

Analyzing Dominance Move (MIP-DoM) Indicator for Multi- and Many-objective Optimization

Claudio Lucio do Val Lopes , Flávio Vinícius Cruzeiro Martins , Elizabeth Fialho Wanner ,
and Kalyanmoy Deb *Fellow, IEEE*

COIN Report Number 2020028

Abstract—Dominance move (DoM) is a binary quality indicator that can be used in multi-objective and many-objective optimization to compare two solution sets obtained from different algorithms. The DoM indicator can differentiate the sets for certain important features, such as *convergence*, *spread*, *uniformity*, and *cardinality*. The DoM indicator has an intuitive and physical meaning, similar to the ϵ -indicator, and calculates the minimum total move of members of one set needed to weakly dominate each member of the other set. Despite the aforementioned properties, DoM is hard to calculate, particularly in higher dimensions. There is an efficient and exact method to calculate it in a bi-objective case only. This work proposes a novel approach to calculate DoM using a mixed integer programming (MIP) approach, which can handle sets with three or more objectives and is shown to overcome the ϵ -indicator’s information loss. Experiments, in the bi-objective space, are done to verify the model’s correctness. Furthermore, other experiments, using 3, 5, 10, 15, 20, 25 and 30-objective problems are performed to show how the model behaves in higher-dimensional cases. Algorithms, such as IBEA, MOEA/D, NSGA-III, NSGA-II, and SPEA2 are used to generate the solution sets (however any other algorithms can also be used with the proposed MIP-DoM indicator). Further extensions are discussed to handle certain idiosyncrasies with some solution sets and also to improve the quality indicator and its use for other situations.

Index Terms—Multi-objective optimization, multi-criteria optimization, quality indicators, performance assessment, mixed-integer programming, evolutionary algorithms.

I. INTRODUCTION

PROBLEMS with conflicting objectives arise in most real-world optimization problems. These problems are solved using evolutionary multi-objective (EMO) or many-objective optimization (EMaO) techniques [1] to find a set of Pareto-optimal solutions in three or less, or more than three-objective problems, respectively. EMO/EMaO algorithms are applied in different domains [2], [3]. The solution sets for these problems

are formed in such a way that each solution represents a *trade-off* among the conflicting objectives.

In contrast to the single-objective optimization, the solutions generated by EMO/EMaO algorithms can be difficult to obtain and compare. Such difficulty grows as the need for having an increasingly large number of candidate solutions with an increase in number of objectives. When there are two or three objectives only, some graphical techniques help to examine the solution set visually. However, when the number of objectives is greater than three, this task is challenging, needing more advanced visualization techniques that can present location, shape, and solution set distribution [4], [5], [6], [7].

Quality indicators are suitable for situations when we need to compare two or more trade-off solution sets [8]. They have been used to compare the outcomes of multi-objective algorithms quantitatively. In a recent paper [9], 100 quality indicators were discussed considering some state-of-the-art indicators focusing on which quality aspects these indicators have, as well as their strengths and weaknesses.

In [10], a new quality measure, called dominance move (DoM), was proposed. This measure is able to capture different aspects of a solution set’s quality, such as spread, uniformity, and cardinality. Moreover, the indicator is also Pareto compliant [8], meaning that a move from a set to another completely dominated set will have a smaller indicator value than the other way around. The DoM measures the minimum ‘effort’ that one solution set has to make in trying to weakly dominate the second set, more specifically the absolute sum of the movements (i.e., Minkowski distance) in objective directions needed to make the first set weakly dominant. Authors to the original study presented an exact approach to calculate DoM for the bi-objective case. Despite the low computational cost in the bi-objective case, the exact DoM computational approach cannot be applied directly or extended easily to problems with three or more objectives.

In [11], as an attempt to circumvent the bi-objective limitation, authors proposed an assignment problem formulation to calculate the DoM in problems with three or more objective functions. They carried out preliminary experiments using problems with three objectives. Results showed that the assignment formulation for DoM is valid and correct, but the cardinality of the solution sets imposed a handicap in preventing its use in solutions sets with more than 20 solution candidates.

Although DoM presents good quality aspects as an indica-

Corresponding author: Cláudio Lúcio V. Lopes is with Postgraduate Program in Mathematical and Computational Modeling, CEFET-MG, Brazil (email: claudiolucio@gmail.com)

Flávio V. C. Martins is with the the Computer Department at the Centro Federal de Educação Tecnológica de Minas Gerais, 30510-000 Brazil (e-mail: flaviocruzeiro@cefetmg.br).

Elizabeth F. Wanner is with the Computer Department at the Centro Federal de Educação Tecnológica de Minas Gerais, 30510-000 Brazil and Computer Science Department, Aston University, Birmingham, UK (e-mail: efwanner@cefetmg.br).

Kalyanmoy Deb is with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA (email: kdeb@msu.edu).

tor, its calculation with a small computational cost poses as a big challenge [12], [13], [9]. This work focuses on a DoM calculation procedure based on a mixed integer programming (MIP) approach [14] aiming to overcome this difficulty. The mixed integer formulation is presented here. Initial experiments on two-objective problems amply show that the MIP-DoM formulation is correct. More specifically, this paper presents the following contributions: (i) the proposed MIP model for DoM calculation is valid for any objective space dimension and cardinality of the supplied solution sets, (ii) extensive evaluation of the proposed MIP-DoM formulation and approach in a number of common problem sets having 3, 5, 10, 15, 20 and 25 objective functions, (iii) comparison with optimal DoM solutions in cases where it is possible to achieve the optimal solution, and (iv) raise of questions and details about the model behaviour for some test sets, including future research paths to tackle other problems.

This paper is organized as follows. In Section II, some problems and definitions are posed, regarding unary, binary, and k -ary quality indicators. Some well-known quality indicators are presented, as well. Dominance Move (DoM), its definition, properties, and features are presented in Section III. Our mixed-integer approach to calculate DoM is presented in Section IV. The MIP-DoM model is introduced with some comments and considerations. Experiments are discussed in Section V, and it starts with bi-objective case presenting the correctness and model validity. Following, some common multi-objective problem sets and algorithms are used to compare MIP-DoM with the ϵ -indicator, IGD^+ , and Hypervolume indicators. Still in this section, the advantages of DoM over the ϵ -indicator are presented, although they share a similar interpretation regarding its use. Still in Section V, the many-objective experiments are discussed using problem sets with 5, 10, 15, 20 and 25 objectives. Finally, in Section VII, some final considerations, recommendations, and future research paths are presented.

II. PROBLEMS AND DEFINITIONS

In general, a multi-objective optimization problem (MOP) includes a decision variable vector, \mathbf{x} , from a feasible decision space $\Omega \subseteq \mathbb{R}^N$, and a set of M objective functions. Without loss of generality, the minimization of an MOP can be simply defined as [15]:

$$\text{Minimize } F(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_M(\mathbf{x})]^T, \quad \mathbf{x} \in \Omega. \quad (1)$$

$F : \Omega \rightarrow \Theta \subseteq \mathbb{R}^M$ is a mapping from the feasible decision space Ω to vectors in the M -dimensional objective space Θ . We are interested in the evaluation of these objective vector (solution) sets, and the comparison relation among them.

Consider two solutions, $\mathbf{p}, \mathbf{q} \in \Theta$. \mathbf{p} is said to *weakly dominate* \mathbf{q} if $p_m < q_m$ for $1 \leq m \leq M$, and it is denoted as $\mathbf{p} \preceq \mathbf{q}$. Moreover, if for all objectives $p_m \leq q_m$ and there is at least one objective m on which $p_m < q_m$, then it is said that \mathbf{p} *dominates* \mathbf{q} , and it is denoted as $\mathbf{p} \prec \mathbf{q}$. Thus, it is clear that if $\mathbf{p} \preceq \mathbf{q}$, \mathbf{p} also dominates \mathbf{q} , or, $\mathbf{p} \prec \mathbf{q}$, but not vice versa. A solution $\mathbf{p} \in \Theta$ is called *efficient* (or, loosely a member of the Pareto-optimal front) if there is no $\mathbf{q} \in \Theta$ that

dominates \mathbf{p} . Such solutions constitutes a Pareto-optimal front (or an efficient set) in the objective space. In the same way, the weak dominance relation can be stated to solution sets:

*Definition 1: Weak Dominance*¹: The set \mathbf{P} weakly dominates \mathbf{Q} , i.e. it is denoted as $\mathbf{P} \preceq \mathbf{Q}$, if every solution $\mathbf{q} \in \mathbf{Q}$ is weakly dominated by at least one solution $\mathbf{p} \in \mathbf{P}$.

The comparison between two or more solution sets is of great importance in EMO algorithms and hence a considerable amount of efforts have been made to define different performance indicators. The indicators can be used to compare the outcomes of multi-objective algorithms or even assist an algorithm during the search process for non-dominated candidates. It is paramount to make more precise statements when applying a quality indicator comparison, for example, if one algorithm is better than another, how much better is it? The following definition formalizes the quality indicators [8]:

Definition 2: Quality indicator: An k -ary quality indicator I is a function $I: \Theta^k \rightarrow \mathbb{R}$, which assigns each vector of k solution sets $(\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_k)$ a real value $I(\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_k)$.

The quality indicators can be unary, binary, or k -ary, defining a value to one solution set, two solution sets, or k solution sets, respectively. In [9], 100 indicators were listed and some were discussed in detail. In this work, four properties of indicator quality are analyzed: *convergence*, *spread*, *uniformity*, and *cardinality*.

It is expected that the Pareto-dominance must be a central criterion in reflecting the *convergence* of solution sets [8]. For two solution sets \mathbf{P} and \mathbf{Q} , for example, if \mathbf{P} weakly dominates \mathbf{Q} , then $I(\mathbf{P}, \mathbf{Q}) = 0$. If \mathbf{P} dominates some points of \mathbf{Q} and \mathbf{Q} does not dominate any point of \mathbf{P} , it is reasonable to expect that the indicator value will be more for $\mathbf{Q} \rightarrow \mathbf{P}$ ($I(\mathbf{Q}, \mathbf{P})$) than $\mathbf{P} \rightarrow \mathbf{Q}$ ($I(\mathbf{P}, \mathbf{Q})$).

The *spread* of a solution set must consider the region that the set is covering. It involves finding the extreme points and a well-distributed set of points in the interior of the Pareto set [16].

The number of solutions in the set is another property, known as *cardinality* (in general, solutions sets with more candidates and generated with the same computational resources are preferred). Finally, a good indicator must prefer a set with uniformly distributed points, *uniformity*, showing an equidistant spacing (either along the Pareto surface or in an Euclidean sense) amongst solutions.

It is plausible to add more details about the indicators due to their usage, such as the computational cost. Some indicators present all quality aspects but are hard to compute, notably in high dimensions or for high-cardinality sets. Other indicator details also deserve to be mentioned, such as the necessity for reference point or set, additional parameters, how to deal with scale, and normalization, etc., but we do not consider them in this study.

There are many indicators available, and they have been used in numerous situations in the literature [9]. Hypervolume (HV) [17], [18], [19], generational distance (GD) [20], inverted

¹There are other common relations defined for solution sets, i.e. *strictly dominance*, *dominance*, *better*, but for DoM context, the weak dominance concept is enough.

generational distance plus (IGD+) [20], KKTPM [21], and ϵ -indicator [8] are some examples. Our supplementary material further present the definition of some indicators.

Hypervolume has been widely applied in the evolutionary community as a unary indicator choice [22]. It is still important to note its drawbacks: the increasing exponential cost (due to the number of objectives) and the necessity to designate an explicit reference point [23]. The reference point affects the ordering of pairs of incomparable sets. Some community efforts have focused on the discovery of new indicators, which can overcome hypervolume's limitations while providing a good set of properties.

An ideal quality indicator must present the four facets. Additionally, it must have a low computational cost, and it does not need a normalization (due to objective scale) and any additional parameters or reference points/sets.

