Impact of Different Discrete Sampling Strategies on Fitness Landscape Analysis Based on Histograms

Vojtěch Uher
vojtech.uher@vsb.cz
Department of Computer Science
VŠB - Technical University of Ostrava
Ostrava, CZE

Pavel Krömer
pavel.kromer@vsb.cz
Department of Computer Science
VŠB - Technical University of Ostrava
Ostrava, CZE

ABSTRACT

Complex problems are frequently tackled using techniques from the realm of computational intelligence and metaheuristic algorithms. Selection of a metaheuristic from the wide range of algorithms possessing various properties to address specific problem types efficiently is a difficult and crucial task to avoid unnecessary blind alleys and computational expenses. Approximation of continuous problem landscapes by a limited number of scattered discrete samples is a widespread problem characterization applied in exploratory landscape analysis (ELA). ELA is a set of methods analyzing the objective and solution spaces of a problem to construct features estimated from the random samples. This paper describes a simple method for fitness landscape analysis based on the normalized histograms of sample fitnesses. Generation of a small number of representative discrete samples is crucial for efficient problem characterization, and therefore, amount of sampling strategies including random generators and low-discrepancy sequences was developed to evenly cover the problem landscapes. The main contribution of this paper is a study examining the impact of different sampling strategies on the distribution of fitness values based on the normalized histogram analysis. The results reveal a strong effect.

CCS CONCEPTS

• **Computing methodologies → Optimization algorithms**; **Bio-inspired approaches**; • **Mathematics of computing → Distribution functions**; • **Theory of computation → Randomness, geometry and discrete structures**.

KEYWORDS

sampling strategies, low-discrepancy sequences, metaheuristics, exploratory landscape analysis, fitness distribution

ACM Reference Format:
Vojtěch Uher and Pavel Krömer. 2023. Impact of Different Discrete Sampling Strategies on Fitness Landscape Analysis Based on Histograms. In *13th International Conference on Advances in Information Technology (IAIT 2023), December 06–09, 2023, Bangkok, Thailand*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3628454.3631563

IAIT 2023, December 06–09, 2023, Bangkok, Thailand

ACM ISBN 979-8-4007-0849-7/23/12.
https://doi.org/10.1145/3628454.3631563

1 INTRODUCTION

The metaheuristic algorithms such as Differential Evolution [5, 34], Particle Swarm Optimization [6], or Genetic Algorithm [10, 22] are methods efficiently dealing with complex problems solution and optimization. However, they are also known to perform differently on various types of problems [19]. A selection of an efficient algorithm for a specific type of problem can significantly improve the optimization performance and save a lot of expensive fitness function evaluations [15, 16, 35]. This issue is even more intensified by multi-objective problems, high dimensions, and heavily constrained problems [12, 14, 17].

A fitness landscape (FL) [27] is defined as a continuous problem function that represents a search space where the 'elevation' over the specific location in the search space is the fitness (objective) value. In other words, a FL is a mountainous region with peaks, plateaus, valleys, and ridges describing the topological features that are crucial for understanding the dynamics of evolutionary algorithms. The fitness landscape analysis investigates the features of FLs (e.g. ruggedness, deceptiveness, multi-modality) to describe and distinguish different types of FLs (problems). This investigation helps to solve the algorithm selection problem [16]. A source of FLs can be some set of artificial benchmark problems [8] and also many real-world problems including quadratic assignment problem [20], dynamic optimization problem [26], knapsack problem [33], traveling salesman problem [23], vehicle routing problem [21] etc.

The exploratory landscape analysis (ELA) [19] is a popular problem-independent method for the characterization of continuous FLs of optimized problems. It encompasses a series of procedures that describe the hypersurfaces created by the fitness and other distinctive attributes of the problem solutions, all based on a finite set of samples. An example of a widely-used collection of features for ELA is the FLACCO library [11] by Kerschke and Trautmann.

The selection of the set of samples and the analysis of the corresponding fitness values are particularly important for fitness landscape characterization as these values represent the input base for features estimation [19, 35]. The more representative is the sampling, the more representative are the computed features. Various strategies can be employed for sampling. In addition to pseudo-random sampling, one can utilize quasi-random techniques like Latin Hypercube Sampling (LHS) [11, 18], or low-discrepancy sequences such as Sobol [30] and Halton [7]. The properties of these sampling strategies vary, but they generally aim at even coverage of the search space, high regularity, and low discrepancy to better capture different areas of FLs.

Empirical investigations conducted on single-objective problems have revealed that the landscape feature values are significantly influenced by the choice of the sampling strategy, as reported by Renau et al. [25]. Furthermore, the variance in feature values implies that features obtained from various sampling strategies serve diverse roles in problem characterization. Recent studies [1, 12] have extended these findings to the realm of bi-objective problems, illustrating that distinct sampling strategies result in varying degrees of accuracy in problem classification. This encourages additional investigation of the effect of various sampling strategies in the context of problem types and the features relevant to landscape analysis.

This paper proposes a fitness landscape analysis method based on the distribution of fitness values computed for a selected set of discrete samples. As the landscapes of different problems have various shapes, smoothness, multi-modality, and local and global optima, the distribution of fitness values is heavily affected by these properties. In statistics, the set of fitness values is considered a random variable and its discrete probability distribution can be represented by a normalized histogram consisting of a user-defined number of bins [29]. The features based on the histograms are often used in image retrieval [2, 28, 32] where similar frequencies of different intensities or colors indicate similar images. We apply this analogy to distributions of continuous variables in the manner of statistical histograms of numerical data. The normalized histograms are used here as feature vectors distinguishing different test functions. The main goal of this paper is to investigate the impact of different sampling strategies on the shape and representativeness of histograms because the samples and their fitness values are crucial for ELA features computation.

