Novel Genotypic Diversity Metrics in Real-coded Optimization on Multi-modal Problems

Anonymous Authors

***Abstract*—This paper presents innovative Genotypic Diversity Metrics (GDMs) designed for real-coded optimization in multi-modal problem contexts. The study introduces Minimum Individual Distance (D_{MID}) and Radius Diversity (D_{RaD}) metrics and compares them with traditional metrics such as Pairwise Distances Metric (D_{PW}) and Minimum Spanning Tree Metric (D_{MST}). The experiments evaluate the GDMs, highlighting their effectiveness and limitations in different scenarios. Through three experiments, the study reveals that while traditional metrics show limitations in complex settings, the newly proposed metrics demonstrate robustness and precision. The research spans simple to complex landscapes, including uni-modal and multi-modal problems, and explores varying numbers of optima and dimensions. The findings emphasize the strengths of D_{RaD} and D_{MID} in capturing population diversity accurately, suggesting their use for developing and evaluating optimization meta-heuristics.**

***Index Terms*—diversity metrics, evolutionary computation, search space**

I. Introduction

Evaluating the performance of Evolutionary Algorithms (EA) is crucial to understanding their effectiveness and guiding their improvement. Performance metrics provide quantitative measures to assess various aspects of an algorithm's behavior. These metrics are divided into two main categories, *Optimality-based* and *Diversity-based*. In this paper we are going to focus on the later one, more specifically, in a subcategory of them which are the *Genotypic Diversity Measures (GDM)*

The solution space alone doesn't give us the full picture of how an algorithm behaves. It can be helpful to also look at the search space, as suggested in [1].

In the context of the EAs, a significant issue is premature convergence, where the algorithm tends to settle for local optima prematurely. Controlling this phenomenon, known as the exploration/exploitation balance (EEB) [2], remains a challenge. Too much exploration can lead to inefficiency, while excessive exploitation results in premature convergence. Adjusting EA parameters, such as population size, evolution models, and restart strategies, offers a potential solution for managing the EEB [3]. It's worth noting that the EEB dilemma is not unique to EAs and can be viewed as a broader resource allocation challenge faced by adaptive systems.

Two paradigms are commonly used to conceptualize the EEB. In the Opposing Forces Paradigm, increasing exploration reduces exploitation and vice versa. In contrast, the Orthogonal Forces Paradigm suggests that both exploration and exploitation can be increased simultaneously, offering potential benefits, particularly in complex, multimodal landscapes [4].

Genotypic diversity, focusing on the diversity of solutions, is a valuable metric for monitoring exploration, while phenotypic diversity, related to the diversity of solution responses (fitness distribution), provides insights into exploitation. Properly assessing genotypic and phenotypic diversity is a challenging task. Nevertheless, the creation of such measures may contribute to the establishment of, for example, the groundwork for a feedback mechanism employed in adaptive methods. In fact, these mechanisms likely present the most intriguing application for diversity measures, as the developed measure could be employed, in part, to evaluate the quality of the EEB driving the optimization process.

This paper is divided into the review of the metrics used so far in the literature in the section II, following by the presentation of the proposed metrics in the section III, three different experiments testing the metrics in Section IV and least the conclusions in the section V.

II. Related Works and Preliminaries

A. Terminology

Before start to go deeply in the metrics, it is important to define some terms that we are going to use throughout the paper. About the optimization process, let $NG \in \mathbb{N}$[1] the total number of generations and $NE \in \mathbb{N}$ the total number of evaluations which our optimization process will go through. As it is possible (and usually do) to have more than one run during the experiments, let $NR \in \mathbb{N}$ the number of total runs and $ND \in \mathbb{N}$ the total number of dimensions (parameters) of the problem to be solved. Therefore, $g \in [1, NG]$, $n \in [1, NE]$, $r \in [1, NR]$ and $d \in [1, ND]$, all of them $\subset \mathbb{N}$, represents a specific generation, evaluation, run and dimension respectively.

About the algorithm, let $NP \in \mathbb{N}$ the total number of individuals in the population, and $NSP \in \mathbb{N}$ the total number of subpopulations, and therefore, $NSP \leq NP$, and each $sp \in [1, NSP] \subset \mathbb{N}$ represents a subpopulation. So, the term $I_{i,sp}(n)$ is the individual[2] i of the subpopulation sp at the evaluation n; The term $I_{best_{i,sp}}(n)$ is used to refer to the position with the highest fitness an individual $I_{i,sp}$ has had up to evaluation n; The term $S_{best_{sp}}(n)$ is the best individual of the subpopulation sp up to evaluation n; and the term $P_{best}(n)$ is used to refer to the best individual among the entire population NP up to evaluation n. Finally, x_i^d is the

[1]For all the $N*$ total numbers, some metrics or algorithm can reference to not the total but the numbers achieved so far, in these cases, the nomenclature is $N*'$.

[2]To make the notation cleaner, sometimes indexes like sp, e, r, g can be omitted when they are obvious.

position of the individual i in the dimension d and v_i^d is the velocity of the individual i in the dimension d.

