Ectropy of diversity measures for populations in Euclidean space

Article *in* Information Sciences · June 2011

DOI: 10.1016/j.ins.2010.12.004 · Source: DBLP

2 authors:

Bakir Lacevic
University of Sarajevo
52 PUBLICATIONS **492** CITATIONS

SEE PROFILE

Edoardo Amaldi
Politecnico di Milano
122 PUBLICATIONS **3,113** CITATIONS

SEE PROFILE

Ectropy of diversity measures for populations in Euclidean space

Bakir Lacevic and Edoardo Amaldi

Politecnico di Milano, Dipartimento di Elettronica e Informazione, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
{*lacevic, amaldi*}*@elet.polimi.it*

Abstract

Measures to evaluate the diversity of a set of points (population) in Euclidean space play an important role in many areas of science and engineering, ranging from evolutionary algorithms and swarm intelligence to classifier systems and biology. Well-known measures are often used without a clear insight into their quality and many of them do not appropriately penalize populations with a few distant groups of collocated or closely located points. To the best of our knowledge, there is a lack of rigorous criteria to compare diversity measures and help select an appropriate one. In this work we define a mathematical notion of ectropy for classifying diversity measures in terms of the extent to which they tend to penalize point collocation, we investigate the advantages and disadvantages of several known measures and we propose some novel ones. In particular, we introduce a quasi-entropy measure based on a geometric covering problem, three measures based on discrepancy from uniform distribution and one based on Euclidean minimum spanning trees. All considered measures are tested and compared on a large set of random and structured populations. Special attention is also devoted to the complexity of computing the measures. The measure based on Euclidean minimum spanning trees turns out to be the most promising one in terms of the tradeoff between the computational complexity and the ectropic behavior.

Keywords: Population diversity measure, Ectropy, Klee's measure, Minimum spanning tree, Singular values, Discrepancy

1. Introduction

Given a population, i.e., a set of points (vectors, individuals) in a metric space, its diversity is undoubtedly one of the most important indicators of its state. Measures to evaluate the diversity of a population play an important role and are studied (or just used) in a variety of areas such as evolutionary algorithms [65, 62, 1, 21, 51, 6, 4, 3, 77, 78, 39, 22, 32, 29, 70, 68, 12, 15, 33, 11, 44, 50, 72, 53, 71, 48, 58]; swarm intelligence [59, 30, 61, 9]; numerical integration, quasi Monte Carlo methods and motion planning [28, 38, 66, 18, 47, 19, 42, 43, 13, 24]; classifier systems [79, 23, 55, 34, 64, 17], biology, environmental sciences and econometrics [69, 73, 67, 52, 14, 45, 63]. Surprisingly, not much effort has been undertaken so far for a systematic analysis and comparison of existing population diversity measures. Moreover, there is a lack of rigorous criteria to assess the quality of population diversity measures and compare them.

In the seminal paper [69], Weitzman discusses the existence of a diversity function that captures the "value of diversity" of a collection of points and proposes a function based on the dissimilarities between points. Assuming ultrametricity of the dissimilarity measure, the diversity function can theoretically be generated recursively using dynamic programming. In [50] Morrison and De Jong give a brief history of population diversity measures and introduce a measure that extends the concept of moment of inertia to arbitrarily high dimensional spaces. Wineberg and Opacher informally define a diversity measure as the answer to the question "how different is everybody from everybody else?" [72]. In [72, 53, 71] they show that commonly used measures (mainly from evolutionary computation) reduce to cumulated difference between all possible pairs of points in the population (this also holds for the measures described in [50]). Lacevic et al. [37] point out that the measures based on the sum (average) of pairwise distances between points in $\mathbb{R}^m$ have substantial shortcomings. For instance, these measures may reach their maximum value when the population consists of very few (sometimes only two) mutually distant clusters of collocated or closely located points. However, the approach presented lacks generality w.r.t. the space dimension. Mattiussi et al. [48] analyze different diversity measures for populations of strings of possibly variable length, paying special attention to