The study is performed on the set of 24 BBOB single-objective test problems available on the COmparing Continuous Optimizers (COCO) platform [8]. The experiments are computed for 4 sampling strategies and 2 distance measures comparing fitness distributions represented by histograms. To assess the impact of sampling strategies, a cluster analysis based on the normalized histograms is accomplished and the histograms for different functions and samplings are compared. The results show clear differences between the tested sampling strategies.

Section 2 describes the pipeline and the methods utilized to describe the test problems with different sampling strategies and corresponding histograms. Section 3 provides the performed experiments comparing sampling strategies and discusses the achieved results.

2 UTILIZED METHODOLOGY

This paper investigates the impact of different sampling strategies on the fitness values distribution obtained by the evaluation of the generated discrete samples. Each test function is represented by a normalized histogram of fitness values. The assumption is that the functions with similar properties should have comparable fitness histograms. The methodology proposed in this section aims to reveal the relation between used sampling strategy and the ability to distinguish different test functions using their fitness histograms.

Section 2.1 describes the used test framework and its basic parameters. Section 2.2 briefly summarizes the used sampling strategies and their properties. Section 2.3 defines the proposed procedure of histogram computation and fitness landscape analysis.

2.1 Test framework

The study is performed on the set of single-objective benchmark problems available on the COmparing Continuous Optimizers (COCO) platform [8]. The set consists of 24 BBOB test problems with different fitness landscapes and properties that represent a simple framework for sampling strategies and corresponding histograms comparison. The COCO is a state-of-the-art publicly available platform (also implemented in Python) often used for ELA, feature extraction and selection, and problem classification upon single- and multi-objective problems. Each problem is defined for various dimensions d that are also considered in experiments. To analyze the continuous functions of test problems, a set of n discrete samples is generated using some sampling strategy which is defined as a power of two $n = 2^m$ as it is convenient for some low-discrepancy sequences. The implementation is based on the Python COCO library, the SciPy and Scikit-learn Python libraries that are all publicly available. The libraries contain all the selected sampling strategies, distance metrics, and analytic tools used in this paper.

2.2 Sampling strategies

In the ELA, the selection of discrete random samples scattered over the continuous FL is crucial as it defines how the objective function and its features are covered [24]. A heavily biased set of samples leads to the systematic loss of information due to under- or oversampling of some regions. Therefore, different sampling strategies have been developed to allow efficient and accurate characterization of the optimized problem focused on even coverage of the search space, high regularity, and low discrepancy to capture different areas of FLs.

The following sampling strategies were studied. **Uniform** random sampling is the baseline method of generating samples using a pseudorandom generator with a uniform probability distribution. **Latin Hypercube Sampling (LHS)** generates near-random samples by dividing values of variables in multi-dimensional spaces into equal intervals and makes sure that only one sample is drawn from each column and row of such a grid [18]. A refined version of LHS, denoted as **LHSO**, employs randomized coordinate permutations to enhance space-filling reliability and reduce centered discrepancy. **Sobol's sequence-based** sampling **(Sobol)** uses Sobol's low-discrepancy sequence which is a quasi-random sequence with base 2 that binarily represents the position of each dimension [30] and can be efficiently implemented by bit-vector operations. In order to improve the sequence's discrepancy, a linear matrix scramble with digital random shifting is utilized. **Halton's sequence-based** sampling **(G-Halton)** is built upon the generalized Halton's low-discrepancy quasi-random sequence and uses the coprime integers as its bases [7]. This technique serves as an extension of the one-dimensional van der Corput sequence [4]. A permutation is applied to the digits when computing the radical inverse that reduces the problem with regular patterns in higher dimensions.

2.3 Histogram-based fitness landscape analysis

To analyze and classify the fitness landscapes of optimization problems, their features have to be computed and compared. This section describes a simple feature based on the statistic of fitness values obtained by evaluating discrete samples scattered over the continuous landscape.

2.3.1 Normalized histogram. Given a fitness function $f : \mathbb{R}^d \to \mathbb{R}$ and a set of samples $S = \{s_1, \ldots, s_n\}$, the set of fitness values is computed as $V = \{v; \forall s \in S : v = f(s)\}$, where d is the problem dimension and n is the number of samples. The set V is used to build a histogram of k bins within the range of values $\langle \min(V), \max(V) \rangle$ such that

$$n = \sum_{j=1}^{k} c_j, \tag{1}$$

where c_j is count of values within the j-th bin. The normalized histogram is defined as

$$1 = \sum_{j=1}^{k} \frac{c_j}{n}. \tag{2}$$

The normalized histogram represents a discrete probability distribution of fitness values $\mathbf{c} = \{c_1/n, \ldots, c_n/n\}$. It is aggregated from local fitness values and forms a simple global feature vector that can be used for ELA. The histogram is affected by landscape properties such as ruggedness, multi-modality, or variance of fitness values. Moreover, the number of histogram bins k gives a fixed length to feature vectors that can be easily compared with each other. The k also controls the level of precision of the captured distribution. Smaller k leads to greater generalization of the contained information.

2.3.2 Distance between histograms. To compare normalized histograms of two test functions, some distance or similarity measures have to be chosen. Generally in a metric space, some distance measuring a spatial distance between points or a degree of difference is defined (e.g. Euclidean, Manhattan, Hamming). However, the meaning of bins of normalized histograms cannot be directly interchanged with point coordinates because histograms cannot generally take on arbitrary values as the sum (2) equals to 1. As we compare distributions of values, two statistical distance measures based on histograms are tested.

The *histogram intersection* [2, 32] is simply defined by a sum of minimum values of corresponding bins of two histograms $\mathbf{a}$ and $\mathbf{b}$:

$$histInt(\mathbf{a}, \mathbf{b}) = \sum_{j=1}^{k} \min(a_j, b_j). \tag{3}$$

The intersection measures the degree of similarity between two histograms. The equality $\mathbf{a} = \mathbf{b}$ means $histInt(\mathbf{a}, \mathbf{b}) = 1$. The *histogram distance* is defined as $histDist(\mathbf{a}, \mathbf{b}) = 1 - histInt(\mathbf{a}, \mathbf{b})$.