TABLE I
SUMMARY OF THE TERMINOLOGY TERMS

Simbol	Description
NG	Total number of Generations
NR	Total number of Runs
NE	Total number of Evaluations
ND	Total number of Dimensions
NEV	Total number of Environments
NP	Total number of Individuals (Population)
NSP	Total number of Subpopulation
$I_{i,sp}$	Individual i of the subpopulation sp
$I_{best_{i,sp}}$	Highest fitness of the individual $I_{i,sp}$
$S_{best_{sp}}$	Best individual of the subpopulation sp
P_{best}	Best individual of the population NP
x_i^d	Position of the Individual i in the dimension d
v_i^d	Velocity of the Individual i in the dimension d

B. Normalization

To normalize GDMs for comparisons, two methods are used. One method involves using the landscape diagonal (LD) as a normalization factor. Another method, as suggested in [5], employs the maximum diversity value observed during the evolutionary process. This approach, known as "Normalized with Maximum Diversity so Far" (NMDF), begins with the initial EA population as a reference, updating as the process progresses. NMDF is akin to a technique in [6].

C. Previous Studies

Corriveau et al. (2012) [5] carried out a detailed study comparing more than 15 GDMs used in EC. Using a controlled evolutionary process and benchmark to control the diversity of the population, the metrics were ranked quantitatively based on some features like stability, sensitivity and effects of outliers. Among the metrics considered in the study, some were eliminated in the preparatory stage as it was observed flaws in their formulations, leaving 5 metrics that were tested using the benchmark presented in the section IV. Among these 5 metrics (D_{PW} [7], D_{MI} [8], D_{VAC} [6], D_{DTAP} [9] and D_{TD} [1]).

The study showed that no measurement was effective in capturing the full diversity of a population within any search process. The main drawback of these metrics happens when the problem trying to be solved is a multi-modal one, where there are more than only one optimum point. Follow are two examples showing why these metrics works good in a uni-modal problem and why they don't reflect good the diversity of the population in a multi-modal problem. However, the study concluded that of the metrics tested, D_{PW} was the one that best met the proposed requirements. Therefore, in this paper, among the 15 metrics, we will only use D_{PW}.

Lacevic and Amaldi (2010) [10] also carried out a comparative study of GDMs, this time considering metric based on alternative methods, not tested by the Corriveau et al. (2012) [5]. The metrics were classified into 6 approches: Volume-based diversity measure (L-diversity), Quasi-entropic diversity measure (E-diversity), Power mean based diversity measure (H-diversity), Diversity measures based on discrepancy, Diversity measure based on Euclidean minimum spanning trees and Diversity measures of binary coded population.

In order to assess the mutual dependencies among the 6 different diversity measures presented in the paper, the Pearson's correlation coefficient was used. The paper highlights that the while L-diversity stands out as an appealing metric from an ectropy perspective, it comes with a significant computational complexity. A comparative analysis was conducted using correlation coefficients across populations of varying types, sizes, and dimensions. Measures that emphasize large cumulative distances between points exhibit weak correlations with those that advocate for populations covering the domain uniformly. Among the metrics with moderate computational costs, the approach based on Euclidean minimum spanning trees (D_{MST}) emerges as a preferable alternative to L-diversity.

Corriveau et al. (2015) [11] in a more recent work, extended his comparison among the metrics in the literature, using now more than the controlled diversity benchmark. The study focuses on assessing the reliability of GDMs as descriptors of population diversity in optimization processes. Four GDMs — mean pairwise measure (D_{PW}), Shannon entropy (D_{GFS}), L-diversity or volume-based measure (D_L), and minimum spanning tree measure (D_{MST}) — are evaluated against three diversity requirements: monotonicity in individual varieties, twinning, and monotonicity in distance. The results indicate that D_{PW}, D_{GFS}, and D_{MST} exhibit improper responses to all three diversity requirements, mainly due to their inability to consider a uniformly distributed population as the most diverse state. D_L, while meeting two of the three requirements, has a prohibitive computational cost and lacks an efficient mechanism for monotonicity in distance. The study highlights the need for theoretical developments supporting practical requirements and emphasizes caution in generalizing univariate diversity indicators into multivariate contexts for GDM purposes. It concludes that the definition of an adequate genotypic diversity formulation for real-coded representation remains an open question, and the proposed GDM validation frameworks can facilitate the evaluation of new proposals by relating them to fundamental diversity requirements. The study suggests that the evaluation tool serves as a preliminary gate, and GDMs should undergo further testing in higher dimensions.

D. Review of the existing GDMs

Based on the comparison studies, we selected the two more suitable GDMs to present in more details and which will be the ones to be tested together with the metrics proposed in this paper.

1) Pairwise Distances Metric (D_{PW}): The concept of calculating the mean of pairwise distances among individuals in a population, denoted as D_{PW}, is an intuitive GDM. While it may be relatively time-consuming, D_{PW} has the potential to provide an effective description of population diversity. It's important to emphasize that prioritizing accuracy is preferable over speed when selecting a diversity measure. The equation 1 shows how to calculate the metric.

$$D_{PW} = \frac{2}{NP(NP-1)} \sum_{i=2}^{NP} \sum_{j=1}^{i-1} \sqrt{\sum_{d=1}^{ND} (x_i^d - x_j^d)^2} \tag{1}$$

where x_i^d is the position of the individual i in the dimension d, x_j^d is the position of the individual i in the dimension d, ND is the total number of dimensions and NP is the total number of individuals.

