be more appropriate.

Using these criteria, we have identified that the first four variables (v_1, v_2, v_3 and v_4) are related between themselves. All objectives ($o_1, \dots, o_5$) are also related between themselves. Each objective is related with all variables except the following cases: the third objective is not related with the fourth variable, the fourth objective is not related with the third and fourth variables, and the fifth objective is not related with the second, third and fourth variables. A graph describing the relationships of the WFG1 problem is shown in Figure 1.

III. Capturing and visualizing relationships in MOEAs

There are a number of questions that should be solved in order to evaluate the accuracy of the correlation measures

1) Determination of the data to be used in the computation of the correlations.
2) Determination of the criteria to classify a candidate relationship as stable.
3) Meaningful way of visualizing and interpreting the relationships.

In this section we analyze these questions.

A. Data selection

One of the issues in the selection of data is that an evolutionary algorithm can evaluate a sizable amount of solutions during the evolution. It would be too costly to evaluate the correlations every generation of the algorithm. Another element is that, we do not know in advance at which stage of the evolution is the information generated by the MOEA more reliable in terms of the accuracy of the detected relationships. Finally, it is expected that the number and type of the dependencies will depend, at least to a certain extent, on the type of MOEA.

Our choice has been to evaluate the correlations after a fixed number of iterations k. However, other criteria could be used in the future, as using parameters that describe the quality of the search as an indicator of when to test for correlations. By conducting a periodical evaluation of the correlations along the evolution we can also track the evolution of the dependencies. Finally, as explained in the previous section, we have selected the MOEA/D-DRA algorithm to evaluate the correlation methods. The question of investigating how different MOEAs behave in terms of the relationships they generate in the data is a relevant one but it is beyond the scope of this paper.

B. Criteria to evaluate the relationships

One of the challenges in the comparison of the correlation measures is the stochastic nature of the data produced by the optimization algorithm. Every run of the algorithm corresponds to a different path of the search and as a result the populations produced by each algorithm in each run will not likely be exactly the same. To deal with the extraordinary variability of the data between runs, it is necessarily to establish criteria to define when a detected relationship between the problem components is likely to be true. We assume that relationships that frequently appear in the different generations and runs of the algorithm are more likely to be true.

The basic block of our analysis is the symmetric correlation matrix M. This matrix has dimension $n + m$ where n is the number of variables and m is the number of objectives. Each entry (i, j) of M stores the correlation between components i and j. The $n \times n$ submatrix M_v in the upper left of M encodes the correlations between variables and the $m \times m$ submatrix M_o in bottom-right of M represents the correlations between objectives. We compute one matrix M for each population (of those for which correlations are computed), each run, and each correlation measure. These sets of matrices are used to identify the stable relationships which can be represented with correlation graphs.

The correlation matrix M is thresholded to remove weak correlations. To threshold the correlation graphs we check if the absolute value of the correlation is greater than a threshold t_1. Associated with matrix M we create a binary matrix B of the same dimension. In B, there is a value 1 in cell (i, j) if the relationship between components i and j is considered to be sufficiently strong, and 0 otherwise. The same threshold t_1 is used to set the entries of B. Therefore, matrix B is essentially a matrix that has value 1 only in those entries where M is non zero.

For a given run of the algorithm, we have a set of matrices $(M^1, \ldots, M^r)$ where r is the number of populations that have been used to compute the correlation ($r < n_{gen}$). Similarly, we have a set of matrices $(B^1, \ldots, B^r)$. Another threshold (t_2) is used at the end of each execution to determine if an interaction is stable. To consider a relationship stable it should have been detected in $(t_2 \times 100)\%$ of the correlation matrices computed. These stable interactions can be represented in a correlation graph associated to the run. This is a binary matrix constructed from the aggregation and thresholding of matrices $(B^1, \ldots, B^r)$. To evaluate the relationships over the search a threshold (t_3) is used. A relationship is considered as present for the calculation of the specificity and sensitivity if it is set as true in more than $(t_3 \times 100)\%$, considering all 30 executions of a specific iteration.

The correlation graph reminds of the frequency matrices used in EDAs [12]. However, frequency matrices do not represent relationships between objectives and are constructed from the graphical models learned by EDAs, not from correlations, as is proposed in this paper.

t_1, t_2 and t_3 are parameters of the analysis or the relationships that can be manipulated to increase or relax the preciseness of the detection process. In Section IV, we present the set of parameters used in our experiments.