The second approach is the *Kullback–Leibler(KL)-divergence* [13] which computes the relative entropy from one probability distribution to another. For two normalized histograms $\mathbf{a}$ and $\mathbf{b}$ it is defined as:

$$KL(\mathbf{a} \parallel \mathbf{b}) = \sum_{j=1}^{k} a_j \log\left(\frac{a_j}{b_j}\right). \tag{4}$$

The statistical divergence measures how one probability distribution differs from another. The divergences can be understood as generalizations of the squared Euclidean distance (SED), they generally do not satisfy symmetry and triangle inequality, and therefore, they are not real distance metrics [3]. However, the SED is often used as a cheaper variant of the Euclidean distance (ED) avoiding the square root as these two distances are proportional [31]. Although the SED is symmetric and does not satisfy the triangle inequality, it can be used instead of ED in applications when comparing distances. The KL-divergence is not symmetric, and thus, it has to be symmetrized to be used similarly:

$$KLDiv(\mathbf{a}, \mathbf{b}) = \frac{KL(\mathbf{a} \parallel \mathbf{b}) + KL(\mathbf{b} \parallel \mathbf{a})}{2}. \tag{5}$$

When mentioning the KL-divergence later in the text its symmetrized form is considered.

3 EXPERIMENTS AND RESULTS

The main aim of the experiments is to investigate the ability of the normalized histograms of fitness values to represent the test functions, the impact of the sampling strategies on histogram shapes, and the comparison of distance measures.

The experimental setup consists of 24 BBOB COCO functions that are tested for three dimensions $d \in \{5, 10, 20\}$. Four sampling strategies (Uniform, Sobol, G-Halton, LHSO) are used to generate random samples. To contribute a robust statistical analysis, 30 sample sets are randomly generated for each strategy and dimension. Three sample set sizes $n \in \{2^{10}, 2^{12}, 2^{14}\}$ are considered. All sample sets are evaluated by 24 test functions of the corresponding dimension to produce sets of fitness values. One normalized histogram of fitness values consisting of $k = 50$ bins is constructed for each function and combination of parameters. Both, histogram distance and KL-divergence are used to compute relevant statistics and visualizations.

First, the cluster analysis is conducted in Section 3.1 to examine how the histograms distinguish the test functions. Section 3.2 compares different histograms to illustrate fitness values distribution.

3.1 Cluster analysis

For each combination of parameters, there are 30 sample sets evaluated and transformed into 30 normalized histograms of fitness values of the same function. In other words, each function (class) is represented by 30 feature vectors that should form a cluster. The better the clusters of different functions are distinguished, the more efficient are the histograms to describe the properties of functions. To evaluate this, the *silhouette score* is computed for all 30 sample sets for each of the 24 test classes as it validates the consistency of points within clusters of data. The silhouette measure ranges from -1 to +1, with a high value signifying strong cohesion within the object's cluster and a weak association with neighboring clusters. The final score is the mean of all silhouette values. The silhouette can be computed with an arbitrary distance measure. As it considers a pairwise distance comparison, we test it with both histogram distance and KL-divergence.

Table 1 and Table 2 show silhouette scores using both measures. The tables compare four sampling strategies. It is obvious that the score systematically rises with the greater number of generated samples (highest for $n = 2^{14}$). Higher dimensions require a higher number of samples to be thoroughly explored as the difference between sampling strategies is blurred. However, there is an obvious trend that LHSO and Uniform overcome the results of Sobol and G-Halton practically in all cases. Considering these statistics, Table 3 and Table 4 represent the Olympic medal ranking of strategies for histogram distance and KL-divergence respectively. The tables show that the 1st place belongs to LHSO, the 2nd to Uniform, the 3rd to G-Halton, and Sobol clearly comes out as the worst one. Comparing the distance measures, the KL-divergence strongly improves the silhouette scores over the histogram distance but the difference between rankings of strategies is minor.

In some cases, the differences between silhouette scores are relatively small. To strengthen the message of the presented experiments, the statistical significance was computed. Figure 1 and Figure 2 represent the critical distance (CD) plots [9] ranking strategies with the application of histogram distance and KL-divergence respectively. The plots consider the four sampling strategies for all combinations of problem dimensions and sample sizes. The non-parametric Friedman test and the post hoc Nemenyi test at significance level $\alpha = 0.05$ were applied [9]. The horizontal line connecting the ranks means that the difference between them is not statistically significant. The CD plots mostly show the superiority of LHSO and Uniform samplings that reach significantly better silhouette scores than G-Halton and Sobol. The plots look very similar for both distance measures and they mostly distinguish significantly LHSO and Uniform from G-Halton and Sobol.

Figure 3 illustrates the t-SNE visualizations of clusters computed for both distance measures. The points are the normalized histograms representing the 30 sample sets for each test function (i.e. $24 \cdot 30$ points). Each color represents one of the 24 test functions. There are 8 plots visualizing the effect of 4 sampling strategies and 2 distance measures. The assumption is that the points of the same class (function) are grouped into well-separated clusters. According to the computed silhouette scores, the parameters are set to $d = 5$ and $n = 2^{14}$ as this configuration best distinguishes the sampling strategies. The plots show the dominance of the LHSO and the Uniform samplings that form well-separated clusters while the G-Halton and the Sobol samplings produce a mixture of points with huge overlaps that lower the efficiency of normalized histogram to represent the underlying problems.

3.2 Distribution comparison

The previous section described the histograms as feature vectors. In this section, the specific normalized histograms are shown to illustrate their difference in various cases. All histograms are computed for $k = 50$ bins.

Figure 4 compares histograms of three selected test functions in dimension $d = 5$ for $n = 2^{14}$ uniform samples. It clearly demonstrates that the histograms differ in their shapes and skewness which indicates that these functions can be well-distinguished on the basis of normalized histograms.