2) Minimum Spanning Tree Metric (D_{MST}): Proposed by Lacevic and Amaldi [10] is a metric based on the minimum spanning tree calculated over the solutions. Consider a complete undirected graph $G(V, C)$, where V is the set of points in the population NP, and C is the set of all $\frac{n(n-1)}{2}$ pairwise connections between the points. The weight of each edge represents the corresponding Euclidean distance between the points. The problem of finding a minimum spanning tree (MST) in this graph is known as the Euclidean minimum spanning tree (EMST) problem. It was introduced a diversity measure expressed in Equation 2.

$$D_{MST} = \mu(MST(G(V, C))) \tag{2}$$

where $MST(G(V, C))$ signifies the MST subgraph of $G(V, C)$, and μ represents the total length of the subgraph, i.e., the sum of the weights of all its edges. The motivation behind this measure is to extract the principal connections (distances) from the complete set of $\frac{n(n-1)}{2}$ connections. If the population consists of distinct clusters, the diversity measure avoids considering inter-cluster distances multiple times. This becomes crucial as multiple considerations of inter-cluster distances pose a challenge for measures reducing to the sum of all pairwise distances. A straightforward approach to compute D_{MST} is to construct a complete graph $G(V, C)$ and apply classical MST algorithms, such as Kruskal's or Prim's algorithm.

III. The Proposed Metrics

A. Minimum Individual Distance (D_{MID})

To address the issues related to the multi-modal problem of the previously mentioned diversity metric, the Minimum Individual Distance (MID) is proposed.

The main idea of this proposed metric is to consider, for each individual, only the distance between that individual and the closest other individual (as opposed to the D_{PW}, where all distances are considered). Figure 1 provides a simplified example of how it works by selecting only the minimum distance of each individual.

This approach ensures that great distances between individuals converged in different optima points are not taken into account. Additionally, distances between individuals with the same positions are not considered to prevent issues in scenarios (rare but possible) where two subpopulations are exactly

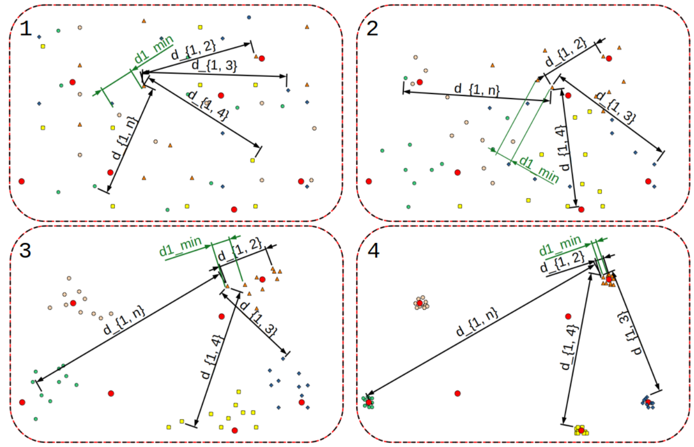

Fig. 1. Evolution of the population with 5 independent subpopulations in order to find the 5 optima points in the search space

the same, canceling each other out because the minimum calculated distance would be 0.

Another way to understand how this metric works is to imagine that as subpopulations tend to create niches around the optimum point they are pursuing, the metric will primarily consider only intra-niche distances and select the smallest among them. This way, large inter-niche distances are disregarded, and they are the main cause of the metrics shown so far not reflecting well the diversity of the population in multi-modal problems. Equation 3 shows how to calculate the metric.

$$D_{MID} = \sum_{i=1}^{NP} min\ d(i) \tag{3}$$

$$\text{with} \quad d(i) = \sqrt{\sum_{d=1}^{ND} (x_i^d - x_j^d)^2} \quad \forall j \neq i \in NP \; and \; d > 0 \tag{4}$$

where $min\ d(i)$ is the minimum euclidean distance found between the individual i and the others individuals in the population, x_i^d is the positions of the individual i in the dth dimension, x_j^d, NP is the total number of individuals in the population and ND is the total number of dimensions.

B. Radius Diversity (D_{RaD})

The Radius Diversity metric (RaD) selects a set of representative individuals from the population according to certain rules, and calculates the diversity based on the distances between these individuals and the rest of the population.

Here is an explanation of the individual selection method. Here, S represents the selected individuals, and P is the population. First, the two individuals in P with the maximum Euclidean distance within the population are added to S. Next, a large sphere is drawn with the individuals in S as the center, and its radius is gradually reduced. Individuals that fall outside the sphere are then added from P to S, and a new sphere is drawn with the updated solution in S as the center. These steps are repeated until P is empty. The length of the radius at the time when an individual is chosen is used as its diversity measure.

A simple, alternative explanation of how RaD chooses individuals is that: Every new individual is "the individual with the largest "smallest distance" to all individuals already selected". Figure 2 shows an example of selecting from a population distributed in a 2D landscape. In this figure, the diamonds represent individuals that have already been selected, and the triangles represent the next individual to be chosen. According to the figure, indeed, the triangle individual is the one with the maximum minimum distance to the diamond individuals.