III. DOMINANCE MOVE

Dominance move (DoM) is an intuitive indicator that is commensurate with all four desirable properties [10]. The first idea presenting DoM came from the performance comparison indicator (PCI) [12]. Examining the PCI proposal, it is quite similar to DoM in its essential purpose. PCI, a binary quality indicator, builds up a reference set using two solution sets, P and Q . This reference set is then split up into clusters, and the indicator calculates the movement distance in order to weakly-dominate the clusters.

Dominance move is a measure for comparing two sets of multi-dimensional points, being classified as a binary indicator. It considers the minimum overall movement of points in one set needed to *weakly-dominate* each and every member of the other set. Recall that a point x weakly-dominate (\preceq) another point y , if $f_i(x) < f_i(y)$ for all $j = 1, \dots, M$. Thus, if $f_i(x) = f_i(y)$ for any i , x cannot weakly-dominate y , but a small reduction in $f_i(x)$ by an infinitesimal $\delta_i \rightarrow 0$ will make the revised x to weakly-dominate y . In such a case, since the move is small, we consider $x \preceq y$ with a vanishing dominance move. The DoM does not need *a priori* problem knowledge, parameters, or a reference set. However, the computational cost to calculate DoM is prohibitive, and the method to obtain DoM for a number of objective functions higher than two is unknown. DoM can be defined as follows.

Definition 3: DoM: Consider that P and Q are sets of points, with p_i points $i \in \{1, \dots, |P|\}$ and q_j points $j \in \{1, \dots, |Q|\}$. The dominance move of P to Q , $DoM(P, Q)$, is the minimum total distance of moving points of P , such that each and every point in Q is weakly dominated by at least one point in P . In fact, the problem aims to obtain $P' = \{p'_1, p'_2, \dots, p'_{|P|}\}$ from $\{p_1, p_2, \dots, p_{|P|}\}$ such that P' weakly dominates Q and the total move from $\{p_1, p_2, \dots, p_{|P|}\}$ to $\{p'_1, p'_2, \dots, p'_{|P|}\}$ must be minimum [10].

The formal expression of DoM can be stated as:

$$DoM(P, Q) = \min_{P' \preceq Q} \sum_{i=1}^{|P|} d(p_i, p'_i), \quad (2)$$

in which $d(p_i, p'_i)$ can be the Euclidean or the Manhattan distance between p_i to p'_i [10].

DoM reflects how far one solution set P needs to move to weakly dominate another solution set Q . Thus, $DoM(P, Q)$ is always larger than or equal to zero. A small value suggests that P and Q are close (e.g., one point of P that only needs a little movement to weakly dominate Q), and a large value indicates that P performs worst than Q (e.g., some points of P must be moved further away in an attempt to weakly dominate Q).

The dominance move indicator is based on the properties of dominance relation among solutions trying to dominate each other, considering all available information (all solutions and objectives). These solutions' efforts scale in a bottom-up manner from the solutions to the set relations. In this sense, the authors claimed that DoM possesses some properties that are enumerated in the following proposition [9]:

Proposition 1: Consider the solution sets $P, Q, R \subset \Theta$ and the points $p, q \in \Theta$:

- $P = Q \iff DoM(P, Q) = DoM(Q, P) = 0$;
- $P \triangleleft Q$ (P is better than Q) $\iff DoM(P, Q) = 0 \wedge DoM(Q, P) > 0$;
- If $P \preceq Q$, then $DoM(P, R) \leq DoM(Q, R)$ and $DoM(R, Q) \leq DoM(R, P)$;
- $DoM(P, Q) + DoM(Q, R) \geq DoM(P, Q \cup R)$ and $DoM(P, Q) + DoM(P, R) \geq DoM(P, Q \cup R)$.

The DoM formulation intends to overcome some ϵ -indicators' weaknesses while possessing a similar interpretation. The ϵ -indicator also measures the minimum value added to one solution set to make it be weakly dominated by another set. Then, it must be observed that DoM and ϵ -indicators have the same general concept and lead to the same conclusions among the solution sets. However, DoM presents the following improvements when compared to ϵ -indicators [11]:

- ϵ -indicators are only related to one particular solution and only one objective in the whole solution set;
- ϵ -additive is not able to capture differences concerning cardinality of solution sets;
- DoM can present greater values than ϵ -indicators. This fact can be explained since DoM takes into account information from all objectives.

The information loss of ϵ -indicators is critical, particularly considering many-objective scenarios (considering only the objective on which one solution performs worst relative to another solution and ignoring the difference on the remaining objective values). A very simple example, proposed in [10], illustrates this information loss: consider two 10-objective solutions, such as $p_1 = \{0, 0, 0, \dots, 1\}$ and $q_1 = \{1, 1, 1, \dots, 0\}$. Whereas p_1 presents a better performance on the first nine objectives and q_1 only on the last objective, ϵ -additive(p_1, q_1) = ϵ -additive(q_1, p_1) = 1.

Considering DoM's definition again, we aim to obtain P' from P that weakly dominates Q . In fact, DoM represents the minimum total of moves from $P = \{p_1, p_2, \dots, p_{|P|}\}$ to $P' = \{p'_1, p'_2, \dots, p'_{|P|}\}$ which is calculated using all available information and not only one objective (as proposed in the additive and multiplicative versions of ϵ -indicator).

P' represents a set of points that are candidates to dominate Q with some update in one or more objectives leading to a better distance such as expressed in Equation (2). In such way, it should be noted that each p'_i with $i \in \{1, \dots, |P|\}$ must be generated from p_i , as proposed in [11].

Let $\mathbb{P}^*(Q)$ be a set composed of all possible subsets from Q , excluding the empty set². For determining each element of P' , p'_i , it is necessary to associate the corresponding element of P to elements of $\mathbb{P}^*(Q)$. The number of all possible associations, η , is the product of $|P|$ and $|\mathbb{P}^*(Q)|$ and is given by

$$\begin{aligned} \eta &= |P| \underbrace{\sum_{g=1}^{|Q|} \binom{|Q|}{g}}_{|\mathbb{P}^*(Q)|}, \\ &= |P| \left(\binom{|Q|}{1} + \binom{|Q|}{2} + \dots + \binom{|Q|}{|Q|} \right), \\ &= |P|(2^{|Q|} - 1). \end{aligned} \quad (3)$$

Let Q_s be an element of $\mathbb{P}^*(Q)$ for $s = \{1, \dots, |\mathbb{P}^*(Q)|\}$. Observe that p_i will be used as a reference point to generate p'_i . In this way, p'_i may weakly dominate Q_s [11].

Regardless of this challenge, the original DoM paper also proposes an exact calculation method for the bi-objective case, which is meticulously detailed in [9]. For the sake of completeness and clarification, an overview of the algorithm is described next. The algorithm employs the concept of *inward neighbor*:

Definition 4: Inward neighbor. Consider P a solution set, $a \in P$, $b \in P$, and $b \neq a$. $n_R(\cdot)$ is a function which determines b as the *inward neighbor* of a , denoted as $b = n_R(a)$, if b has the smallest dominance move distance to a , that is $b = \arg \min_{p \in P/a} d(p, a)$.

Using the *inward neighbor* concept, the DoM algorithm, as presented in [9], can be outlined as:

Step 1: Remove the dominated points in both P and Q , separately. Remove the points of Q that are dominated by at least one point in P .

Step 2: Denoting $R = P \cup Q$, each point of Q in R is considered as a subset. For each subset, find its *inward neighbor* $r = n_R(q_j)$ in R . If the point $r \in P$, then merge r into the subset, otherwise r does not belong to the subset, or it is already owned by a subset. At the first case, q_j and r are merged into the subset. At the second case, there is nothing to do.

Step 3: If there exists no point $q_j \in Q$ such that $q_j = n_R(n_R(q_j))$ (i.e., there is a loop between the points) in any subset, then the procedure ends and there is an optimal solution to the case.

Step 4: Otherwise, there is a loop in one or more subset, then replace these solutions by their ideal solution (formed of the best of each objective in each solution inside the loop or subset). This leads to a new Q , then find the *inward neighbor* of such ideal point and subset then. Return to step 3 until convergence.

²It is worthwhile to mention that $\mathbb{P}^*(Q)$ represents the power set excluding the empty set ($\mathbb{P}^*(Q) \cup \emptyset = \mathbb{P}(Q)$).

All definitions, theorems, and corollaries to prove the algorithm correctness for the bi-objective case can be seen in [9].

Another attempt to solve the problem in the multi-objective and many-objective case was presented in [11]. The DoM calculation was treated as an assignment problem, and some experiments in three objective spaces were also presented. However, the proposal has some drawbacks that prevent its use in solutions sets with more than 20 solution candidates. The problem comes from the fact that the Equation (3) is calculated using a full combinatorial approach. Nevertheless, the method presented optimal solutions to some problems, and it is showed that the dimension of the objective space barely influences it.

IV. THE MIP DOMINANCE MOVE CALCULATION APPROACH

Our goal is to modify the DoM formulation to deal with problems having three or more objective functions. Our DoM calculation proposal is based on the perspective that the problem is, in fact, a particular instance of an assignment problem with two levels and some restrictions, as stated in [11]. It is considered that to treat the problem, we have to find an assignment of P to Q with the restrictions that each q_j must be assigned to one p_i with the minimum distance. Nevertheless, in classic assignment problems, it is not possible that points in P change their own features, such as changing positions to alter their distances. In the proposed DoM calculation, this issue is considered.

A simple example to clarify the situation is given next. Let set P be $\{(2.0, 2.5), (3.0, 1.9)\}$ and set Q be $\{(2.2, 2.0), (3.0, 1.5)\}$. The *inward neighbors* $p_i = n_R(q_j)$ of points q_1 and q_2 are p_1 and p_2 , respectively. This creates an assignment of P to Q with the minimum $DoM(P, Q)$. In this sense, a search to find p'_1 and p'_2 should be done, in a way that they must be generated to weakly dominate q_1 and q_2 , respectively. Let P' be $\{(2.0, 2.0), (3.0, 1.5)\}$, for example, the $DoM(P, Q) = d(p_1, p'_1) + d(p_2, p'_2) = 0.5 + 0.4 = 0.9$.

The problem can be modeled as a mixed integer programming (MIP) approach and using the DoM definition. P and Q are sets of points, with p_i points $i \in \{1, \dots, |P|\}$ and q_j points $j \in \{1, \dots, |Q|\}$, respectively. Each p_i is composed of $p_{(i,m)}$ components, $1 \leq m \leq M$ in which M is the number of objective functions. Analogously, each q_j is composed of $q_{(j,m)}$ components, $1 \leq m \leq M$.

Given P and Q , P' is a set of points, in which each p'_i is generated from p_i after some updates in one or more objectives to weakly dominate a Q_s . It is important to note that if P already weakly dominates Q , then $p'_i = p_i$. Observe that P' must weakly dominate Q , resulting in a better distance such as expressed in Equation (2).

The proposed MIP model calculates the distance $d(p_i, p'_i)$. In order to calculate this distance, a new point, \hat{p}_i , is obtained and there is strict relation between \hat{p}_i and p'_i . At the end, the MIP-DoM model calculates the final p'_i using \hat{p}_i .