The next experiment displays the effect of sampling strategies on two selected COCO test functions f002 (Figure 5) and f008 (Figure 5). The experiments were performed for two dimensions $d \in \{5, 20\}$ and sample size $n = 2^{14}$. The figures show that there are obvious differences between histograms produced by different sampling strategies, especially in 5-dimensional space. The higher dimension smooths the distribution of fitness values but the impact is also visible e.g. in the case of the f008 function. The Uniform and LHSO samplings lead to very similar smooth distributions in both dimensions, while the Sobol and G-Halton lead to disrupted histograms that strongly vary from the others. These plots correspond to the results obtained from the silhouette score analysis that pointed out the LHSO as the best sampling strategy for fitness landscape representation by normalized histograms.

4 CONCLUSION

The paper proposed a straightforward method for fitness landscape analysis based on the normalized histograms computed from fitness values gained by the evaluation of random samples scattered over the test functions. The test framework was based on the BBOB single-objective problems from the COCO library. The experiments tested the impact of different sampling strategies on the shape and representativeness of histograms. The results computed for different dimensions, sample sizes, and distances (histogram distance, KL-divergence) showed significant differences between histograms produced by different sampling strategies generating random samples. In this perspective, the best-performing strategy is the LHSO and the Uniform sampling significantly overcoming the G-Halton and Sobol samplings. This was verified by cluster analysis based on the silhouette score and distribution comparison using the visualization of selected histograms. To reveal the reason for this requires further research. However, we offer one hypothetical explanation. The G-Halton and Sobol low-discrepancy sequences were developed to maximally evenly sample the space which may lead to patterns that systematically miss out on some areas. This can be a strong effect, especially in the case of heavily rugged landscapes.

The presented results indicate that the usage of histograms as feature vectors representing the underlying continuous problems is efficient and it can distinguish different functions and types of problems with properly chosen sampling. It can also be used as an evaluation method for comparing discrete samplings. A bad sampling leads to a vague representation of landscapes that lowers the ability of a normalized histogram to serve as a feature vector for classification.

The future work will focus on a more comprehensive study of test problems, sampling strategies, other distance metrics, multi-objective functions, and the extension of this approach to other features from the ELA family. As the number of histogram bins is a user-defined parameter that controls the precision of the captured distribution, its effect should be studied as well.

ACKNOWLEDGMENTS

This work was supported by the project “Constrained multi-objective Optimization Based on Problem Landscape Analysis” funded by the Czech Science Foundation (grant no. GF22-34873K) and in part by the grant of the Student Grant System no. SP2023/12, VSB - Technical University of Ostrava.

Table 1: Silhouette score computed for the histogram distance

	$d = 5$			$d = 10$			$d = 20$		
n	2^{10}	2^{12}	2^{14}	2^{10}	2^{12}	2^{14}	2^{10}	2^{12}	2^{14}
Uniform	0.0710	0.1755	0.3349	0.0447	0.1114	0.2068	0.0269	0.0764	0.1729
Sobol	0.0335	0.0699	0.0872	0.0224	0.0641	0.1247	0.0231	0.0497	0.1009
G-Halton	0.0387	0.0742	0.1447	0.0419	0.1030	0.1844	0.0271	0.0781	0.1586
LHSO	0.0842	0.1897	0.3446	0.0493	0.1169	0.2115	0.0308	0.0774	0.1773

Table 2: Silhouette score computed for the KL-divergence

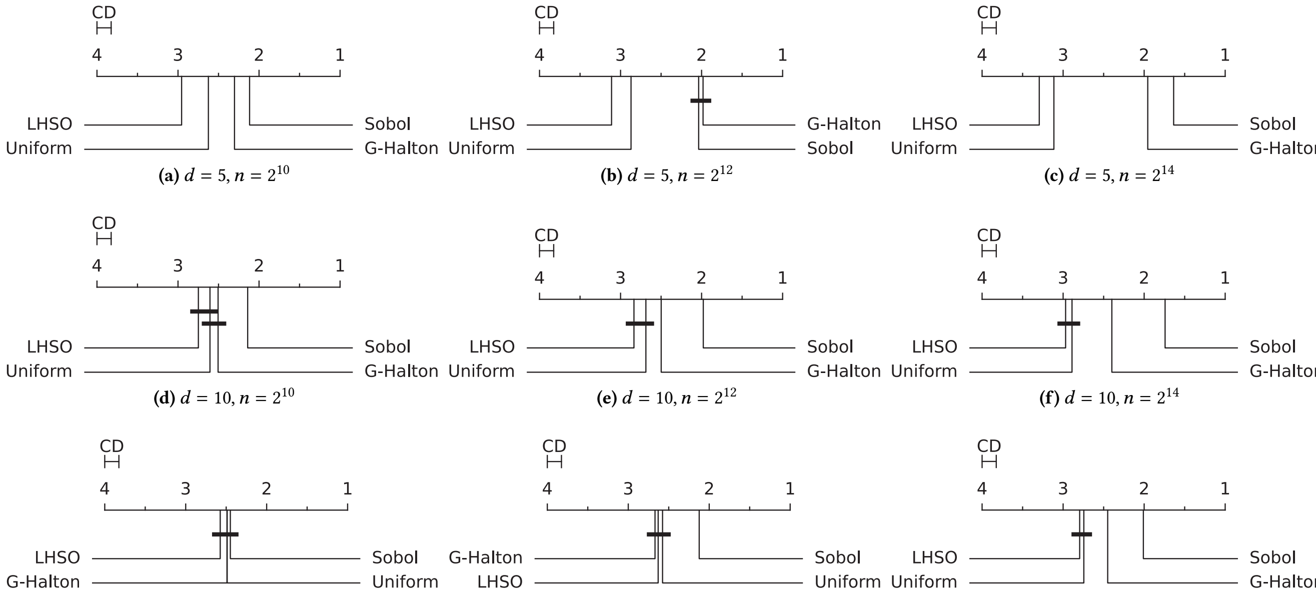
	$d = 5$			$d = 10$			$d = 20$		
n	2^{10}	2^{12}	2^{14}	2^{10}	2^{12}	2^{14}	2^{10}	2^{12}	2^{14}
Uniform	0.0736	0.2718	0.4792	0.0393	0.1707	0.2912	0.0152	0.1223	0.2648
Sobol	0.0194	0.0981	0.0806	0.0117	0.0730	0.1299	0.0011	0.0606	0.1160
G-Halton	0.0012	0.0963	0.1831	-0.0077	0.1582	0.2569	0.0154	0.1238	0.2323
LHSO	0.0963	0.2907	0.4933	0.0011	0.1790	0.2992	0.0252	0.1240	0.2683