the computational complexity issue. In [34] and [64], the authors provide an extensive analysis of various diversity measures for constructing classifier ensembles and a deeper understanding of the principle of diversity, but the specific nature of the field substantially mitigates the generality of the concepts. Similarly, the notion of diversity is ubiquitous in ecology and related disciplines, and various (prevalently entropy-based) indices are used to estimate biodiversity [73, 67, 52, 14, 45, 63]. The very interesting problem of uniform design is considerably studied in the literature (see e.g. [28, 38, 66, 18, 47, 19, 42, 43]) with a focus on uniform sampling (used for numerical integration, quasi Monte Carlo method, computer graphics, etc.). Various discrepancy functions are designed to measure the deviation of a given point set from a uniform distribution and plenty of generators of "small-discrepancy" sequences are available, e.g. [38, 66]. However, we know of no attempt to explicitly bring this theory together with the related concept of diversity.

In many applications, a high-quality diversity measure is needed to evaluate the state of a population. Diversity measures are typically used for pure a posteriori analysis or estimates are used to predict some suitable parameter setups that implicitly affect diversity [6, 4, 3, 77]. An accurate diversity measure is also particularly important when some decisions are influenced by the diversity via feedback, for instance, when parameters or the structure of an algorithm change (in terms of measured diversity) during the execution [65, 62, 77, 59, 30, 79]. Furthermore, a proper insight into the nature of a diversity measure could help improve algorithms that aim at diversity enhancement [57, 76, 46].

Several approaches are widely used for measuring population diversity. In a first approach the diversity is evaluated in genotypic space (when the population consists of binary strings or strings of "genes" from any alphabet) using either deterministic formulas (e.g. Hamming distance or another metric) [62, 1, 21, 72, 53, 71, 48] or entropy calculations [51, 15, 49, 72, 53, 71, 48, 73, 67, 52, 14, 45, 63]. In a second approach the diversity is computed in parameter (phenotypic) space, usually in $\mathbb{R}^m$, where $m \in \mathbb{N}$. The two most frequently used measures are based on summing (averaging) the distances of all the points to the centroid point (see [65, 6, 72, 53, 71, 59, 30, 61, 46]) or the distances between all pairs of points in the population (see [51, 4, 3, 72, 53, 71, 48, 61, 46]). Similar techniques include column-based variance and moment of inertia diversity measure [50]. In spite of the different expressions, all of these measures rely on the distances between all possible pairs of points [72]. Apparently, this is also true for the Shannon entropy-based diversity measure [72].

The main focus of this paper is on diversity measures in $\mathbb{R}^m$, although some attention is also devoted to measures for binary coded populations. The main contributions are the following[1]:

- We introduce a mathematical notion of ectropy that allows classification of different diversity measures. The level of ectropy indicates to what extent the diversity measure tends to penalize the collocation/proximity of points. The definition of ectropy is then used to point out some drawbacks of the classical measures.
- We propose several novel measures that can overcome the drawbacks of the classical ones. They are designed to capture how uniformly the population is distributed in the domain space. In particular, we introduce a quasi-entropy measure based on a geometric covering problem, three measures based on discrepancy from uniform distribution and one based on Euclidean minimum spanning trees. Furthermore, a special attention is devoted to the complexity of computing the diversity measures.
- An extensive computational study is conducted in order to establish the mutual dependencies among different measures. This study supports the theoretical analysis performed for a variety of measures. On the other hand, it provides valuable insight into the behavior of those measures for which the rigorous analysis could not be carried out.

The remainder of the paper is organized as follows. In Section 2, the ectropic property of the diversity measure is defined, while in Section 3, some drawbacks of the classical diversity measures are pointed out. In Section 4 we analyze several alternative diversity measures and prove that they overcome some of the shortcomings of classical measures. Our setup for the extensive computational study is presented in Section 5, along with the numerical results and comments. Finally, Section 6 contains some concluding remarks and future work directions.

[1] Selected preliminary results of this paper can be found in [36].