Figure 1 illustrates the intuition behind our proposal in a bi-objective problem. The $d(p_i, p'_i)$ is composed of two

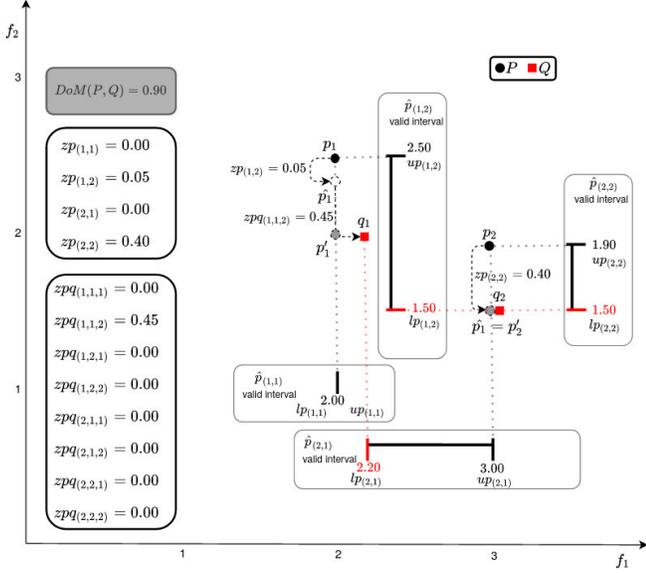


Figure 1: An example that shows up the intuition behind the MIP-DoM approach. Consider two sets, $P = \{(2.0, 2.5), (3.0, 1.9)\}$ and $Q = \{(2.2, 2.0), (3.0, 1.5)\}$. The DoM value indicates the total movement for P to weakly dominate Q . A typical vertical movement, $zp_{(1,2)} = 0.05$ represents an improvement in f_2 objective generating \hat{p}_1 . However, \hat{p}_1 does not dominate q_1 . There is still a movement to be done in f_2 objective, represented by $zpq_{(1,1,2)} = 0.45$, so that p'_1 will weakly dominate q_1 . Observe that point p_2 does not weakly dominate q_2 , so it is vertically moved obtaining \hat{p}_2 , in an attempt to weakly dominate q_2 . With this vertical movement, indicated by $zp_{(2,2)} = 0.40$, \hat{p}_2 dominates q_2 , becoming p'_2 . The DoM value is the total distance using the terms $zp_{(1,2)}$, $zp_{(2,2)}$, and $zpq_{(1,1,2)}$. Limits on p' movements (lower ($lp_{(i,m)}$) and upper bounds ($up_{(i,m)}$)) are also indicated.

terms, $zp_{(i,m)}$ and $zpq_{(i,j,m)}$. The summation of all $zp_{(i,m)}$, $\forall m \in \{1, \dots, M\}$ results in $d(p_i, \hat{p}_i)$. Analogously, the summation of all $zpq_{(i,j,m)}$, $\forall m \in \{1, \dots, M\} \forall j \in \{1, \dots, |Q|\}$ results in $d(\hat{p}_i, p'_i)$. \hat{p}_i , indicated as unfilled circles, represents the point after the first movement. For each p_i and each m -th objective, there are a lower ($lp_{(i,m)}$) and upper ($up_{(i,m)}$) bounds constraining \hat{p}_i 's movement. The bounds are generated considering that the $\hat{p}_{(i,m)}$ needs to be smaller or equal to the minimum value in the Q taking the m -th objective into account (lower bound), and cannot be greater than $p_{(i,m)}$ (upper bound). The component $\hat{p}_{(i,m)}$ acts as an extra variable in our MIP model that leverages the calculation. After finding a $\hat{p}_{(i,m)}$ within the corresponding bounds, $zp_{(i,m)}$ represents the extent of movement. In Figure 1, there are \hat{p}_1 and \hat{p}_2 , represented as unfilled circles and there are some improvements in f_2 objective generating the $zp_{(1,2)} = 0.05$ and $zp_{(2,2)} = 0.40$.

The $zpq_{(i,j,m)}$ defines the move in m -th objective for the i -th component of $\hat{p}_{(i,m)}$ to generate the $p'_{(i,m)}$. In Figure 1, the improvement generated by the \hat{p}_2 is good enough to weakly dominate q_2 . In that case, $zpq_{(2,2,1)}$ and $zpq_{(2,2,2)}$ are equal to zero, considering that $\hat{p}_2 = p'_2$. The same does not happen

for \hat{p}_1 ; there is still a distance such that \hat{p}_1 must be moved to dominate q_1 . This distance is indicated by the $zpq_{(1,1,2)} = 0.45$. The final p'_1 is obtained using \hat{p}_1 and $zpq_{(1,1,2)}$.

The distances are calculated using an MIP model expressed from Equation (4) to (11). The model contains continuous z -variables (zp and zpq) and binary x -variables (discussed later), and a number of constraint sets, which are combined in an attempt to find the minimum DoM value. The objective function uses z -variables. There is a constraint set related to $zpq_{(i,m)}$, and it is expressed in the set of Equations (5). Essentially, it must be greater than or equal to zero and it must be less than $p_{(i,m)}$ and, finally, it must be greater than or equal to the difference between $p_{(i,m)}$ and $\hat{p}_{(i,m)}$.

$$\text{minimize } \sum_{i=1}^{|P|} \sum_{m=1}^M zp_{(i,m)} + \sum_{i=1}^{|P|} \sum_{j=1}^{|Q|} \sum_{m=1}^M zpq_{(i,j,m)} \quad (4)$$

subject to:

$$\begin{cases} zp_{(i,m)} \geq 0, \\ zp_{(i,m)} \geq p_{(i,m)}xp_{(i)} - \hat{p}_{(i,m)}, & \forall i, \forall j, \forall m \\ zp_{(i,m)} \leq p_{(i,m)}xp_{(i)}, \\ \begin{cases} zpq_{(i,j,m)} \geq 0, \\ zpq_{(i,j,m)} \geq \hat{p}_{(i,m)} - q_{(j,m)} - p_{(i,m)}(1 - xpq_{(i,j)}), \\ M_{(i,j,m)} = \text{Max}(0, p_{(i,m)} - q_{(j,m)}), \\ zpq_{(i,j,m)} \leq \hat{p}_{(i,m)} - q_{(j,m)} + \dots \\ \dots + (M_{(i,j,m)} - lp_{(i,m)} - q_{(j,m)}) \cdot (1 - xpq_{(i,j,m)}), \\ zpq_{(i,j,m)} \leq M_{(i,j,m)}(xpq_{(i,j,m)}), \end{cases} & \forall i, \forall j, \forall m \end{cases} \quad (5)$$

$$\begin{cases} up_{(i,m)} = p_{(i,m)}, \\ lp_{(i,m)} = \text{Min}(p_{(i,m)}, \text{Min}(q_{(1..|Q|,m)})), \forall i, \forall m \\ lp_{(i,m)} \leq \hat{p}_{(i,m)} \leq up_{(i,m)}, \end{cases} \quad (6)$$

$$\{xp_{(i)} \geq xpq_{(i,j)}, \quad \forall i, \forall j \quad (7)$$

$$\left\{ xp_{(i)} \leq \sum_{j=1}^{|Q|} xpq_{(i,j)}, \quad \forall i \quad (8)$$

$$\left\{ \sum_{i=1}^{|P|} xpq_{(i,j)} = 1, \quad \forall j \quad (9)$$

$$\begin{cases} xp_{(i)} \in \{0, 1\}, \\ xpq_{(i,j)} \in \{0, 1\}, \\ xpq_{(i,j,m)} \in \{0, 1\}, \end{cases} \quad \forall i, \forall j, \forall m \quad (10)$$

where, $i = 1, \dots, |P|, j = 1, \dots, |Q|, m = 1, \dots, M$.

In the same way, the variable $zpq_{(i,j,m)}$ represents the difference that each $\hat{p}_{(i,m)}$ still have to improve to generate the $p'_{(i,m)}$ candidate in an attempt to be less than or equal to some $q_{(j,m)}$. The $zpq_{(i,j,m)}$ can be zero if $\hat{p}_{(i,m)}$ is equal to $p'_{(i,m)}$, or greater than zero if $\hat{p}_{(i,m)}$ is greater than $p'_{(i,m)}$. In fact, $zpq_{(i,j,m)}$ variable can be seen as a penalty value for those solutions that still do not dominate a Q_s .

The constraint set in Equation (6) ensures that if $\hat{p}_{(i,m)}$ is not less than or equal to $q_{(j,m)}$, then there is a valid value in $zpq_{(i,j,m)}$. It will receive $\max(0, \hat{p}_{(i,m)} - q_{(j,m)})$, if the difference to be less than $q_{(j,m)}$ or 0, otherwise. A binary variable, $xpq_{(i,j)}$ is used to reach out this model condition.

Thus, $xpqd_{(i,j,m)}$ assumes 1 to guarantee the maximum value; otherwise, the solution will be infeasible. Moreover, we have used a linearization technique to obtain the maximum function.

The $\hat{p}_{(i,m)}$ is a continuous variable and the constraint set expressed in Equation (7) presents the interval in which $\hat{p}_{(i,m)}$ lies within.

Constraints represented in Equations (8) and (9) associates the binary variables $xp_{(i)}$ and $xpq_{(i,j)}$ guaranteeing that if $xp_{(i)}$ is ‘active’, at least one $xpq_{(i,j)}$ would be ‘active’ as well. Equation (8) asserts that a candidate \hat{p}_i coming from p_i will try to weakly dominate Q_s , and in this sense, $xp_{(i)}$ will be equal to one and some $xpq_{(i,j)}$ will be also equal to one.

Finally, the constraint set in Equation (11) presents all the binary variables used in the model. The $xp_{(i)}$ is used to guarantee that if a $\hat{p}_{(i,m)}$ is used in the model, we will know exactly which $p_{(i,m)}$ has generated it. In the same way, $xpq_{(i,j)}$ indicates that p_i generated \hat{p}_i and is trying to weakly dominate a q_j . These two binary variables are used to guarantee that at least one p_i will be associated with a q_j .

Different MIP solvers can compute the proposed model such as CPLEX [24], GUROBI [25], and SCIP [26], just to name a few. However, there is no guarantee that the problems can be solved optimally in a non-prohibitive computational time. Just to have an idea of the model size regarding the number of variables and constraints, consider the problem having $P = \{(2.0, 2.5), (3.0, 1.9)\}$ and $Q = \{(2.2, 2.0), (3.1, 1.5)\}$, depicted in Figure 1. In this case, the model has 16 continuous variables, 22 binary variables, and 60 constraints (including non-negativity constraints). Considering a generic problem, the number of continuous variables, binary variables, and constraints are detailed in the Equations (12) to (14), respectively:

$$\# \text{ continuous variables} = (2 + |Q|)(|P|M), \quad (12)$$

$$\# \text{ binary variables} = (|P|(1 + |Q|(1 + 2M))), \quad (13)$$

$$\begin{aligned} \# \text{ constraints} &= |Q| + |P|(1 + 3M) + \dots \\ &\dots + (|P||Q|(3 + 4M)). \end{aligned} \quad (14)$$

As an example, considering $M = 5$, $|P| = |Q| = 200$, there are 202,002 continuous variables, 440,200 binary variables, and 923,400 constraints, making the MIP problem a relatively large-sized optimization problem.

V. EXPERIMENTS

A. Bi-objective case

The first test aims to show how MIP-DoM calculation addresses the four quality indicator facets: convergence, spread, uniformity, and cardinality [9]. The same experiments proposed in [10] to solve DoM in the bi-objective case are adopted to verify the model correctness³.

Figure 2 presents the controlled experiments for assessing the model correctness. Each facet of quality indicators, convergence, cardinality, uniformity, and spread, is represented in a ‘row’ in Figure 2. Each row has two plots corresponding to the two examples. The examples are slightly different from

one another. In each plot, there are two solution sets, P and Q .

Convergence is an important factor to reflect Pareto-dominance compliance of sets. This behavior is shown considering two examples in Figure 2, first row: *test = convergence*. In Examples *I* and *II*, the number of points in P and Q are equal, but some points of P dominate some points of Q . In plot *I*, P completely dominates Q , hence, dominance move from P to Q , $\text{DoM}(P,Q)$, is zero. In plot *II*, since some points of Q are dominated by P , $\text{DOM}(P,Q)$ is smaller than (Q,P) . There is still a movement sign (gray arrow) in plot *II* indicating the movement for one set to weakly dominate the other. Observe that q_4 needs to move to q'_4 , thereby weakly dominating P . On the same way, p_1 needs to move to p'_1 to weakly dominate Q . Importantly, these are the *minimum* movement needs to achieve a weak-dominance between the two sets.