Table 3: Olympic medal ranking of the sampling strategies based on the histogram distance

	Ranking			
	1st	2nd	3rd	4th
Uniform	0	7	2	0
Sobol	0	0	0	9
G-Halton	1	1	7	0
LHSO	8	1	0	0

Table 4: Olympic medal ranking of the sampling strategies based on the KL-divergence

	Ranking			
	1st	2nd	3rd	4th
Uniform	1	6	2	0
Sobol	0	1	2	6
G-Halton	0	2	4	3
LHSO	8	0	1	0

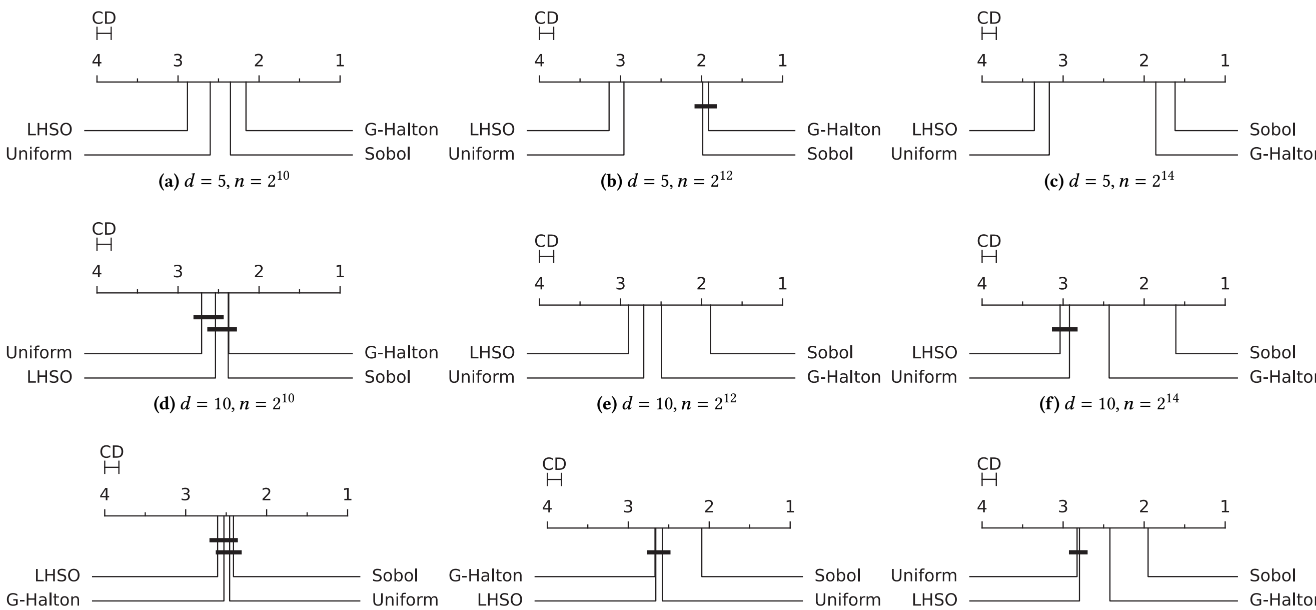


(a) $d = 5, n = 2^{10}$ (b) $d = 5, n = 2^{12}$ (c) $d = 5, n = 2^{14}$

(d) $d = 10, n = 2^{10}$ (e) $d = 10, n = 2^{12}$ (f) $d = 10, n = 2^{14}$

(g) $d = 20, n = 2^{10}$ (h) $d = 20, n = 2^{12}$ (i) $d = 20, n = 2^{14}$

Figure 1: Critical distance plot for histogram distance. Average ranks of sampling strategies computed for different sample sizes and dimensions based on the silhouette score. The difference between ranks connected by a horizontal line is not statistically significant at $\alpha = 0.05$.