2. Ectropic property of diversity measures

2.1. Notation

Let $X \subseteq \mathbb{R}^m$, with $m \in \mathbb{N}$ and let $\mathbf{P}(X,n) = (\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n)\,,\ n \in \mathbb{N}$ be the population of n points, given in a matrix form. A column $\mathbf{x}_k$ represents a k-th point $\mathbf{x}_k = (x_{1k}\ x_{2k}\ \ldots\ x_{mk})^T \in X,\ k = 1, 2, \ldots, n$. Whenever convenient, a population $\mathbf{P}$ is represented in the form of the set $\mathbf{P} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$. Then the cardinality $|\mathbf{P}(X,n)|$ of population $\mathbf{P}(X,n)$ is at most n and is strictly smaller than n when some points are collocated. We say that $\mathbf{P}(X,n) \in E(X,n) \subseteq \mathbb{R}^{m\times n}$, where $E(X,n)$ represents the set of all $m\times n$ matrices whose columns (points) belong to X. Further, let $D(\mathbf{P}(X,n))$ be a diversity measure ($D : E(X,n) \to \mathbb{R}$) and let $E^*_D(X,n) \subseteq E(X,n)$ be the set of all populations $\mathbf{P}(X,n) \in E(X,n)$ for which the diversity measure $D(\mathbf{P}(X,n))$ reaches its maximum value, i.e.,

$$E^*_D(X,n) = \{\mathbf{P}(X,n) \in E(X,n) \mid \forall \mathbf{Q}(X,n) \in E(X,n) : D(\mathbf{P}(X,n)) \geq D(\mathbf{Q}(X,n))\} \tag{1}$$

2.2. Definition of ectropic property

In this section we define a new property of diversity measures that provides some insight into their inherent bias towards order/disorder within the population. Order implies emphasized occurrence of collocation between points, while disorder means that the population is scattered more uniformly over its domain. The notion of *ectropy* draws its origin from its antagonistic counterpart *entropy* that can be viewed as the measure of disorder in the population.

Definition 2.1. *A diversity measure $D(\boldsymbol{P}(X,n))$ is (ρ,ε)-ectropic on X if the following holds:*

$$\sup_{|\boldsymbol{P}(X,n)|=\rho n} D(\boldsymbol{P}(X,n)) = (1-\varepsilon) \sup_{\boldsymbol{P}(X,n)} D(\boldsymbol{P}(X,n))$$

where $\rho \in \left\{\frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}$ and $\varepsilon \in [0,1]$.

From now on, we assume that the set X is a compact and that the suprema in the equation above exist. The intuition behind this definition is as follows. Assume the population $\mathbf{P}_c(X,n)$ is a set of n points in X that is constrained to be of cardinality ρn. If $\rho < 1$, some of the points have to be collocated. If the maximum of $D(\mathbf{P}_c(X,n))$ over all such populations, equals $(1-\varepsilon)$ times the maximum value of $D(\mathbf{P}_u(X,n))$, where $\mathbf{P}_u(X,n)$ is an unconstrained population, we say that the measure $D(\mathbf{P}(X,n))$ is (ρ,ε)-*ectropic* on X. If the measure $D(\mathbf{P}(X,n))$ is (ρ,ε)-*ectropic* while both ρ and ε are small (say close to zero), it means that the diversity measure might reach its near-maximum value for a population that has many collocated points (e.g. consists of several clusters of collocated points). This property is clearly in contrast with the intuitive notion of diversity measure because the collocation of points is rewarded instead of being penalized.

An example of how informative the ectropic property can be is given in Fig. 1. The left figure depicts the population of size $n = 10$ that may have arbitrary cardinality, i.e., $\rho \leq 1$. The right figure shows a population of the same size $n = 10$ but with at least one collocation and hence a cardinality at most 9, yielding $\rho \leq \frac{9}{10}$. Now assume the measure D is computed for the two types of populations. If a measure D is $(\frac{9}{10}, 0)$-ectropic, the collocation is not penalized and D can reach its maximum value for the population of the second type. Thus such a diversity measure does not fully capture the extent to which the points are uniformly distributed in the domain. On the other hand, a diversity measure D that is $(\frac{9}{10}, \frac{1}{10})$-ectropic would mean that with 90% of the resources (in terms of cardinality) the measure D can go up to 90% of its full range. Such a measure better suits the applications where we aim at capturing the uniformity of the points distribution.