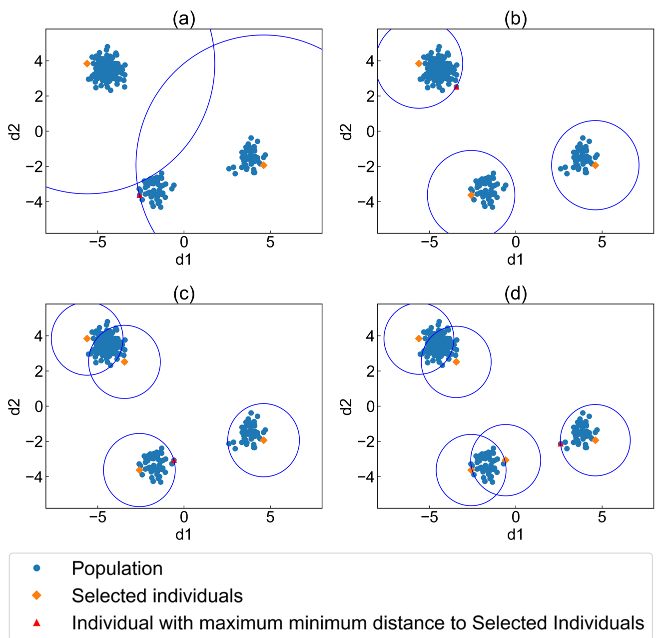

Fig. 2. Transition of individual selection by RaD. Select individuals in the order (a) → (b) → (c) → (d)

This selection method implies that once an individual is selected, individuals distributed around it are not selected for some time. Thus, in a converging population with multimodal optimization, the representative individuals of a niche in the population are selected first, followed by the individuals surrounding that niche. In other words, the representative solution of a niche is highly valued and the individuals surrounding that representative solution are valued low. The RaD individual selection algorithm is shown in Algorithm 1, and the metric calculation formula is shown in Equation 5.

$$D_{RaD} = \sum_{i=1}^{NP} d_{\text{radius}}(i) \tag{5}$$

where $d_{\text{radius}}(i)$ is the diversity value for individual i, signifying the length of the radius at the time when individual i is chosen, and NP is the total number of individuals in the population.

Algorithm 1 Individual selection

```
1: P ← Population of individuals
2: S ← ∅
3: Select individuals x_i and x_j ∈ P with the maximum
   distance
4: d_radius(i) ← 0
5: d_radius(j) ← maximum distance
6: S ← S ∪ {x_i, x_j}
7: Remove x_i, x_j from P
8: while P ≠ ∅ do
9:    Select individual x_k ∈ P with the maximum minimum
      distance to S
10:   d_radius(k) ← maximum minimum distance to S
11:   S ← S ∪ {x_k}
12:   Remove x_k from P
13: end while
```

IV. Experiments and Results

The Table II summarize the equation to calculate the GDMs presented so far. Using these metrics we are now going to run some experiments in order to compare their behaviors in different scenarios.

TABLE II
GENOTYPIC DIVERSITY METRICS (GDMS) TO BE TESTED AND ITS MATHEMATICAL EQUATIONS

GDM	Equation
D_{PW}	$\frac{2}{NP(NP-1)} \sum_{i=2}^{NP} \sum_{j=1}^{i-1} \sqrt{\sum_{d=1}^{ND}(x_i^d - x_j^d)^2}$
D_{MST}	$\mu(MST(G(X,E)))$
D_{RaD}	$\sum_{i=1}^{NP} d_{\text{radius}}(i)$
D_{MID}	$\sum_{i=1}^{NP} min\ d(i)$

A. Dataset with frozen population diversity

We conducted a mathematical assessment of the metrics' performance on a simple dataset with a fixed population diversity. This evaluation was based on the seven cases proposed by Corriveau et al [11]. In these cases, with the exception of cases 6 and 7, all solutions converged perfectly to predefined optimal points. This convergence allows us to observe diversity without being influenced by variations in values due to data variability.

All these cases are situated on a 2D landscape within the range of $[-1, 1]$, with a constant population size of 100. In Case 1, a single optimal point exists at the center of the landscape, and the population is concentrated at this central point. In Case 2, the optimal point lies between the center and the corner of the landscape, and the population is uniformly distributed. Case 3 shares the same optimal point as Case 2 but exhibits non-uniform population distribution. Case 4 features the optimal point at the corner of the landscape, with an even distribution of the population. Similarly, Case 5 has the same optimal point as Case 4 but also displays non-uniform population distribution. In both Case 3 and Case 5, 70% of the individuals are concentrated at one optimal point, while the remaining 30% are evenly distributed among the other optimal points. In Case 6, individuals are evenly distributed along the diagonal of the landscape. Lastly, in Case 7, individuals are uniformly distributed at regular intervals across the landscape. Figure 3 illustrates the population distributions.

In these frozen cases, the diversity ranking must be case1 $<$ case2 $=$ case3 $<$ case4 $=$ case5 $<$ case6 $<$ case7 in order to satisfy the three diversity requirements proposed by Corriveau.

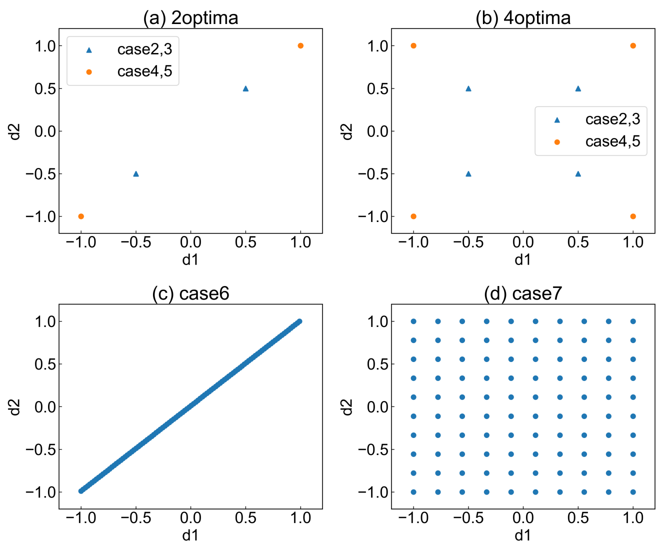

Fig. 3. Position of the population in cases 2 to 7. (a) cases 2 to 5 in 2-optima landscape , (b) cases 2 to 5 in 4-optima landscape, (c) case 6, (d) case 7

1) Results: When applying the metrics listed in Table II to the seven previously described cases, the results are presented in Table III. Upon initial examination of the results, it becomes that all metrics performed as anticipated in Case 1, where we observed a population entirely converged to a single point within the search space.