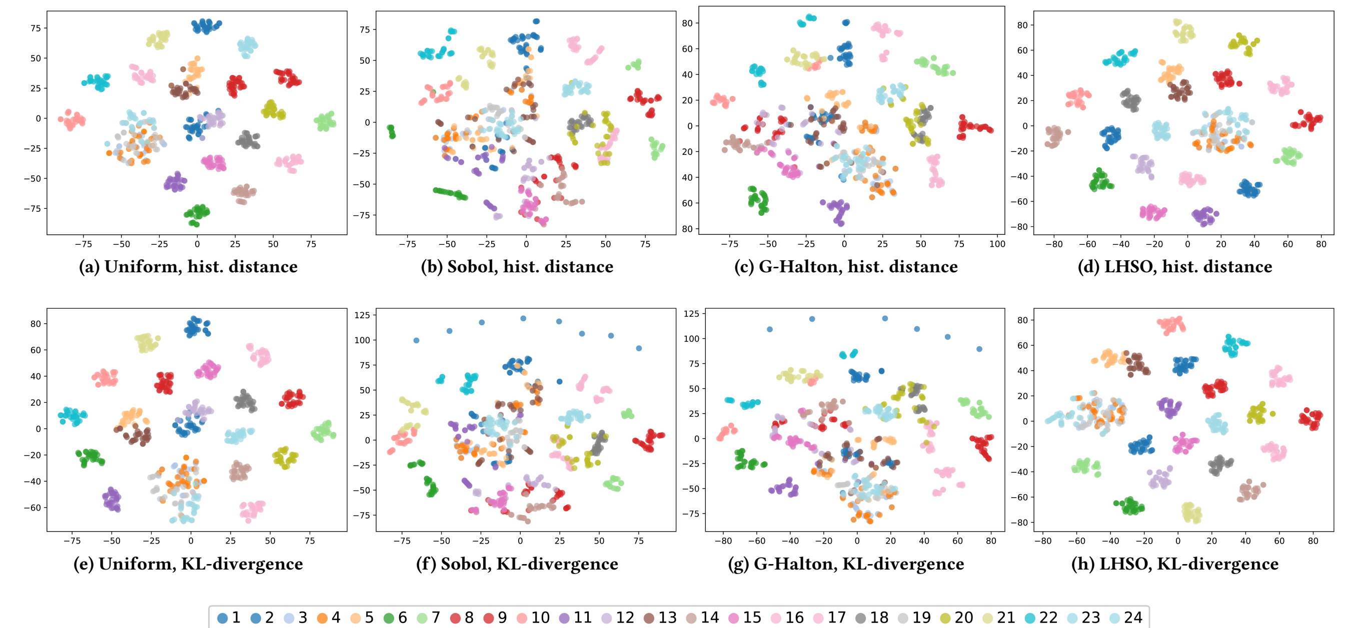


(a) $d = 5, n = 2^{10}$ (b) $d = 5, n = 2^{12}$ (c) $d = 5, n = 2^{14}$

(d) $d = 10, n = 2^{10}$ (e) $d = 10, n = 2^{12}$ (f) $d = 10, n = 2^{14}$

(g) $d = 20, n = 2^{10}$ (h) $d = 20, n = 2^{12}$ (i) $d = 20, n = 2^{14}$

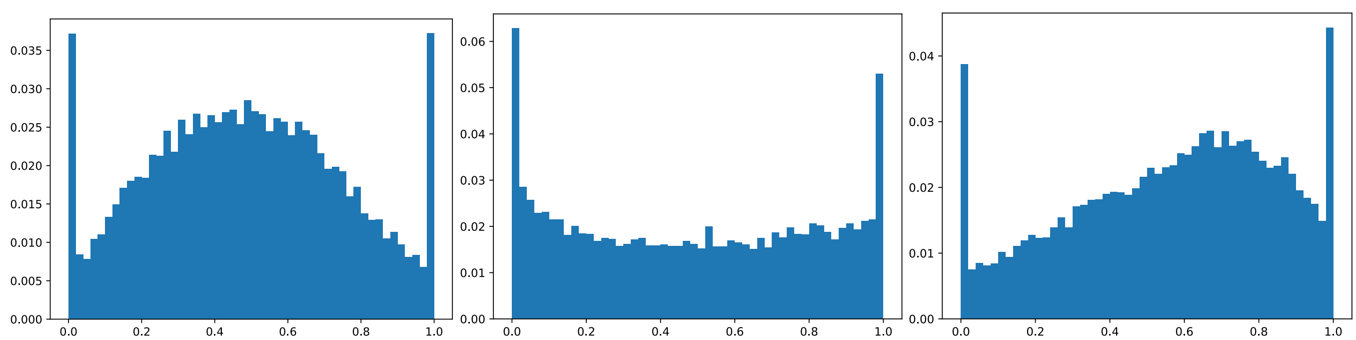
Figure 2: Critical distance plot for KL-divergence. Average ranks of sampling strategies computed for different sample sizes and dimensions based on the silhouette score. The difference between ranks connected by a horizontal line is not statistically significant at $\alpha = 0.05$.



(a) Uniform, hist. distance (b) Sobol, hist. distance (c) G-Halton, hist. distance (d) LHSO, hist. distance

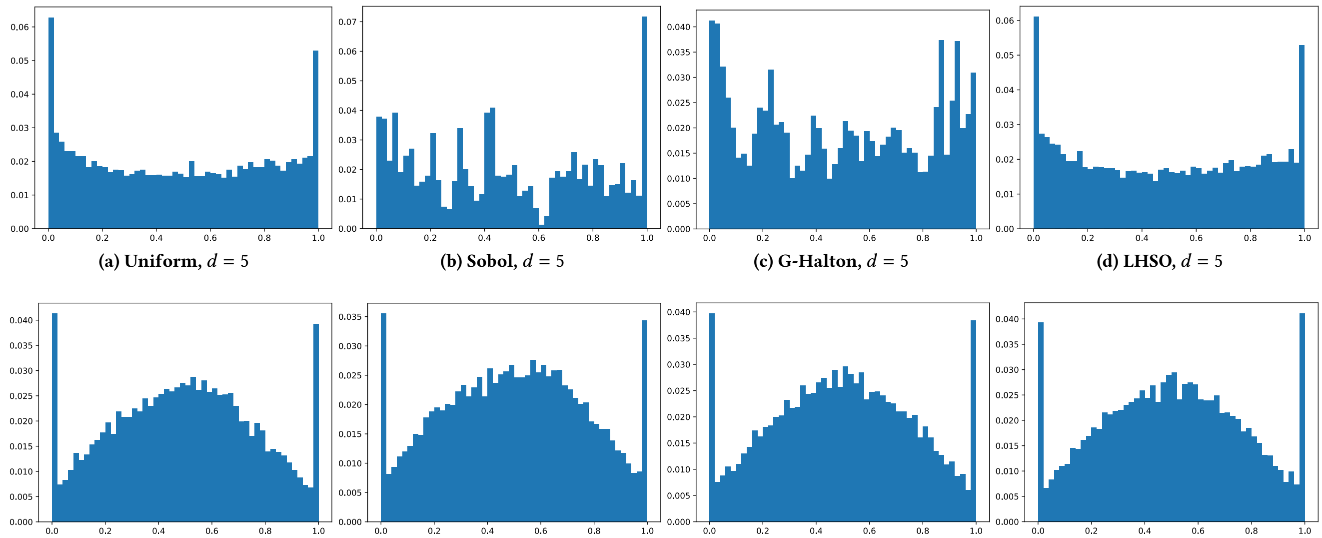
(e) Uniform, KL-divergence (f) Sobol, KL-divergence (g) G-Halton, KL-divergence (h) LHSO, KL-divergence

Figure 3: t-SNE visualization using histogram distance and KL-divergence measures of normalized fitness histograms computed from samples generated by 4 different samplings. The test functions have dimension $d = 5$ and the number of samples is $n = 2^{14}$. The colors represent the 24 COCO functions.