DoM prefers solutions with a different cardinalities in P and Q . In Figure 2, *test = cardinality*, graphics *III* and *IV*, it can be observed that the solution sets have the same convergence, and spread. In graphic *III*, Q has one more point than P . In graphic *IV*, P has one more point than Q . The DoM values of both graphics reflected the cardinality aspect.

Uniformity indicates the preference for evenly distributed points. The solution sets in Figure 2, *test = uniformity*, presented this feature. The sets have the same convergence, spread, and cardinality. In graphics *V* and *VI*, the set P is distributed uniformly, and Q has a random distribution. The DoM values show that Q needs a bigger move value (DoM) to weakly dominate P . In graphic *VI*, the density of points in set Q increased gradually from bottom to top (considering the f_1). Again, a more uniformly distributed set of points needs a smaller DoM to move to a more non-uniformly distributed set of points.

Finally, DoM must exhibit its preference for solutions having a better spread. A set with better extensity (more extreme points) has a smaller dominance move compared to its competitor. In Figure 2, *test = spread*, considering graphic *VII*, set Q was generated by shrinking P a little (more concentrated in the middle). In graphic *VIII*, set P was distributed uniformly in the range, while Q assumed five right bottom points. The graphic *VIII* still presents the q_1 movement generating q'_1 to weakly dominate some elements from P , which were not yet dominated. On the same way p_1 , p_2 and p_4 , generate p'_1 , p'_2 and p'_4 to weakly dominate Q .

The MIP-DoM approach results are same as the results reported in [10] using the proposed bi-objective algorithm.

B. Multi- and Many-objective Problems

Following the initial and controlled test problems, it is crucial to validate the MIP-DoM model with standard test problem sets. First, a problem set with three objectives is compared with existing indicators, such as the additive ϵ -indicator, hypervolume (HV), IGD+, and also with visual graphics to assess the results. Second, an attempt to solve problems with five, ten and more objectives is made. In all tests, algorithms such as IBEA, MOEA/D, NSGA-III,

³We thank the authors for providing the data set.

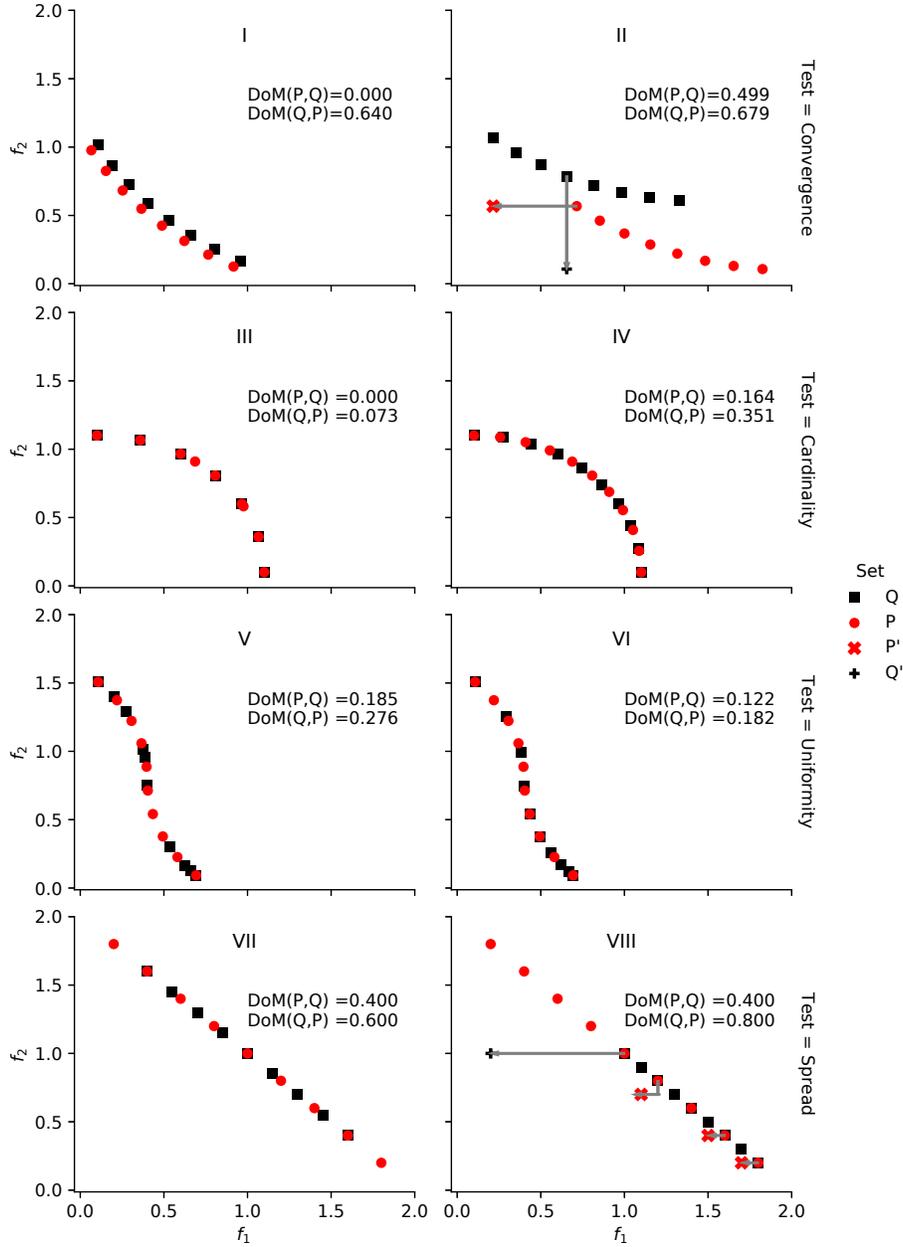


Figure 2: Experiments proposed in [10] to assess the four facets of quality indicators: convergence, cardinality, uniformity, and spread.

NSGA-II, and SPEA2 are used to generate the solution sets. It is important to highlight that the goal is to assess the effectiveness of the proposed MIP-DoM approach and not to compare the algorithm's performance, so any other algorithms could have been applied to generate the solution sets.

For each problem set, the population size for the algorithms needs to be initially chosen. We argue that this choice is crucial and is closely related not only to Equation (3) but also to the cardinality of the solution set. Considering a good approximation set of the Pareto front, in terms of convergence, spread, and uniformity, the number of non-dominated solutions grows exponentially concerning the dimension of the objective space. In [27], the shape of the Pareto front was discussed

in the niche size definition. A useful limit to the number of individuals in the population, given by $L = Mr^{M-1}$ is provided, where r is the resolution required or the number of points needed to represent the Pareto front. This expression makes it clear how $|P|$, or $|Q|$, must be increased as the problem dimension grows, showing an exponential relation between L and M . However, other works in many-objective optimization do not strictly follow this rule. In [28], for example, the number of objectives M is 3, 5, and 10, and the population size L is 105, 126, and 275, respectively. In [18], an efficient hypervolume calculation is provided and some tests are done with M set to 3, 4, 5 and L between 10 and 200. Likewise, in CEC'2018, a competition on many-objective

optimization [29] M was chosen to be 5, 10, and 15, and the maximum population size was set to 240.

Since Equation (3) plays an important role in the proposed model (Equations (4) to (11)), we have decided to validate DoM using $M = 3, 5, 10,$ and 15 and $|\mathcal{P}| = |\mathcal{Q}| = 50, 100, 170,$ and 240 indicating the final Pareto front approximation size.

All experiments are done using *Platypus* [30], *PyGMO* [31], and *pymoo* [32] to generate the problem sets and to calculate the IGD+ and ϵ -indicator. Hypervolume is calculated using the Walking Fish Group based algorithm [33] (considering its worst-case complexity as $O(M \times 2^L)$). The model (Equations (4) to (11)) is implemented using *Python* and it is solved using *GUROBI* [25] (version 9.0.0 build v9.0.0rc2) running on a Linux 64 bits operational system with 8 CPU's (Intel Xeon E5-2630 v4 2.2GHz) and 32Gb of RAM. Some *GUROBI* solver parameters [25] are altered: the minimum $gap = 10^{-6}\%$, and the $MIPFocus = 3$, which help to improve the best bound during the execution. For all models tested, the optimal solutions are always reached with the established gap.

1) *Multi-objective Problems*: A review of multi-objective benchmark problems is presented in [34]. To validate our proposed approach, we select some well-known problem test sets in three dimensions, which come from *DTLZ* and *WFG* families. Some test sets are initially chosen based on some characteristics: convexity/concavity, disconnection, multimodality, and degeneracy. *DTLZ1*, *DTLZ2*, *DTLZ3*, and *DTLZ7*, which are linear, concave/multimodality, concave, and disconnected, respectively, are selected. In the same manner, *WFG1*, *WFG2*, *WFG3*, and *WFG9* are also chosen to take into account the properties: convex/mixed, convex/disconnected, linear/degenerate, and concave, respectively.

Some algorithms are used to generate the approximated Pareto front, and using the outcomes, the MIP-DoM indicator is compared to other quality indicators. Embedded with a similar purpose of [11] and [35], the *NSGA-III* and *MOEA/D* algorithms (Pareto-based and decomposition approach, respectively) are firstly chosen, and, afterwards, *IBEA* [36], *SPEA2* [37], and *NSGA-II* [19]. All the approximation sets are shown in Figure 3.

In our experiments, the maximum number of fitness evaluations is set to 10,000, and each algorithm is executed 21 times. Three unary quality indicators, the additive ϵ -indicator, hypervolume (HV), and the inverted generational distance plus (IGD+), are calculated for each run. It is mandatory to have a reference point or a reference set to calculate the indicators, and this task is a challenging one [9], [23], but is sometimes provided by the user [18]. We use the maximum values amongst all algorithm solutions for HV. For additive ϵ -indicator and IGD+, the reference set is a joint Pareto front, which comes from the algorithm's results (it can be viewed in the last column of Figure 3).

The MIP-DoM quality measure is a binary indicator. Using MIP-DoM as a unary indicator is still feasible; a straightforward idea is merely using a joint Pareto (created from other algorithm solution sets) or the real Pareto front. The MIP-DoM is then calculated using the joint Pareto as a reference set for

each problem. We use this approach to facilitate a comparison with some unary indicators, such as HV and IGD+.

Table I shows average values of three unary quality indicators, the additive ϵ -indicator, hypervolume (HV), and the inverted generational distance plus (IGD+). MIP-DoM has also been calculated. In this table, the algorithms are sorted taking into account the ascending order of MIP-DoM results. It can

Table I: MIP-DoM(\mathcal{P}, \mathcal{Q}) values for the *DTLZ* and *WFG* families for comparison among *IBEA*, *MOEA/D*, *NSGA-III*, *NSGA-II*, and *SPEA2* algorithms. It must be noted that \mathcal{P} is the solution set generated by the algorithm and \mathcal{Q} is the non-dominated solution set from all algorithm's results combined. HV, IGD+, and additive ϵ -indicator are also presented. The algorithms are sorted according to worse performance by MIP-DoM.