For Cases 2 to 5, the only metric that did not meet the specified requirements was D_{PW}, which failed to satisfy the condition that Case 2 should equal Case 3 and Case 4 should equal Case 5.

In Case 6, the only metric that satisfied this requirement was D_{RaD}, consistent with the premise that Case 6 exhibits greater diversity compared to Cases 1 through 5.

Lastly, in Case 7, where individuals are randomly distributed across the landscape, implying the highest expected diversity, only D_{PW} once again failed to accurately reflect this diversity.

In summary, it becomes apparent that only D_{RaD} managed to meet all the specified requirements.

B. Controlled Diversity Benchmark

To gain a deeper understanding, let's delve into how these metrics behave in both uni-modal and multi-modal problems.

TABLE III
BEHAVIOR OF THE GDMS OVER THE SEVEN FROZEN CASES

GDM	Landscape	Genotypic distribution cases 1	2	3	4	5	6	7
D_{PW}	2opt	0.00	0.71	0.60	1.43	1.20	0.96	1.16
	4opt		0.86	0.55	1.72	1.10		
D_{MST}	2opt	0.00	1.41	1.41	2.83	2.83	2.83	22.00
	4opt		3.00	3.00	6.00	6.00		
D_{RaD}	2opt	0.00	1.41	1.41	2.83	2.83	11.74	37.38
	4opt		3.41	3.41	6.83	6.83		
D_{MID}	2opt	0.00	2.83	2.83	5.66	5.66	2.86	22.22
	4opt		4.00	4.00	8.00	8.00		

To facilitate this exploration, an appropriate benchmark is essential. Common benchmarks in Evolutionary Algorithms (EAs) often lack a controlled variable for population diversity. Consequently, information about the actual diversity status of a population is scarce, except through the comparison of Genotypic Diversity Metrics (GDMs). This creates an ill-defined problem, as different GDMs may yield varied estimations, complicating the determination of the most accurate representation of authentic diversity.

The benchmark's definition is also crucial. Genotypic diversity focuses solely on the distribution of individuals across a landscape, irrespective of their respective fitness functions. Thus, creating an environment for GDMs necessitates transitioning the population from a fully dispersed state to a completely aggregated one. Additionally, the number of optimal solutions across the landscape is expected to impact GDMs. A well-defined benchmark should simulate the influence of modalities. In this context, we adopt the benchmark proposed by Corriveau et al [5].

The benchmark's concept involves creating an environment where the diversity of individuals in the population can be known and controlled, ensuring a clear evaluation of diversity metrics. In each generation, the landscape linearly decreases, and a new population is randomly created.

To assess the performance of GDMs on a uni-modal problem, Figure 4 illustrates the benchmark's evolution over generations. The landscape progressively reduces by 1% of the initial size in each dimension in every generation, and the population becomes less spread until converging in generation 100.

To test the performance of the GMDs on a multi-modal problem, Figure 5 shows how the benchmark evolve over the generations, making the landscape 1% of the initial size smaller in each dimension in each generation and randomly recreating the population less spread until the convergence on the generation 100. But in this case, instead of only one optimum point, we have 5 and the population was divided in 5, where each of these subpopulations is in charge to converge to a different optimum point.

Finally in order to evaluate the metric's consistency under different numbers of optima points and dimensions we are going now to consider only the value of the last generation.

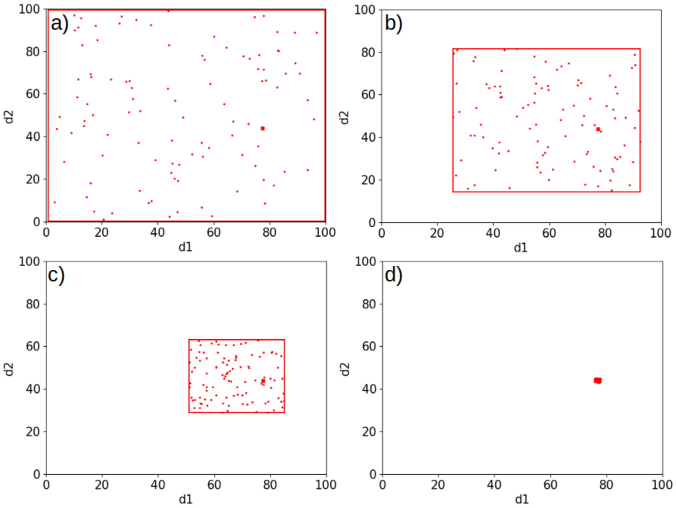

Fig. 4. Diversity controlled benchmark over the generations until the population converge to the optimum point. a) Generation: 1; b) Generation: 33; c) Generation: 66; d) Generation: 99