(a) f001 (b) f002 (c) f021

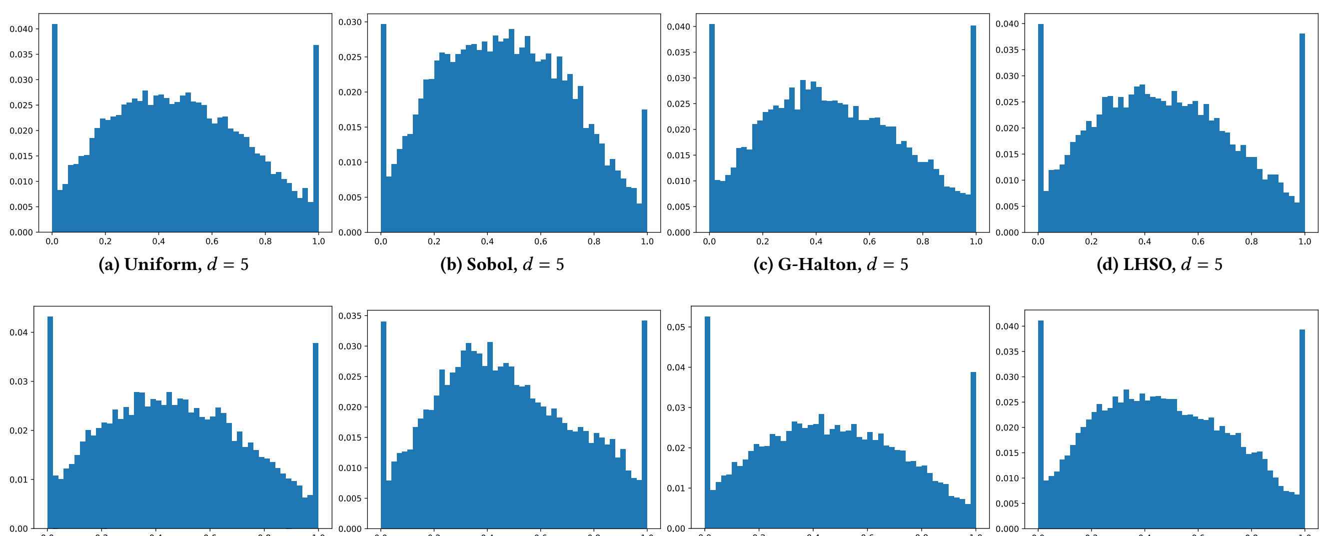
Figure 4: Comparison of fitness distribution of three COCO functions (normalized histograms with 50 bins, uniform sampling, $d = 5$, $n = 2^{14}$).



(a) Uniform, $d = 5$ **(b) Sobol, $d = 5$** **(c) G-Halton, $d = 5$** **(d) LHSO, $d = 5$**

(e) Uniform, $d = 20$ **(f) Sobol, $d = 20$** **(g) G-Halton, $d = 20$** **(h) LHSO, $d = 20$**

Figure 5: Comparison of fitness distributions of f002 COCO function (normalized histograms with 50 bins, 4 samplings, $d \in \{5, 20\}$, $n = 2^{14}$).



(a) Uniform, $d = 5$ **(b) Sobol, $d = 5$** **(c) G-Halton, $d = 5$** **(d) LHSO, $d = 5$**

(e) Uniform, $d = 20$ **(f) Sobol, $d = 20$** **(g) G-Halton, $d = 20$** **(h) LHSO, $d = 20$**

Figure 6: Comparison of fitness distributions of f008 COCO function (normalized histograms with 50 bins, 4 samplings, $d \in \{5, 20\}$, $n = 2^{14}$).

REFERENCES

[1] Andrejaana Andova, Aljoša Vodopija, Pavel Krömer, Vojtěch Uher, Tea Tušar, and Bogdan Filipič. 2022. Initial Results in Predicting High-Level Features of Constrained Multi-Objective Optimization Problems. In *Slovenian Conference on Artificial Intelligence, Proc. of the 25th Int. Multiconference Information Society*, Vol. A. 7–10.

[2] Annalisa Barla, Francesca Odone, and Alessandro Verri. 2003. Histogram intersection kernel for image classification. In *Proceedings 2003 international conference on image processing (Cat. No. 03CH37429)*, Vol. 3. IEEE, III–513.

[3] Sung-Hyuk Cha and Sargur N Srihari. 2002. On measuring the distance between histograms. *Pattern Recognition* 35, 6 (2002), 1355–1370.

[4] H. Chi, M. Mascagni, and T. Warnock. 2005. On the Optimal Halton Sequence. *Mathematics and Computers in Simulation* 70, 1 (2005), 9–21.

[5] Swagatam Das and Ponnuthurai Nagaratnam Suganthan. 2010. Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation* 15, 1 (2010), 4–31.

[6] Mahdiyeh Eslami, Hussain Shareef, Mohammad Khajehzadeh, and Azah Mohamed. 2012. A survey of the state of the art in particle swarm optimization. *Research Journal of Applied Sciences, Engineering and Technology* 4, 9 (2012), 1181–1197.