C. Meaningful visualization of relationships

Knowledge about the relationships encoded in the correlation matrices could be used by methods that improve the MOEA search using problem information. This representation could be also beneficial for the end-users, for instance, to unveil previously unknown characteristics of the problem, as done in the innovization approach [8]. Therefore, in addition of computing the correlation matrices and correlation graphs, it is important to design appropriate ways for the visualization of this information. One of the goals of our work was to investigate this issue. First, we identified what sort of information could provide knowledge about the characteristics of the problem. The following aspects were identified:

- Which components interact?
- What type of interaction is described by the correlation (negative or positive correlation)?
- What is the strength of the correlation?

To address these issues, the following conventions were taken regarding the visualization:

- Each component of the problem is represented by a node.
- Objectives and variables are represented in different colors. Green color is used for variables, yellow for objectives.
- An arrow between components defines the existence of a relationship between them.
- The color of the relationship defines the type of interaction. Positive correlations are represented in blue. Negative correlations are defined in red.
- The width of the arrow between two nodes defines the strength of the relationship. Thicker arrows correspond to stronger interactions.
- Weights can be added to the arrows to further describe the strength of the correlation.

Figure 2 shows an example of the visualization used to present the correlation matrices learned. This type of graphical representation allows to convey a significant amount of information about the problem. When graphs from different generations of the algorithm are simultaneously shown, it is possible to infer not only properties of the problem but also to get information about the behavior of the algorithm. Notice, that the type of representation we propose is more expressive than those previously used to describe the behavior of other evolutionary algorithms [12], [13].

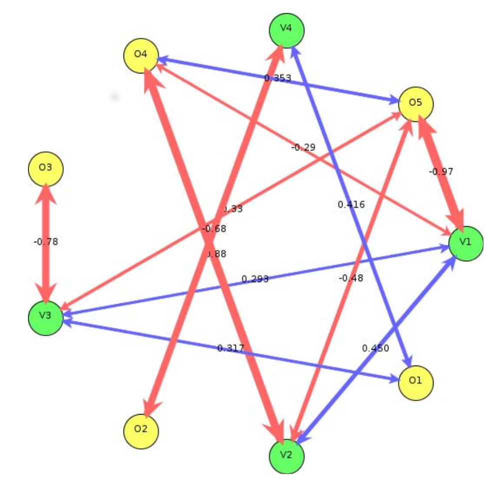

Fig. 2. Example of the visualization of the correlation matrices recovered by the algorithm.

IV. Experiments

We remind the main elements of our experimental pipeline:

- Execute MOEA/D-DRA algorithm to optimize the WFG1 problem. Store the sampled solutions and objectives.
- Apply Kendall, Spearman and Pearson metrics to the stored data to detect correlations between the problem components.
- Mine the correlation matrices to identify the stable relationships.
- Create and visualize the correlation graphs showing the sign and strength of the correlations.

A. Experimental framework

To collect the data used for the analysis, 30 independent runs of MOEA/D-DRA were executed. The correlation matrices for the three metrics were computed using populations collected at intervals of 26 generations. The parameters used in the computation of the correlation matrices and correlation graphs were $t_1 = 0.25$, $t_2 = 0.2$ and $t_3 = 0.3$.

B. Computation of structural accuracy

A key issue in the comparison of the metrics is how to evaluate their accuracy at capturing the original structure. We used the sensitivity and specificity measures [14], extensively applied in the machine learning domain. Both measures are computed using the following auxiliary parameters: True Positive (TP), the number of the real relationships correctly classified; False Negative (FN), the number of non-existing relationship and classified as existing; True Negative (TN), the number of relations that do not exist and are correctly classified; and False Positive (FP), the number of existing relationships classified as non-existing.

The sensitivity (Equation 9) of a metric reflects how much this is effective in properly identifying, among all the evaluated individuals, those who really have the characteristic of interest. In other words, correctly classified relationship through the metric that also exist in the real structure.

The specificity (Equation 10) of a metric reflects how much this is effective in correctly identifying individuals that do not have the condition of interest. That is, correctly classified relationships that do not exist in metric and did not exist in the real structure either.

$$SEN = \left(\frac{TP}{TP + FN}\right) \quad (9)$$

$$SPE = \left(\frac{TN}{TN + FP}\right) \quad (10)$$

The sensitivity and specificity measures are in the range $[0, 1]$, where 0 expresses that the metric does not capture any of the relationships, and returning 1 means that the entire set of relationships was captured. These measures were previously applied in [15] to investigate the accuracy of the MOEA introduced in that paper.