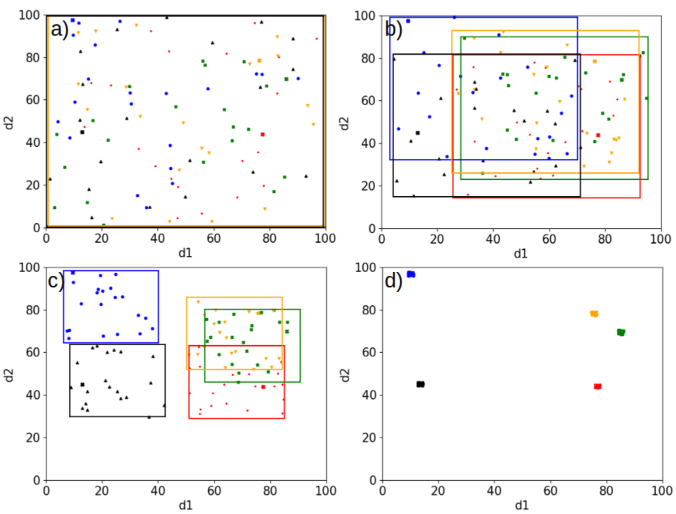

Fig. 5. Diversity controlled benchmark over the generations until the population converge to 5 optima points. a) Generation: 1; b) Generation: 33; c) Generation: 66; d) Generation: 99

As a controlled system, we know that this is the generation where the populations have converged.

All the experiments were ran 50 times, the population size set in 100 individuals and along 100 generations.

1) Results: Regarding the uni-modal problem, the results are illustrated in Figure 6-a, and it is worth noting that all the metrics behaved as expected. As the size of the landscape decreases linearly over the course of 100 generations, the expected pattern is a straight line representing the transition from the most diverse case to a fully converged one.

More intriguing results are observed in Figure 6-b, where the metrics were applied to a multi-modal problem featuring 5 optimal points. It is noteworthy that the only metric deviating from the expected behavior was D_{PW}. After approximately the 50th generation, its values begin to rise again, even as the population converges towards the optimal points.

In contrast, the behavior of the other metrics closely resembles the uni-modal scenario, characterized by a smooth curve in the middle of the process. This smooth curve in the middle is also anticipated, as, at a certain point in the process, the subpopulations will intersect, leading to a brief ”stable moment” lasting a few generations before returning to the linear decrease. Ultimately, this results in different values, but significantly smaller ones compared to the initial diversity value.

Now, in terms of metric stability while varying the number of optima and dimensions, we examine Figures 6-c and 6-d. When maintaining a constant number of optima (5) and varying the number of dimensions from 2 to 50, it becomes evident by Figure 6-c that all metrics exhibit a decline in values up to 10 dimensions. However, beyond this point, the values remain nearly constant, indicating robust stability across a range of dimensions for all metrics.

Conversely, when keeping the number of dimensions fixed (10) and varying the number of optima, Figure 6-d illustrates that only D_{MID} displayed robust stability. It's worth noting that the stability of D_{PW} was not observed, as its values consistently approached the diversity limit.

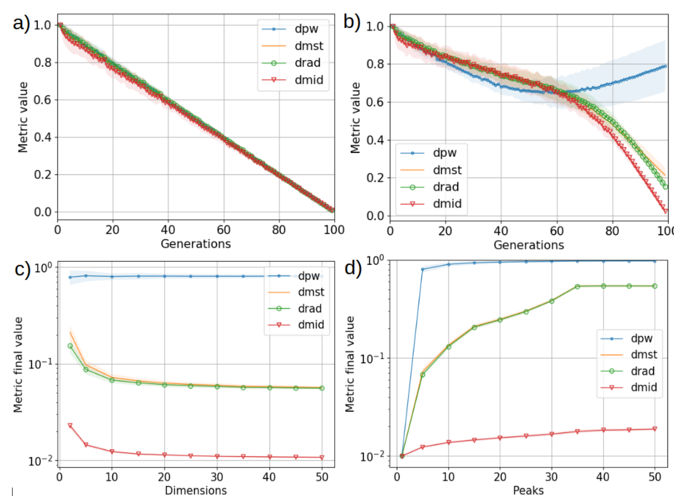

Fig. 6. The four GDMs presented in the Table II with NP: 100, NR: 50, NG: 10 applied to: a) An uni-modal problem; b) A multi-modal problem (5 optima); c) Fixed number of optima (5) varying the number of dimensions; d) Fixed number of dimensions (10) varying the number of optima

C. Moving Peaks Benchmark

In these final experiments, we put the metrics to the test in real optimization processes, providing insights into their practical performance alongside actual algorithms. This approach not only helps us explore the metrics theoretically but also reveals their behavior in practical optimization scenarios. To conduct these experiments, we have selected the Moving Peaks Benchmark (MPB) [12] as our testing ground. We chose the MPB due to its ability to finely control various problem characteristics, including the number of optima, their positions, dimensions, height, and more.

The MPB offers three predefined scenarios, and throughout our experiments, we will thoroughly investigate each one. This entails subjecting the metrics to testing under these three distinct scenarios. Table IV provides an overview of the key differences between these scenarios, with the remaining configurations adjusted to match each specific scenario. To further illustrate, Figure 7 showcases an example configuration employing scenario 1, featuring 5 optima points in a 2-dimensional problem.