[7] John H Halton. 1964. Algorithm 247: Radical-inverse quasi-random point sequence. *Commun. ACM* 7, 12 (1964), 701–702.

[8] Nikolaus Hansen, Anne Auger, Raymond Ros, Olaf Mersmann, Tea Tušar, and Dimo Brockhoff. 2021. COCO: A platform for comparing continuous optimizers in a black-box setting. *Optimization Methods and Software* 36, 1 (2021), 114–144.

[9] Steffen Herbold. 2020. Autorank: A Python package for automated ranking of classifiers. *Journal of Open Source Software* 5, 48 (2020), 2173.

[10] Sourabh Katoch, Sumit Singh Chauhan, and Vijay Kumar. 2021. A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications* 80, 5 (2021), 8091–8126.

[11] Pascal Kerschke and Heike Trautmann. 2019. Automated algorithm selection on continuous black-box problems by combining exploratory landscape analysis and machine learning. *Evolutionary Computation* 27, 1 (2019), 99–127.

[12] P. Kromer, V. Uher, A. Andova, T. Tusar, and B. Filipic. 2022. Sampling Strategies for Exploratory Landscape Analysis of Bi-Objective Problems. In *2022 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE Computer Society, Los Alamitos, CA, USA, 336–342.

[13] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22, 1 (1951), 79 – 86.

[14] Yaxin Li, Jing Liang, Kunjie Yu, Ke Chen, Yinan Guo, Caitong Yue, and Leiyu Zhang. 2022. Adaptive local landscape feature vector for problem classification and algorithm selection. *Applied Soft Computing* 131 (2022), 109751.

[15] Katherine Mary Malan. 2021. A Survey of Advances in Landscape Analysis for Optimisation. *Algorithms* 14, 2 (Feb. 2021), 40.

[16] Katherine M Malan and Andries P Engelbrecht. 2014. Fitness landscape analysis for metaheuristic performance prediction. In *Recent advances in the theory and application of fitness landscapes*. Springer, 103–132.

[17] Katherine M. Malan and I. Moser. 2019. Constraint Handling Guided by Landscape Analysis in Combinatorial and Continuous Search Spaces. *Evolutionary Computation* 27, 2 (2019), 267–289.

[18] Michael D McKay, Richard J Beckman, and William J Conover. 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 42, 1 (2000), 55–61.

[19] Olaf Mersmann, Bernd Bischl, Heike Trautmann, Mike Preuss, Claus Weihs, and Günter Rudolph. 2011. Exploratory landscape analysis. In *Proceedings of the 13th Annual Genetic and Evolutionary Computation Conference (GECCO)*. ACM, 829–836.

[20] Peter Merz and Bernd Freisleben. 2000. Fitness landscape analysis and memetic algorithms for the quadratic assignment problem. *IEEE Transactions on Evolutionary Computation* 4, 4 (2000), 337 – 352.

[21] Sebastián Muñoz-Herrera and Karol Suchan. 2021. Constrained fitness landscape analysis of capacitated vehicle routing problems. *Entropy* 24, 1 (2021), 53.

[22] Jana Nowaková and Miroslav Pokornỳ. 2014. System identification using genetic algorithms. In *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014*. Springer, 413–418.

[23] Gabriela Ochoa and Nadarajen Veerapen. 2018. Mapping the global structure of TSP fitness landscapes. *Journal of Heuristics* 24, 3 (2018), 265–294.

[24] Quentin Renau, Carola Doerr, Johann Dreo, and Benjamin Doerr. 2020. Exploratory landscape analysis is strongly sensitives to the sampling strategy. In *Parallel Problem Solving from Nature – PPSN XVI*, Thomas Bäck, Mike Preuss, André Deutz, Hao Wang, Carola Doerr, Michael Emmerich, and Heike Trautmann (Eds.). Vol. 12270. Springer, 139–153.

[25] Quentin Renau, Johann Dreo, Carola Doerr, and Benjamin Doerr. 2021. Towards explainable exploratory landscape analysis: Extreme feature selection for classifying BBOB functions. In *Applications of Evolutionary Computation*, Pedro A. Castillo and Juan Luis Jiménez Laredo (Eds.). Springer, 17–33.

[26] Hendrik Richter. 2013. Dynamic Fitness Landscape Analysis. In *Evolutionary Computation for Dynamic Optimization Problems*, Shengxiang Yang and Xin Yao (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 269–297.

[27] Hendrik Richter and Andries Engelbrecht. 2014. *Recent advances in the theory and application of fitness landscapes*. Springer.

[28] Szabolcs Sergyan. 2008. Color histogram features based image classification in content-based image retrieval systems. In *2008 6th International Symposium on Applied Machine Intelligence and Informatics*. 221–224.

[29] Shinichi Shirakawa and Tomoharu Nagao. 2016. Bag of local landscape features for fitness landscape analysis. *Soft Computing* 20, 10 (2016), 3787–3802.

[30] Il'ya Meerovich Sobol. 1967. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7, 4 (1967), 784–802.

[31] Neil H Spencer. 2013. *Essentials of multivariate data analysis*. CRC press.

[32] Michael J Swain and Dana H Ballard. 1991. Color indexing. *International journal of computer vision* 7, 1 (1991), 11–32.

[33] Jorge Tavares, Francisco B. Pereira, and Ernesto Costa. 2008. Multidimensional knapsack problem: A fitness landscape analysis. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 38, 3 (2008), 604 – 616.

[34] Vojtěch Uher, Petr Gajdoš, Michal Radecký, and Václav Snášel. 2016. Utilization of the discrete differential evolution for optimization in multidimensional point clouds. *Computational Intelligence and Neuroscience* 2016 (2016), 6329530.

[35] Feng Zou, Debao Chen, Hui Liu, Siyu Cao, Xuying Ji, and Yan Zhang. 2022. A survey of fitness landscape analysis for optimization. *Neurocomputing* 503 (2022), 129–139.