C. Results

In this section we present the results of the conducted experiments.

1) Comparison between the correlation measures: In Table I we present the sensitivity and specificity obtained for each of the correlation metrics investigated. We evaluate the accuracy of the metrics to detect three different types of relationships: 1) Between variables, 2) between variables and objectives, and 3) between objectives separately. In addition, we evaluate the overall performance of the algorithm using a global measure 4) all relationships together. At the time of evaluating the relationships between objectives we do not consider the specificity since all objectives are related.

TABLE I
SENSITIVITY AND SPECIFICITY OBTAINED USING KENDALL, PEARSON, AND SPEARMAN CORRELATION MEASURES

	Kendall	Pearson	Spearman
var/var	0.17/**1.00**	**0.83**/0.66	0.50/0.80
var/obj	0.08/**0.83**	0.10/0.66	**0.12**/0.50
obj/obj	0.00/ *	**0.40/** *	0.10/ *
	0.08/**0.99**	**0.19**/0.66	0.14/0.79

* all five objectives are related, specificity are not applicable.

An analysis of Table I reveals that the Kendall's correlation obtains the best specificity value in all cases, but at the expense of achieving the worst sensitivity too. Pearson's correlation achieves the best sensitivity in most cases. However, for the same cases it obtains the worst specificity. The overall best behavior is that of the Spearman's correlation that achieves the best balance. In most cases it has a better specificity than Pearson's and better sensitivity than Kendall's.

2) Evolution of the relationships during the search: We also analyze the correlation metrics specificity and sensitivity along the search process. Our objective is to verify if it is possible to capture the relationships along the search process, and if the metrics ability is the same in all stages of the search. Figure 3 shows the evolution of the specificity in the the different generations.

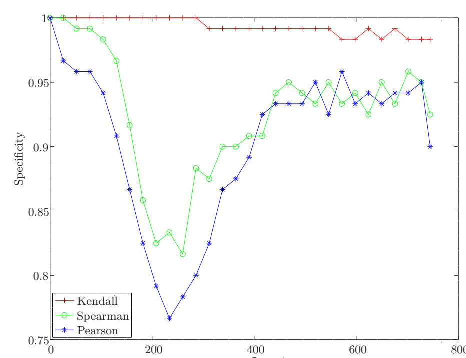

Fig. 3. Specificity along the search process

An analysis of Figure 3 reveals that the Kendall's correlation reaches a value very close to the maximum throughout the search. It is maximized around the first stages of the search. The other two metrics start with very high specificity values but considerably decrease until generation 250, where the value is increased again. The final values of the specificity are stabilized at the end of the search between 0.90 and 0.95, approximately. It is possible to note that at the initial stage of the search few relationships are detected and they are all present in the problem. Then, all correlation metrics begin with high specificity values (1.0). Along the search, more relationships are detected, some of them false positives, decreasing the specificity.

Figure 4 shows the behavior of the sensitivity measure for the three correlation measures. The figure illustrates that the ability of the metrics to recover the original relationships of the problem steadily increases along the search process. This behavior is probably due to the fact that the exploration used along initial stages of the search can generate many erratic solutions, that do not represent the problem characteristics, but as the solutions evolve, it becomes easier to detect the true relationships. Our analysis shows that, if the analysis is made considering the different stages of the search separately, the metrics have difficulty to capture a very accurate model of the true relationships. This can be observed because the sensitivity is under 0.25 for all metrics. In a supplementary material[1] we show the mean relationships along the search process.

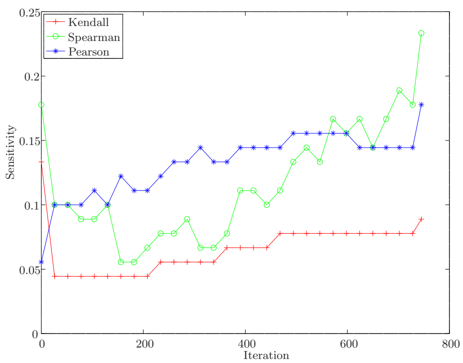

Fig. 4. Sensitivity along the search process

V. Conclusions

In this work we have evaluated the suitability of different correlation measures to capture relationships in multi-objective optimization. We have proposed a way to visualize these correlations supporting information about the characteristics of the problem but also about the behavior of the algorithm in terms of the accuracy of the detected interactions.