TABLE IV
MPB SETTING VALUES FOR THE EXPERIMENTS. THE REMAINING CONFIGURATION IS THE ONES OF ITS SCENARIO [13]

Parameters	Scenario 1	Scenario 2	Scenario 3
Peak function	$function1()$ [14]	$cone()$ [14]	$cone()$
Number of peaks	5	10	50
Number of dimensions	2	2	2
Peak heights	$\in [30, 70]$	$\in [30, 70]$	$\in [30, 70]$
Peak widths	$\in [0.0001, 0.2]$	$\in [1, 12]$	$\in [1, 12]$
Min coordinate	0	0	0
Max coordinate	100	100	100

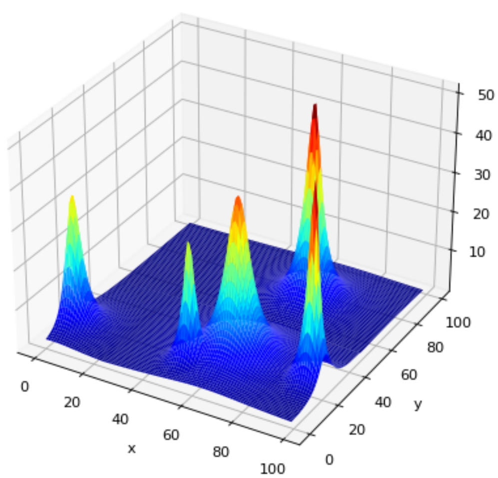

Fig. 7. Example of the MPB with the configuration set up as defined by the scenario 1

1) Results: All experiments were conducted 50 times, each with a population of 100 individuals and in 50000 number of evaluations (NR = 50, NP = 100 and NE = 50000).

Our results suggest that diversity metrics provide not only theoretical insights but also reveal distinct behavioral patterns when applied to practical optimization problems. In each scenario, the diversity of algorithms such as PSO [15], DE [16], $EA + SBX$ [17] and $ANDE$ [18] varied, with their convergence rates becoming particularly noticeable under the dynamic conditions presented by the MPB. The Figures 8, 9, and 10 show the results for each scenario for each metric.

Based on prior knowledge of how these algorithms operate, it was expected that in all scenarios and across all metrics, the algorithms $EA + SBX$ and $ANDE$ would exhibit the greatest diversity. This can be attributed to these algorithms' use of a multi-population strategy, which inherently enhances diversity across the landscape. The rapid convergence of PSO and DE is also apparent in all the scenarios and metrics.

Despite these practical results showing the rate at which the population of each algorithm converges, we can draw attention to D_{MID} for its ability to more distinctly differentiate between each algorithm. Excluding the completely converged situations of PSO and DE, in all three scenarios, $EA + SBX$ and $ANDE$ can be distinguished from one another, even when considering their variance over the 50 runs.

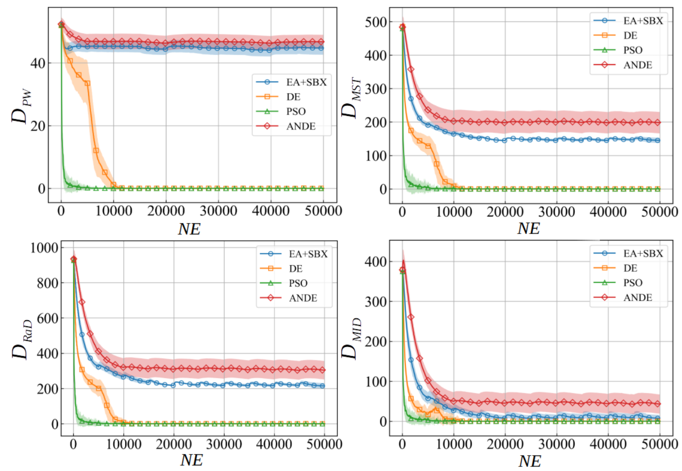

Fig. 8. The four GDMs running in the MPB set in scenario 1

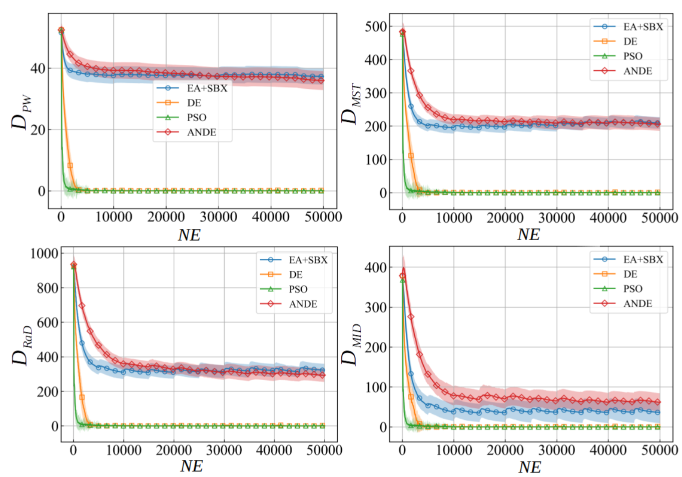

Fig. 9. The four GDMs running in the MPB set in scenario 2

V. CONCLUSION

Based on the detailed results from the three experiments, the conclusion can be summarized as follows:

The experiments comprehensively evaluated various diversity metrics across different scenarios, revealing significant insights. In simple scenarios (Case 1), all metrics performed as expected, indicating their basic effectiveness. However, in