<i>Problem</i>	<i>Algorithms</i>	MIP-DoM	HV	IGD+	additive ϵ-indicator
DTLZ1	<i>IBEA</i>	0.312	0.999	0.089	0.001
	<i>NSGA-III</i>	0.551	0.999	0.032	0.001
	<i>MOEA/D</i>	1.630	0.999	0.052	0.009
	<i>NSGA-II</i>	6.422	0.998	0.041	0.027
	<i>SPEA2</i>	11.139	0.984	0.089	0.064
DTLZ2	<i>MOEA/D</i>	0.970	0.641	0.065	0.089
	<i>IBEA</i>	1.000	0.632	0.074	0.130
	<i>NSGA-III</i>	1.004	0.641	0.065	0.106
	<i>NSGA-II</i>	1.013	0.607	0.076	0.193
	<i>SPEA2</i>	1.021	0.595	0.082	0.162
DTLZ3	<i>IBEA</i>	0.0491	0.999	0.129	0.000
	<i>NSGA-III</i>	18.652	0.999	0.017	0.045
	<i>MOEA/D</i>	25.374	0.998	0.077	0.065
	<i>NSGA-II</i>	51.208	0.984	0.099	0.142
	<i>SPEA2</i>	150.870	0.706	0.342	0.369
DTLZ7	<i>NSGA-III</i>	1.402	0.198	0.085	0.175
	<i>IBEA</i>	1.468	0.210	0.064	0.082
	<i>MOEA/D</i>	1.695	0.196	0.073	0.123
	<i>NSGA-II</i>	1.782	0.181	0.071	0.167
	<i>SPEA2</i>	2.184	0.141	0.112	0.185
WFG1	<i>IBEA</i>	1.230	0.342	0.192	0.257
	<i>MOEA/D</i>	1.557	0.371	0.105	0.243
	<i>SPEA2</i>	1.659	0.256	0.289	0.425
	<i>NSGA-III</i>	1.822	0.286	0.248	0.326
	<i>NSGA-II</i>	1.944	0.343	0.134	0.272
WFG2	<i>IBEA</i>	1.503	0.864	0.054	0.110
	<i>NSGA-III</i>	1.571	0.824	0.062	0.201
	<i>MOEA/D</i>	1.683	0.834	0.057	0.134
	<i>NSGA-II</i>	1.687	0.790	0.075	0.184
	<i>SPEA2</i>	1.859	0.777	0.080	0.216
WFG3	<i>IBEA</i>	1.889	0.410	0.060	0.111
	<i>NSGA-III</i>	2.609	0.383	0.069	0.264
	<i>NSGA-II</i>	2.630	0.392	0.060	0.241
	<i>MOEA/D</i>	2.820	0.375	0.083	0.282
	<i>SPEA2</i>	3.178	0.373	0.079	0.211
WFG9	<i>IBEA</i>	1.965	0.312	0.085	0.156
	<i>NSGA-III</i>	2.194	0.324	0.074	0.095
	<i>MOEA/D</i>	2.331	0.290	0.089	0.161
	<i>NSGA-II</i>	2.601	0.304	0.074	0.123
	<i>SPEA2</i>	2.759	0.289	0.085	0.171

be noted from the table that *DTLZ1* and *DTLZ3* have the HV values inflated by the presence of the dominance resistant solutions [9], which are non-dominated solutions with a poor value in one objective but with good values in others.

For the *DTLZ1*, the algorithms presenting the best HV, ϵ -indicator, and MIP-DoM values are *IBEA* and *NSGA-III*. For the HV indicator, the comparison is difficult due to inflated

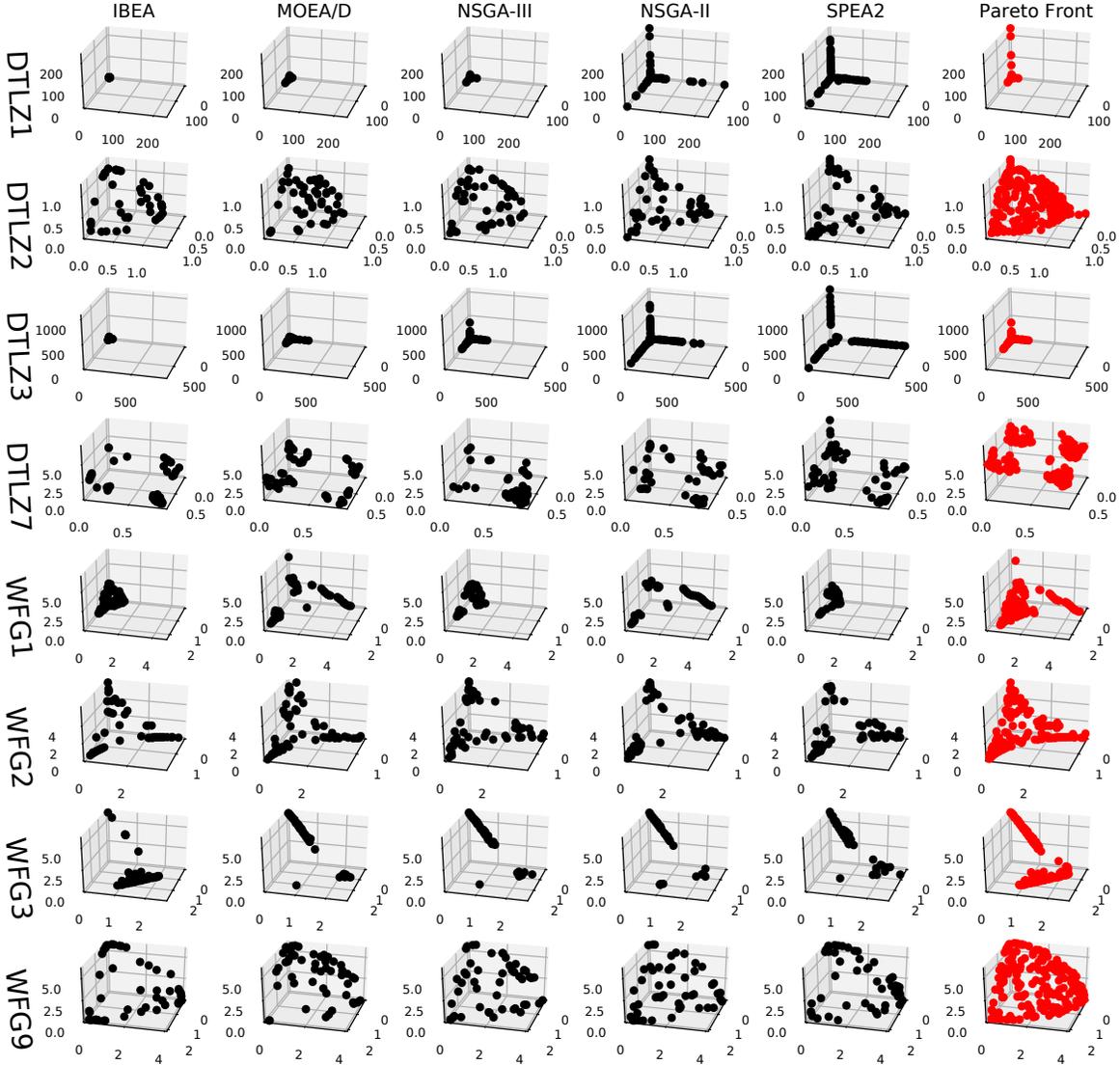


Figure 3: Solution sets with $|P| = |Q| = 50$ solutions and three objectives, $M = 3$, generated by IBEA, MOEA/D, NSGA-III, NSGA-II, and SPEA2 algorithms applied to DTLZ1, DTLZ2, DTLZ3, WFG1, WFG2, WFG3 and WFG9 problem sets. The Pareto front is generated by the collecting non-dominated solutions from all algorithms.

values. For IGD+, the results indicate NSGA-III and NSGA-II are the best algorithms. For MIP-DoM, the top two algorithms are IBEA and NSGA-III with 0.312 and 0.551, respectively. In the DTLZ2 problem, HV, IGD+, ϵ -indicator, and MIP-DoM indicate MOEA/D as one of the best results. For HV and IGD+, there is a tie between NSGA-III and MOEA/D. DTLZ3 results for ϵ -indicator, HV, and MIP-DoM classify IBEA and NSGA-III as the best algorithms. Again, considering HV, due to some extreme points, the values are next to each other. The best algorithms are indicated as IBEA, NSGA-III, and MOEA/D (there is a full agreement between MIP-DoM and HV, considering algorithms' ranking). For IGD+, there is an indication of NSGA-III and MOEA/D as the best ones. Finally, for the DTLZ7 problem set, IBEA is pointed as the best one for HV, IGD+, and ϵ -indicator. However, NSGA-III is the best one, according to MIP-DoM, and the second one considering

HV.

The same experiment is done for the *WFG* family. For WFG1 in Table I, the best algorithm is MOEA/D for HV, IGD+, and ϵ -indicator. For MIP-DoM, the best algorithms are IBEA and MOEA/D. For WFG2, the best values for HV, IGD+, and ϵ -indicator and MIP-DoM indicate IBEA as one of the best algorithms. HV, IGD+, and MIP-DoM rank these three algorithms as the best ones: IBEA, MOEA/D, and NSGA-III. In the same way, for WFG3, IBEA generate the best solution sets for all indicators. Again, there is an agreement amongst them. For IGD+, the best algorithms are IBEA and NSGA-II, presenting a tie between these two algorithms. Finally, for the WFG9 problem set, considering HV, IGD+, and additive ϵ -indicator NSGA-II is the best algorithm. MIP-DoM point out IBEA and NSGA-III as the best ones. For IGD+, there is a tie between NSGA-III and NSGA-II.

Correlation with Existing Indicators: The goal of this paper is to introduce the MIP-DoM as an alternate performance indicator. Next, we calculate Pearson correlation coefficient of the MIP-Dom indicator with HV, IGD+ and additive ϵ -indicator. A Pearson correlation coefficient (ρ) is calculated in Table II to assess if there is a linear relation among the indicators. We measure the correlation coefficient using the problem sets individually, the problem set families, and the general correlation among all indicators. To assess the statistical significance of the results, a hypothesis test with a significance level of 0.05 is carried out. Values which are not statistically significance are indicated as * in Table II. All values are normalized before calculation the correlation coefficients. For simplicity, we use the negative HV for calcu-

Table II: The Pearson correlation coefficient (ρ) to assess the linear relationship between MIP-DoM and other indicators (HV, IGD+, and ϵ -indicator).

Problem set	MIP DoM correlation with		
	-HV	IGD+	additive ϵ -indicator
DTLZ1	0.896	0.347*	0.989
DTLZ2	0.770*	0.769*	0.783*
DTLZ3	0.964	0.887	0.998
DTLZ7	0.942	0.720*	0.502*
Combined DTLZ	0.895	0.681	0.818
WFG1	0.265*	0.019*	0.256*
WFG2	0.860	0.815*	0.626*
WFG3	0.945	0.747*	0.684*
WFG9	0.627*	0.093*	0.226*
Combined WFG	0.674	0.372*	0.448
Combined All	0.784	0.526	0.633

* We applied a hypothesis test for the correlation coefficient with a significance level of 0.05. The * shows that the calculated p-value is over 0.05, and it is not possible to reject the null hypothesis that $\rho = 0$.

lating the correlation, since it is the only indicator for which a larger value means better. It is observed that there is reasonably strong correlation between MIP-DoM and HV, and between MIP-DoM and additive ϵ -indicator taking into account all problems. Concerning the DTLZ family, the correlation between MIP-DoM and HV and additive ϵ -indicator is even stronger, 0.90 and 0.82, respectively. For the WFG family, the correlations between MIP-DoM and HV and additive ϵ -indicator is different from the DTLZ family. We argue that WFG1 is responsible for the decrease in the correlation values since the correlation between MIP-DoM and HV is high for WFG2, and WFG3.

In general, the correlation between MIP-DoM and HV is high for most problems. This gives us confidence of its use as an alternate performance indicator for set-based multi-objective optimization algorithms.