We conclude that it is possible to capture the relations along the evolutionary process using correlation metrics and that different correlation metrics have different characteristics. In our experiments the Kendall's correlation was the best in terms of specificity, it means, the ability of not reporting false positives relations. Pearson's had better results in terms of sensitivity than the others because is was better at recovering the known interactions. Spearman's correlation showed the best balance between specificity and sensitivity. We did not find strong evidence suggesting that any of the metrics was better than the others in both: sensitivity and specificity

As lines of future research we consider the investigation of other multi-objective algorithms. This would help to understand whether and how different MOEAs behave differently in terms of their accuracy to recover the problem interactions. We need also to investigate other types of functions that exhibit different patterns of relationships between their components. Finally, our long term goal is designing ways to use the captured information to improve the evolutionary search and its results.

[1]Available at http://www.pdf-archive.com/2015/07/28/bare-conf/bare-conf.pdf

Acknowledgments

This work was supported by CNPq (Productivity Grant Nos. 305986/2012-0 and Program Science Without Borders Nos. : 400125/2014-5), by CAPES, by IT-609-13 program (Basque Government) and TIN2013-41272P (Spanish Ministry of Science and Innovation).

References

[1] C. Coello, G. Lamont, and D. Van Veldhuizen, *Evolutionary Algorithms for Solving Multi-objective Problems*. Springer-Verlag New York Inc, 2007.

[2] A. López Jaimes, C. A. Coello Coello, H. Aguirre, and K. Tanaka, "Objective space partitioning using conflict information for solving many-objective problems," *Information Sciences*, vol. 268, pp. 305–327, 2014.

[3] R. Shang, K. Zhang, L. Jiao, W. Fang, X. Zhang, and X. Tian, "A novel algorithm for many-objective dimension reductions: Pareto-PCA-NSGA-II," in *Evolutionary Computation (CEC), 2014 IEEE Congress on*. IEEE, 2014, pp. 1974–1981.

[4] D. Brockhoff and E. Zitzler, "Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization," in *Parallel Problem Solving from Nature-PPSN IX*. Springer, 2006, pp. 533–542.

[5] L. Marti, J. Garcia, A. Berlanga, C. A. Coello, and J. M. Molina, "On current model-building methods for multi-objective estimation of distribution algorithms: Shortcommings and directions for improvement," Department of Informatics of the Universidad Carlos III de Madrid, Madrid, Spain, Tech. Rep. GIAA2010E001, 2010.

[6] H. Karshenas, R. Santana, C. Bielza, and P. Larrañaga, "Multi-objective optimization based on joint probabilistic modeling of objectives and variables," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 519–542, 2014.

[7] R. Cheng, Y. Jin, and K. Narukawa, "Adaptive reference vector generation for inverse model based evolutionary multiobjective optimization with degenerate and disconnected pareto fronts," in *Evolutionary Multi-Criterion Optimization*. Springer, 2015, pp. 127–140.

[8] K. Deb, "Unveiling innovative design principles by means of multiple conflicting objectives," *Engineering Optimization*, vol. 35, no. 5, pp. 445–470, 2003.

[9] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained mop test instances," School of CS & EE, University of Essex, Tech. Rep. CES-491, February 2009.

[10] J. L. Myers, A. Well, and R. F. Lorch, *Research design and statistical analysis*. Routledge, 2010.

[11] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *Evolutionary Computation, IEEE Transactions on*, vol. 10, no. 5, pp. 477–506, Oct 2006.

[12] R. Santana, P. Larrañaga, and J. A. Lozano, "The role of a priori information in the minimization of contact potentials by means of estimation of distribution algorithms," in *Proceedings of the Fifth European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, ser. Lecture Notes in Computer Science, E. Marchiori, J. H. Moore, and J. C. Rajapakse, Eds., vol. 4447. Valencia, Spain: Springer, 2007, pp. 247–257.

[13] H. Karshenas, A. Nikanjam, B. Helmi, and A. Rahmani, "Model accuracy for hierarchical problems," in *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, vol. 1. IEEE, 2009, pp. 852–856.

[14] D. G. Altman and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity." *BMJ: British Medical Journal*, vol. 308, no. 6943, p. 1552, 1994.

[15] H. Karshenas, R. Santana, C. Bielza, and P. Larranaga, "Multi-objective optimization with joint probabilistic modeling of objectives and variables," in *Evolutionary Multi-Criterion Optimization*. Springer, 2011, pp. 298 – 312.