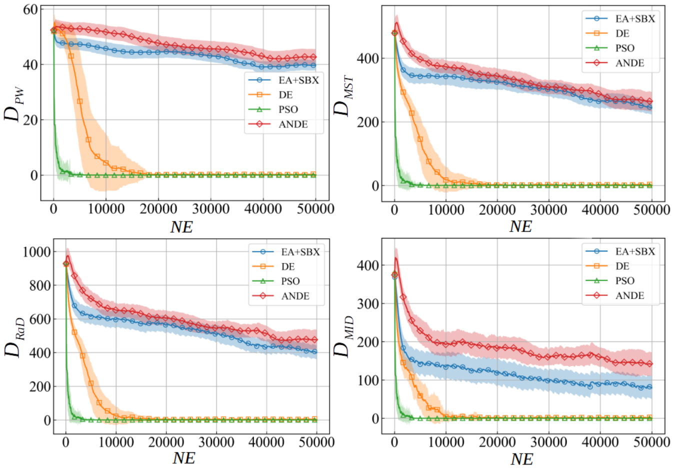

Fig. 10. The four GDMs running in the MPB set in scenario 3

more complex situations (Cases 2-5), the Pairwise Distances Metric (D_{PW}) failed to meet specific conditions, suggesting its limitations. The Radius Diversity (D_{RaD}) metric consistently met all requirements across these cases, indicating its robustness and accuracy.

In the uni-modal problem, all metrics behaved as anticipated, depicting a linear decrease in diversity. However, in the multi-modal problem, D_{PW} deviated unexpectedly, suggesting its inadequacy in certain conditions. The other metrics showed a temporary stability phase before continuing their decline, reflecting the dynamics of population convergence.

The final experiment, varying the number of optima and dimensions, demonstrated that while all metrics were stable up to a certain dimensionality, only the Minimum Individual Distance (D_{MID}) metric showed consistent stability across different numbers of optima.

Overall, these experiments highlight the effectiveness and limitations of different diversity metrics in evolutionary algorithms, with particular strengths observed in the Radius Diversity and Minimum Individual Distance metrics under varying conditions and problem complexities.

References

[1] M. Wineberg and F. Oppacher, "The underlying similarity of diversity measures used in evolutionary computation," in *Genetic and Evolutionary Computation — GECCO 2003*, E. Cantú-Paz and J. Miller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 1493–1504.

[2] A. E. Eiben and C. A. Schippers, "On evolutionary exploration and exploitation," *Fundamenta Informaticae*, vol. 35, no. 1-4, pp. 35–50, 1998.

[3] Á. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms," *IEEE Transactions on evolutionary computation*, vol. 3, no. 2, pp. 124–141, 1999.

[4] A. K. Gupta, K. G. Smith, and C. E. Shalley, "The interplay between exploration and exploitation," *Academy of management journal*, vol. 49, no. 4, pp. 693–706, 2006.

[5] G. Corriveau, R. Guilbault, A. Tahan, and R. Sabourin, "Review and study of genotypic diversity measures for real-coded representations," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 5, pp. 695–710, 2012.

[6] F. Herrera, E. Herrera-Viedma, M. Lozano, and J. L. Verdegay, "Fuzzy tools to improve genetic algorithms," in *Second European Congress on Intelligent Techniques and Soft Computing*, vol. 3, 1994, pp. 1532–1539.

[7] A. Barker and W. Martin, "Dynamics of a distance-based population diversity measure," in *Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512)*, vol. 2, 2000, pp. 1002–1009 vol.2.

[8] D. J. Morrison, R.W. and K.A., "Measurement of population diversity," in *Artificial Evolution. EA 2001. Lecture Notes in Computer Science, vol 2310. Springer, Berlin, Heidelberg*, vol. 2310, 2002, pp. 1128–1134.

[9] R. K. Ursem, "Diversity-guided evolutionary algorithms," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2002, pp. 462–471.

[10] B. Lacevic and E. Amaldi, "Ectropy of diversity measures for populations in euclidean space," *Information Sciences*, vol. 181, no. 11, pp. 2316–2339, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025510006043

[11] G. Corriveau, R. Guilbault, A. Tahan, and R. Sabourin, "Evaluation of genotypic diversity measurements exploited in real-coded representation," *ArXiv*, vol. abs/1507.00088, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:2995563

[12] J. Branke, "Memory enhanced evolutionary algorithms for changing optimization problems," in *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, vol. 3, 1999, pp. 1875–1882 Vol. 3.

[13] F.-M. De Rainville, F.-A. Fortin, M.-A. Gardner, M. Parizeau, and C. Gagné, "Deap: A python framework for evolutionary algorithms," in *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation*, ser. GECCO '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 85–92. [Online]. Available: https://doi.org/10.1145/2330784.2330799

[14] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: Evolutionary algorithms made easy," *Journal of Machine Learning Research*, vol. 13, pp. 2171–2175, jul 2012.

[15] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, 1995, pp. 1942–1948 vol.4.

[16] R. Storn and K. V. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359, 1997. [Online]. Available: https://api.semanticscholar.org/CorpusID:5297867

[17] K. K. Bali, Y.-S. Ong, A. Gupta, and P. S. Tan, "Multifactorial evolutionary algorithm with online transfer parameter estimation: Mfea-ii," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 1, pp. 69–83, 2020.

[18] Z.-J. Wang, Z.-H. Zhan, Y. Lin, W.-J. Yu, H. Wang, S. Kwong, and J. Zhang, "Automatic niching differential evolution with contour prediction approach for multimodal optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 1, pp. 114–128, 2020.