2) *Many-Objective Problems:* Next, the goal is to verify if the MIP-DoM approach can be applied in many-objective scenarios. Inspired by the issues raised in the multi-objective experiments and the well-known quality indicators' weaknesses, our motivation to apply DoM in many-objective scenarios can be summed up as:

- 1) HV is hard to exactly calculate in many-objective problems [38]. Yet, it is sensitive to extreme points and

dominance resistant solutions, such as observed in the last experiment for DTLZ1 and DTLZ3 cases;

- 2) In many-objective problem sets and real cases, it is hard to obtain a Pareto front or reference set that is evenly distributed such as vital to IGD+, for example; The same happens to HV, in [23], for example, some experiments have shown that a slightly worse point than the nadir point is not always appropriate for problem sets comparison;
- 3) Considering the DoM definition and its calculation proposal, it is still relevant to note that there is no information loss in DoM definition. This is an essential feature to a quality indicator, mainly considering problem sets with high number of objective functions and cardinality.

In this sense, we would like to observe how the MIP-Dom model behaves in such scenarios, and analyze its characteristics. The same problem sets from the previous experiment are used, and two algorithms are chosen to compare and observe the quality indicator features. We choose MOEA/D and NSGA-III to generate the approximation sets. It is relevant to emphasize that any other algorithm could have been selected, however, the rationale behind this choice is due to specificity of these algorithms to deal with many-objective problems. The object of our study here is the quality indicator, and the aim is to assess how it behaves with different many-objective test cases.

In the same manner as the CEC'2018 competition [29] we establish $L = 100, 170, \text{ and } 240$ and $M = 3, 5, 10, \text{ and } 15$, respectively. In Table III the MIP-DoM values are presented.

Table III: Using the solution sets generated by MOEA/D and NSGA-III algorithms, MIP-DoM value for the many-objective experiments for *DTLZ* and *WFG* families. The number of points in the final non-dominated set is set to $L = 100, 170, \text{ and } 240$, and the number of objectives, $M = 5, 10, \text{ and } 15$.

Problem	MIP-DoM(P, Q)	Value		
		$L=100$ $M=5$	$L=170$ $M=10$	$L=240$ $M=15$
DTLZ1 (MOEA/D, NSGA-III)		1463.461	560.182	594.894
	(NSGA-III, MOEA/D)	3401.332	6315.371	9944.266
DTLZ2 (MOEA/D, NSGA-III)		3.943	1.175	1.065
	(NSGA-III, MOEA/D)	3.804	8.972	16.665
DTLZ3 (MOEA/D, NSGA-III)		2839.942	1878.871	2040.851
	(NSGA-III, MOEA/D)	6418.077	15532.120	23735.510
DTLZ7 (MOEA/D, NSGA-III)		2.502	5.543	4.745
	(NSGA-III, MOEA/D)	1.148	2.948	2.950
WFG1 (MOEA/D, NSGA-III)		0.270	0.279	0.124
	(NSGA-III, MOEA/D)	0.464	0.707	0.775
WFG2 (MOEA/D, NSGA-III)		2.001	1.179	1.137
	(NSGA-III, MOEA/D)	1.571	2.423	3.665
WFG3 (MOEA/D, NSGA-III)		4.524	4.678	1.955
	(NSGA-III, MOEA/D)	5.170	14.160	22.981
WFG9 (MOEA/D, NSGA-III)		3.183	2.177	0.524
	(NSGA-III, MOEA/D)	4.450	13.583	31.296

The algorithm solution sets are compared in pairs and both directions. For DTLZ1 with $L = 100$ and $M = 5$, the DoM(MOEA/D, NSGA-III) = 1463.461 and the DoM(NSGA-III, MOEA/D) = 3401.332, indicating that the approximation set generated by MOEA/D is better than that by NSGA-III. For DTLZ2, while DoM(NSGA-III, MOEA/D) is smaller than DoM(MOEA/D, NSGA-III), but for higher objectives, MOEA/D's performance is better. From the table, we observe

that MIP-DoM has an agreement about the best algorithm for some problem sets, independent to the parameter M and L . In DTLZ1, DTLZ3, WFG1, WFG3, and WFG9, for example, MOEA/D set is better than that in NSGA-III. On the other hand, for DTLZ7, NSGA-III provides lower values. There is a disagreement found in DTLZ2 and WFG2 for $L = 100$ and $M = 5$. Furthermore, for the DTLZ2, for example, the values are similar for both algorithms. However, our goal here is to show that it is possible to use and calculate MIP-DoM regardless the number of points (L) in the solution set and the number of objective functions (M) and still compare two or more EMaO algorithms.

One interesting fact observed in the experiments is related to the number of p'_i different from the original p_i . The minimum number in this case was one, thus the MIP-DoM recommendation is to change only one point. This is always the case behind ϵ -indicators, which exhibit information loss, as previously discussed. On the other side, the maximum number is 9, 8, and 26 in the experiments for $L = 100$ and $M = 5$, $L = 170$ and $M = 10$, and $L = 240$ and $M = 15$, respectively. This fact denotes that MIP-DoM value is more suitable for many-objective scenarios.

VI. FURTHER STUDIES WITH MIP-DoM

Having compared MIP-DoM indicator with popular existing indicators on multi- and many-objective optimization problems, we now analyze and reveal several other properties of the proposed MIP-DoM indicator.

A. Computational Time

Every MIP-DoM computation requires solving a mixed-integer programming (MIP) problem stated in Equations (4) to (11). In cases in which most points of Q are weakly-dominated by some point in P exist, the number of required MIP iterations is small and the computational time to solve the MIP problem is also small. However, in our experiment, the worst-case scenario spends approximately ~ 208 hours, for example, using MOEA/D to solve WFG1 with $L = 240$ and $M = 15$. In this case, the pre-processing method is not feasible, and the number of simplex iterations is very large, as well as the number of explored nodes.

The time spent in seconds by the solver is an important aspect to analyze. Figure 4 presents an experiment using DTLZ1 problem, as an example. We generate 50 random approximation set as P and the Pareto front as Q to calculate $MIP-DoM(P, Q)$ (the distance of movement which a random approximation set needs to dominate the Pareto front). Our intention is to assess how the computational time increases with an increase in number of population size for 5, 10, and 15 objective problems. In Figure 4, the y axis is in log scale, and the x axis is represented with $L = 50, 100, 200, 300$ and 400 solutions.

In Figure 4, considering $M = 5, 10$ and 15, we can observe a polynomial time behavior for DTLZ1 considering the solutions number ($\approx O(L^{3.311})$, $O(L^{2.888})$ and $O(L^{2.833})$, respectively).

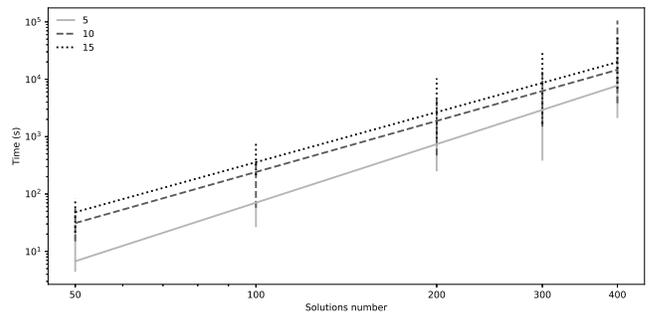


Figure 4: A polynomial increase in computational time with population size. Fifty experiments are done with $L = 50, 100, 200, 300$ and 400 solutions using the DTLZ1 problem. The data set is randomly generated, and the objective function is analytically calculated.

Another aspect to verify is how the time behaves if M is increased for many-objective problems. In Figure 5, using the DTLZ1, values of $M = 7, 10, 12, 15, 20, 25$ and 30 are adopted and we generate random solutions, and calculate its objective functions using its analytical function. The Pareto front is generated, and our experiments measure the time spent to calculate $MIP-DoM(P, Q)$, with random approximation set as P , and the Pareto front as Q . For each M , fifty experiments are run with $L = 200$ solutions in the approximation set. In Figure 5, both axes are in log scale, representing a polynomial time behavior ($\approx O(M^{0.686})$). Each box-plot graph is a test set with its associated M . The figure reveals that the MIP-DoM

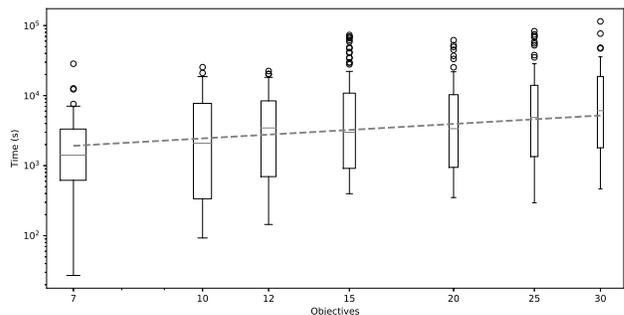


Figure 5: A boxplot with the time spent to solve the MIP-DoM model as a function of number of objectives. For each box-plot, fifty experiments are done with $L = 200$ using the DTLZ1 problem set definition. The data set is randomly generated, and the objective function is analytically calculated.

indicator can be useful for large number of objectives, as it demands a small increase in computational time for an unit increase in number of objectives.

A burdensome scenario for MIP-DoM is when there is no a priori dominance between the points in each solution set, and the density of the two solution sets used are high. In areas with high density, there are many options among p_i trying to dominate some Q_s . The branch-and-bound dynamic becomes more challenging in these situations, combining the \hat{p} possibilities, between its lower and upper bound, and the

binary variables. This fact must be better investigated and exploited in a future work.

B. MIP-DoM as a Running Performance Indicator

Popular performance metrics (such as, GD, IGD, IGD+, and also HV in some sense) requires the knowledge of true Pareto front so that a reference set can be obtained. This severely restrict their usage in arbitrary problems. But, a recently proposed running performance metric method [39] can be used for handling an arbitrary problem. At a generation t , all non-dominated (ND) solutions from all past generations from the initial population is first collected and the resulting ND set is used as the reference set for metric computation. Then, at generation $t + \Delta t$, the whole process can be repeated (Δt can be suitably chosen) and the time varying plot can be made to indicate how the metric changes with generation.

In [39], IGD+ is used as an example of the running quality indicator. The IGD+ indicator demands a reference set, such as the true Pareto front. We show the use of running metric for two purposes: (i) termination criteria in a pair algorithm comparison and (ii) a detailed algorithm behavior view in the quality indicator perspective. Our goal is to verify if MIP-DoM can be used as a running quality indicator, and what are their characteristics. Figure 6 presents six curves. We calculate each one using the solution set from each generation as P and, at the generation 10, 20, 30, 40, and 50, solution sets from its generation as Q in the $MIP-DoM(P, Q)$ calculation. The final solid line is computed using the known Pareto front solutions, to demonstrate the behavior of the running indicator.

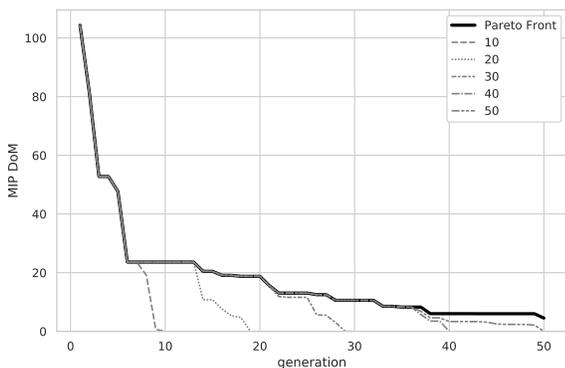


Figure 6: MIP-DoM as a running performance quality indicator. The data refers to DTLZ1 problem set, with two objectives using NSGA-II algorithm with 50 generations. We calculate MIP-DoM using P as each generation solution set and Q as solution sets at generations 10, 20, 30, 40, 50 and the final Pareto front.

The figure can be analysed from different manners. First, there is a monotonic decrease in each MIP-DoM curve, indicating that MIP-DoM is able to correctly indicate the monotonic improvement of IGD+ with generations. It is not surprising that each intermediate running indicator curve drops to zero at the pivot generation at which the indicator was calculated. Since the population at this pivot generation is expected to be best compared to earlier generation populations,

$MIP-DoM(P_t, P_t) = 0$. Second, consider the two consecutive running indicator curve (say at generations 30 and 40). The DoM value calculated at generation 30 using 40-th generation ND set is higher than that calculated using 30-th generation ND set. This indicates that 40-th generation ND set is better than 30-th generation ND set. Third, notice that for up to about 35 generations, DoM calculated using the Pareto front is almost equal to DoM calculated using ND set at generation 50. By checking with a threshold value of MIP-DoM, a termination of an EMO/EMaO run can be adopted.

Another use of the running MIP-DoM indicator can come in comparing two or more EMO/EMaO algorithms generation by generation. In Figure 7, a comparison between NSGA-II and NSGA-III is made. We calculate $MIP-DoM(NSGA-II, \text{combined solutions})$ in a dashed line and $MIP-DoM(NSGA-III, \text{combined solutions})$ in a solid line. The combined solutions represent the non-dominated solutions from NSGA-II and NSGA-III results at generation 50. We use the DTLZ1 with two-objective functions and 30 solutions. Each point in each algorithm curve indicates the MIP-DoM value at the specific generation calculated using the combined solutions. We can observe a similar monotonic drop behavior. In the beginning, until generation 4, NSGA-II presents a better value than NSGA-III. From generation 8 onward, there is a change, and NSGA-III is better than NSGA-II.

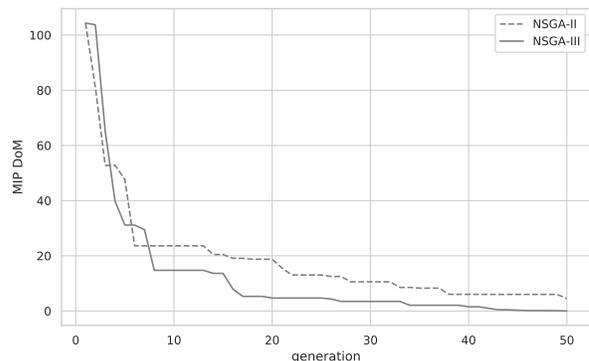


Figure 7: An algorithm comparison between NSGA-II and NSGA-III using the MIP-DoM running quality indicator. The data refers to DTLZ1 problem set, with two objective functions and 30 solutions in each set. We calculate MIP-DoM using P as each generation solution set and Q as a joint solution set(non-dominant solutions) from the NSGA-II and NSGA-II 50 generation.

These plots indicate how the MIP-DoM can be employed as a running quality indicator, as termination criteria and as a comparison indicator without the knowledge of the true Pareto solutions.

C. Parametric Study with Gap in Computing MIP-DoM

In the branch-and-bound formulation, some information can be obtained using the function value of the linear programming relaxation and the function value of the integer program. A valid upper bound on the optimal solution is a proper integer solution with the best value found in the branch-and-bound

procedure in a minimization problem. On the opposite side, we also can have a valid lower bound during the branch-and-bound. It comes from the linear programming relaxation, taking the minimum of all current leaf nodes' optimal objective values. The difference between the upper and lower bounds is known as *gap*. In general, when the *gap* = 0, we have reached optimality.

The question raised is whether MIP-DoM is affected by a gap variation, such as *gap* = 5% or 10%? To investigate this issue, we repeat the same experiment in the multi-objective section, in which the Pearson correlation coefficient between MIP-DoM and some indicators are calculated. The idea is to observe how the gap will influence the MIP-DoM results and how a change affects the correlation coefficient.

We vary the gap parameter, such as posed in our initial question. In Table IV, not only the original results are presented, but also the results with *gap* = 5% and 10% are added. We can observe that the gap parameter introduces a small perturbation in the Pearson correlation coefficient. For example, the HV has a change in the correlation coefficient of less than one percent comparing the values at *gap* = 0% and 10%. Similar behavior happens to IGD+ and additive ϵ -indicator.

On the other hand, when we analyze the effect in the model's time, we can observe a clear impact in all cases. In Table IV, the mean and standard deviation (std deviation) of times for each problem set are presented. The gap has good impacts on the time spent. In some cases, the difference lies in one order of magnitude.

Additionally, there are differences in the time spent amongst the problem sets. At this experiment, it must be noted that the reference set has different sizes (number of solutions in the joint Pareto), and this influences the time spent by the model, as discussed before. For DTLZ1, DTLZ2, DTLZ3 and DTLZ7 are 79, 207, 93 and 154 solutions, respectively. Considering WFG1, WFG2, WFG3, and WFG9, the values are 104, 124, 151, and 179. Even considering just one problem set, the times spent by MIP-DoM still can show relevant variations among the algorithms. In our experiments IBEA and NSGAIII present the highest times, considering the DTLZ2 and WFG3.

The gap information can help in the binary quality indicator context. If one is interested in a provably optimal solution, it is necessary to wait until the gap reaches zero. Usually, in the quality indicator context, optimality is not an issue and "*P* is better than *Q*" is the only question to be answered. Moreover, the faster the answer, the better. Motivated by these issues, different gaps are analysed and the time spent to decide whether "*P* is better than *Q*" is computed in each case. Considering Table III in the many-objective experiments, some high times are also found, for example, DTLZ7 and WFG1 for MOEA/D with $L = 240$ and $M = 15$. For these cases, and in the context of a binary quality indicator, it is not mandatory to obtain an optimal solution, providing the statement is valid. Taking, for example, DTLZ7 for NSGA-III and MOEA/D with $L = 240$ and $M = 15$, it took $\sim 1,818$ seconds to calculate MIP-DoM(NSGA-III, MOEA/D) = 4.475. Similarly, the solver took $\sim 198,795$ seconds to calculate MIP-DoM(MOEA/D, NSGA-III) = 2.950. However, the solver presented a MIP-DoM valid interval search between 6.188 and 3.101, with

gap = 50% at 98,294 seconds. With this interval, it is possible to assert that NSGA-III is better than MOEA/D, even before the branch-and-bound conclusion.

VII. CONCLUSIONS

The DoM binary quality indicator considers the minimum move of one set to weakly dominate the other set. The indicator is Pareto compliant and does not demand any parameters or reference sets. The handicap about DoM is its requirement to solve a mixed-integer programming (MIP) problem, making it apparently time consuming to be applied to problems with many objectives and a large population size.

In this paper, we have provided a detailed MIP formulation of the DoM calculation indicating how the number of variables increase with the number of objectives and population size. The results of our approach have been verified with the original proposal on artificial bi-objective examples. The results have also make the sensitivity of the DoM indicator on dominance, diversity, cardinality, and other issues related to two non-dominated set comparison.

We have also reported that the MIP-DoM indicator values are highly correlated with other common performance indicators, such as hypervolume (HV), IGD+ and additive ϵ -indicator. The proposed indicator is more correlated with the HV indicator. MIP-DoM indicator has been successfully applied to indicate the performance on three to 15-objective test problems from DTLZ and WFG test suites with population sizes varying from 30 to 400, making an extensive evaluation of the MIP-DoM indicator.

Based on this extensive study, we have observed that MIP-DoM brings following advantages as a binary quality indicator:

- It does not require any pre-defined set of points, such as, a reference point or a reference set;
- It does not demand any normalization of the objective functions;
- It is not affected by dominance resistant solutions, unlike that in HV;
- It considers all the solutions of both sets being used. There is no lack of information.
- Presumably, MIP-DoM is compatible in indicating four performance facets between two sets: convergence, spread, uniformity, and cardinality;
- It appears to have a highly Pearson correlation with HV;
- It presents a monotonic decrease in its value, when the first set (*P*) is fixed and the second set (*Q*) approaches towards the Pareto front;

However, MIP-DoM quality indicator still present some drawbacks:

- MIP-DoM time calculation is polynomial to number of variables and objectives, and is more affected by the number of solutions in the set;
- Although it can be used as a running quality indicator and presents a correct ranking among curves, in some cases, the difference in MIP-DoM values are in decimal places, thereby making it difficult to visualize the difference;
- It demands an efficient MIP solver in its calculation.

Table IV: The first part of the table presents the Pearson correlation coefficient between MIP-DoM and other indicators. The coefficient is re-calculated with the new MiP-DoM value considering $gap = 5\%$ and 10% . The second part of the table is focusing on the time spent by the model for the given gap. We show the mean and standard deviation for each problem set and gap.

Problem set	MIP DoM correlation with									Time spent (seconds)					
	-HV			IGD+			additive ϵ -indicator			mean		std deviation			
	gap			gap			gap			gap		gap			
	0% ¹	5%	10%	0% ¹	5%	10%	0% ¹	5%	10%	0% ¹	5%	10%	0% ¹	5%	10%
DTLZ1	0.896	0.896	0.896	0.347*	0.347*	0.347*	0.989	0.989	0.989	2.4e+01	2.5e+01	1.4e+01	1.2e+01	2.9e+01	1.2e+01
DTLZ2	0.770*	0.773*	0.770*	0.769*	0.763*	0.769*	0.783*	0.784*	0.783*	2.8e+05	2.0e+05	1.3e+05	3.9e+05	3.3e+05	2.6e+05
DTLZ3	0.964	0.964	0.963	0.887	0.887	0.886	0.998	0.998	0.998	3.8e+01	2.8e+01	1.6e+01	4.7e+01	2.9e+01	1.8e+01
DTLZ7	0.942	0.942	0.942	0.720*	0.720*	0.720*	0.502*	0.502*	0.502*	1.6e+04	3.1e+03	9.9e+02	2.8e+04	3.3e+03	8.3e+02
Combined DTLZ	0.895	0.894	0.895	0.681	0.679	0.681	0.818	0.818	0.818	7.5e+04	5.1e+04	3.3e+04	2.2e+05	1.7e+05	1.3e+05
WFG1	0.265*	0.250*	0.242*	0.019*	0.003*	0.002*	0.256*	0.242*	0.224*	3.3e+02	1.6e+02	8.7e+01	3.6e+02	1.2e+02	6.1e+01
WFG2	0.860	0.877*	0.877*	0.815*	0.833*	0.833*	0.626*	0.631*	0.631*	3.6e+02	2.0e+02	1.2e+02	3.0e+02	1.3e+02	7.4e+01
WFG3	0.945	0.945	0.946	0.747*	0.747*	0.747*	0.684*	0.684*	0.684*	1.6e+04	2.2e+03	7.4e+02	1.7e+04	1.3e+03	2.8e+02
WFG9	0.627*	0.627*	0.649*	0.093*	0.093*	0.009*	0.226*	0.226*	0.333*	9.9e+02	4.7e+02	2.4e+02	1.1e+03	3.4e+02	1.2e+02
Combined WFG	0.674	0.675	0.678	0.372*	0.372*	0.398*	0.448	0.446	0.468	4.4e+03	7.5e+02	3.0e+02	1.1e+04	1.1e+03	3.1e+02
Combined All	0.784	0.784	0.786	0.526	0.526	0.539	0.633	0.632	0.643	4.0e+04	2.6e+04	1.7e+04	1.6e+05	1.2e+05	9.2e+04

¹ In fact, the 0% means a $gap = 10^{-6}\%$.

* We use a hypothesis test for the correlation coefficient with a significance level of 0.05. The * shows that the calculated p-value is over 0.05, and it is not possible to reject the null hypothesis that $\rho = 0$.

Despite this extensive study on analyzing properties of the MIP-DoM indicator, there are a number of future studies which are worth pursuing. For some experiments in our study, we observed a high variability in the computational time from generation to generation for the same problem size and population size. It was observed that this fact was inherent to the distribution or density, and some inner characteristics involving the two solution sets, P and Q . This issue is something that deserves to be better investigated in order to improve the MIP model, creating, for example, new constraints in the MIP model.

In the MIP approach for DoM, other directions also deserve to be investigated: i) How to efficiently calculate DoM using MIP and exploiting some inherent solution set features, such as solution set density; ii) How to define a priori the minimum number of $p_i \neq p'_i$ and what benefits it can bring to the MIP model; iii) How to improve the convergence time controlling the ill-conditioning of the coefficient matrix induced by the problem formulation presented in IV. Such questions have been posed here to emphasize possible future research directions.

Finally, besides its use as quality indicator, MIP-DoM generates a new solution set which weakly dominates another set, P' . This information can be used within genetic operators to create better solutions for a faster convergence. All these are interesting extensions, nevertheless, we have revealed useful features of the MIP-DoM indicator which should encourage EMO/EMaO researchers to pay more attention to it.

ACKNOWLEDGEMENTS

Flávio Martins and Elizabeth Wanner would like to thank the support from the Brazilian funding agencies: CAPES, FAPEMIG and CNPq. Flávio Martins also acknowledges the support from Michigan State University (MSU) (under Koenig Endowed Chair grant of Prof. Kalyanmoy Deb) for his visit to MSU.

REFERENCES

- [1] T. Chugh, K. Sindhya, J. Hakanen, and K. Miettinen, "A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms," *Soft Comput.*, vol. 23, no. 9, pp. 3137–3166, May 2019. [Online]. Available: <https://doi.org/10.1007/s00500-017-2965-0>
- [2] D. L. González-Álvarez and M. A. Vega-Rodríguez, "Analysing the scalability of multiobjective evolutionary algorithms when solving the motif discovery problem," *Journal of Global Optimization*, vol. 57, pp. 467–497, 2013.
- [3] K. Deb, "Scope of stationary multi-objective evolutionary optimization: A case study on a hydro-thermal power dispatch problem," *Journal of Global Optimization*, vol. 41, pp. 479–515, 08 2008.
- [4] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," *ACM Comput. Surv.*, vol. 48, no. 1, Sep. 2015. [Online]. Available: <https://doi.org/10.1145/2792984>
- [5] A. Ibrahim, S. Rahnamayan, M. V. Martin, and K. Deb, "3d-radviz antenna: Visualization and performance measure for many-objective optimization," *Swarm and Evolutionary Computation*, vol. 39, pp. 157–176, 2018. [Online]. Available: <https://doi.org/10.1016/j.swevo.2017.09.011>
- [6] K. S. Bhattacharjee, H. K. Singh, and T. Ray, *Many-Objective Optimization with Limited Computing Budget*. Cham: Springer International Publishing, 2020, pp. 17–46. [Online]. Available: https://doi.org/10.1007/978-3-030-18764-4_2
- [7] A. K. A. Talukder and K. Deb, "Paletviz: A visualization method for functional understanding of high-dimensional pareto-optimal data-sets to aid multi-criteria decision making," *IEEE Computational Intelligence Magazine*, vol. 15, no. 2, pp. 36–48, 2020.
- [8] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. da Fonseca, "Performance assessment of multiobjective optimizers: an analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, April 2003.
- [9] M. Li and X. Yao, "Quality evaluation of solution sets in multiobjective optimisation: A survey," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 26:1–26:38, Mar. 2019. [Online]. Available: <http://doi.acm.org/10.1145/3300148>
- [10] —, "Dominance move: A measure of comparing solution sets in multiobjective optimization," *CoRR*, vol. abs/1702.00477, 2017. [Online]. Available: <http://arxiv.org/abs/1702.00477>
- [11] C. L. d. V. Lopes, F. V. C. Martins, and E. F. Wanner, "An assignment problem formulation for dominance move indicator," in *2020 IEEE Congress on Evolutionary Computation (CEC)*, 2020, pp. 1–8.
- [12] M. Li, S. Yang, and X. Liu, "A performance comparison indicator for pareto front approximations in many-objective optimization," in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '15. New York, NY, USA:

- ACM, 2015, pp. 703–710. [Online]. Available: <http://doi.acm.org/10.1145/2739480.2754687>
- [13] M. Li, “Evolutionary many-objective optimisation: Pushing the boundaries.” Ph.D. dissertation, Department of Computer Science, Brunel University London, London, UB8 3PH, U.K., 2015.
- [14] M. Jünger, T. M. Lieblich, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, Eds., *50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art*. Springer, 2010. [Online]. Available: <https://doi.org/10.1007/978-3-540-68279-0>
- [15] Y. Yuan, Y. S. Ong, A. Gupta, and H. Xu, “Objective Reduction in Many-Objective Optimization: Evolutionary Multiobjective Approaches and Comprehensive Analysis,” *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 189–210, 2018.
- [16] K. Deb and D. Kalyanmoy, *Multi-Objective Optimization Using Evolutionary Algorithms*. USA: John Wiley Sons, Inc., 2001.
- [17] J. Deng and Q. Zhang, “Approximating hypervolume and hypervolume contributions using polar coordinate,” *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 913–918, Oct 2019.
- [18] K. Yang, M. Emmerich, A. Deutz, and T. Bäck, “Efficient computation of expected hypervolume improvement using box decomposition algorithms,” *Journal of Global Optimization*, vol. 75, no. 1, pp. 3–34, Sep 2019. [Online]. Available: <https://doi.org/10.1007/s10898-019-00798-7>
- [19] E. Bradford, A. Schweidtmann, and A. Lapkin, “Efficient multiobjective optimization employing gaussian processes, spectral sampling and a genetic algorithm,” *Journal of Global Optimization*, vol. 71, 02 2018.
- [20] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, “Modified distance calculation in generational distance and inverted generational distance,” in *Evolutionary Multi-Criterion Optimization*, A. Gaspar-Cunha, C. Henggeler Antunes, and C. C. Coello, Eds. Cham: Springer International Publishing, 2015, pp. 110–125.
- [21] K. Deb and M. Abouhawwash, “An optimality theory-based proximity measure for set-based multiobjective optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 515–528, Aug 2016.
- [22] C. Audet, S. L. Digabel, D. Cartier, J. Bibeon, and L. Salomon, “Performance indicators in multiobjective optimization,” 2018.
- [23] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, “How to specify a reference point in hypervolume calculation for fair performance comparison,” *Evol. Comput.*, vol. 26, no. 3, p. 411–440, Sep. 2018. [Online]. Available: https://doi.org/10.1162/evco_a_00226
- [24] E. Klotz and A. Newman, “Practical guidelines for solving difficult mixed integer linear programs,” *Surveys in Operations Research and Management Science*, vol. 18, p. 18–32, 10 2013.
- [25] L. Gurobi Optimization, “Gurobi optimizer reference manual,” 2019. [Online]. Available: <http://www.gurobi.com>
- [26] G. Gamrath, D. Anderson, K. Bestuzheva, W.-K. Chen, L. Eifler, M. Gasse, P. Gemander, A. Gleixner, L. Gottwald, K. Halbig, G. Hendel, C. Hojny, T. Koch, P. L. Bodic, S. J. Maher, F. Matter, M. Miltenberger, E. Mühmer, B. Müller, M. Pfetsch, F. Schlösser, F. Serrano, Y. Shinano, C. Tawfik, S. Vigerske, F. Wegscheider, D. Weninger, and J. Witzig, “The SCIP Optimization Suite 7.0,” Optimization Online, Technical Report, March 2020. [Online]. Available: http://www.optimization-online.org/DB_HTML/2020/03/7705.html
- [27] P. Sen and J.-B. Yang, “Multiple criteria decision support in engineering design,” 1998.
- [28] Y. Tian, R. Cheng, X. Zhang, F. Cheng, and Y. Jin, “An indicator-based multiobjective evolutionary algorithm with reference point adaptation for better versatility,” *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 4, pp. 609–622, Aug 2018.
- [29] R. Cheng, M. Li, Y. Tian, X. Xiang, X. Zhang, S. Yang, Y. Jin, and X. Yao, “Benchmark functions for the CEC’2018 competition on many-objective optimization,” Tech. Rep., 2018.
- [30] D. Brockhoff and T. Tušar, “Benchmarking algorithms from the platypus framework on the biobjective bbob-bioj testbed,” in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO ’19. New York, NY, USA: ACM, 2019, pp. 1905–1911. [Online]. Available: <http://doi.acm.org/10.1145/3319619.3326896>
- [31] D. Izzo, “Pygmo and pykep: Open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization),” 01 2012, pp. –.
- [32] J. Blank and K. Deb, “Pymoo: Multi-objective optimization in python,” *IEEE Access*, vol. 8, pp. 89 497–89 509, 2020.
- [33] R. While, L. Bradstreet, and L. Barone, “A fast way of calculating exact hypervolumes,” *IEEE Trans. Evolutionary Computation*, vol. 16, pp. 86–95, 02 2012.
- [34] S. Huband, P. Hingston, and L. Barone, “A Review of Multi-objective Test Problems and a Scalable Test Problem Toolkit A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit,” *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 2006, pp. 477–506, 2011. [Online]. Available: <http://ro.ecu.edu.au/ecuworks/2022>
- [35] H. Li, K. Deb, Q. Zhang, P. Suganthan, and L. Chen, “Comparison between MOEA/D and NSGA-III on a set of many and multi-objective benchmark problems with challenging difficulties,” *Swarm and Evolutionary Computation*, vol. 46, pp. 104–117, 02 2019.
- [36] E. Zitzler and S. Künzli, “Indicator-based selection in multiobjective search,” in *Proc. 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII)*. Springer, 2004, pp. 832–842.
- [37] E. Zitzler, M. Laumanns, and L. Thiele, “SPEA2: Improving the strength pareto evolutionary algorithm,” Tech. Rep., 2001.
- [38] A. P. Guerreiro, C. M. Fonseca, and L. Paquete, “The hypervolume indicator: Problems and algorithms,” 2020.
- [39] J. Blank and K. Deb, “A running performance metric and termination criterion for evaluating evolutionary multi- and many-objective optimization algorithms,” in *2020 IEEE Congress on Evolutionary Computation (CEC)*, 2020, pp. 1–8.



Cláudio Lúcio V. Lopes is a computer scientist and Ph.D. student at Computational and Mathematical Modelling program at CEFET-MG, Belo Horizonte, Brazil. He is a teaching assistant at PUC-MINAS, CTO, and founder at A3Data. His research interests include evolutionary computation, multi and many-objective optimization, quality indicators, machine learning and artificial intelligence.



Flávio Vinícius Cruzeiro Martins is graduated in Computer Science from the Federal University of Minas Gerais (2007), master's degree (2009), and a Ph.D. degree (2012) in Electrical Engineering from the same institution. He was an assistant professor at the Federal University of Ouro Preto from 2010 to 2014. He is currently an assistant professor with the Computer Engineering Department at the Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brazil. His current research interests include evolutionary algorithms,

mono and multi-objective optimization, combinatorial optimization, and optimization applied in real-world problems.



Elizabeth F. Wanner received the B.S. degree in mathematics in 1994 and the M.Sc. degree in pure mathematics in 2002 from the Universidade Federal de Minas Gerais (UFMG) and a Ph.D. degree in electrical engineering UFMG, in 2006. She is currently an Associate Professor with the Department of Computer Engineering, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, Brazil. Her current research interests include evolutionary computation, global optimization, constraint handling techniques for evolutionary algorithms, and multiobjective optimization.



Kalyanmoy Deb is Fellow, IEEE and the Koenig Endowed Chair Professor with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, USA. He received his Bachelor's degree in Mechanical Engineering from IIT Kharagpur in India, and his Master's and Ph.D. degrees from the University of Alabama, Tuscaloosa, USA, in 1989 and 1991, respectively. He is largely known for his seminal research in evolutionary multi-criterion optimization. He has published 549 international journal and conference research papers to date. His current research interests include evolutionary optimization and its application in design, modeling, AI, and machine learning. He is one of the top cited EC researchers with more than 149,000 Google Scholar citations and a h-index of 122.