At first, the concept of collocation may appear too radical when trying to highlight the tendencies of some diversity measures to reward non-uniformly distributed populations. In particular, some optimization algorithms that deal with populations in Euclidean space (e.g., real-valued evolutionary algorithms or particle swarm optimizers) do not usually generate collocated points. Nevertheless, by pointing out that the measure $D(\mathbf{P}(X,n))$ reaches its (near)maximum value while having some points in $\mathbf{P}(X,n)$ collocated we can also expect a similar behavior for the populations with full cardinality n that have some points very close to each other. By introducing small perturbations in the population with collocated points such that cardinality becomes equal to n, the measure $D(\mathbf{P}(X,n))$ is still close to its maximum value provided that the mapping $D : E(X,n) \to \mathbb{R}$ is continuous.

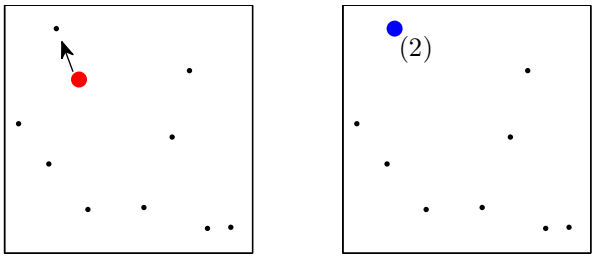

Figure 1: The intuition behind the ectropic property of the diversity measure. The population of cardinality ≤ 10 (left) and the population of the same size but of cardinality ≤ 9 (right). If a diversity measure D can reach its maximum value for the population on the right, then it is $(\frac{9}{10}, 0)$-ectropic.

In this paper, we will often assume that the set X is a hyperrectangle, i.e., $X = \bigotimes_{i=1}^{m} [a_i, b_i]$, where "$\bigotimes$" denotes the Cartesian product. This corresponds to the scenario where each component x_{ij} from matrix $\mathbf{P}(X, n)$ belongs to a predefined interval $[a_i, b_i]$, which is a very common case in a variety of fields of application (EAs, PSO, etc.). A suitable affine coordinate transformation can be applied to map X into a unit hypercube $[0, 1]^m$.

3. Classical diversity measures

One of the two most frequently used diversity measures is based on summing (averaging) the Euclidean distances from every point to the center-point [65, 6, 72, 53, 71, 59, 30, 61, 46]:

$$D_v(\mathbf{P}(X, n)) = \sum_{i=1}^{n} \left\| \mathbf{x}_i - \overline{\mathbf{x}} \right\|, \tag{2}$$

where $\overline{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i = [M_1 \ M_2 \ \ldots \ M_m]^T$ is the centroid of the population with $M_k = \frac{1}{n} \sum_{j=1}^{n} x_{kj}$, $k = 1, 2, \ldots, m$. The computational complexity of $D_v(\mathbf{P}(X, n))$ is obviously $O(mn)$.

Another popular measure is based on summing (averaging) the Euclidean distances between all pairs of points [51, 4, 3, 72, 53, 71, 48, 61, 46]:

$$D_d(\mathbf{P}(X, n)) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left\| \mathbf{x}_i - \mathbf{x}_j \right\|, \tag{3}$$

with a complexity of $D_d(\mathbf{P}(X, n))$ is $O(mn^2)$.

Theorem 3.1. *The diversity measure $D_v(\boldsymbol{P}(X, 2p))$ is $\left(\frac{1}{p}, 0\right)$-ectropic on $X = \bigotimes_{i=1}^{m} [a_i, b_i]$, where $a_i, b_i \in \mathbb{R}$, $a_i < b_i$, $(i = 1, 2, \ldots, m)$, $p \in \mathbb{N}$, $p \geq 2$.*

Proof. We derive an upper bound on $D_v(\mathbf{P}(X, n))$. Using the well-known inequality between arithmetic and quadratic mean, we have:

$$D_v^2(\mathbf{P}(X, n)) = \left(\sum_{j=1}^{n} \sqrt{\sum_{i=1}^{m} \left(x_{ij} - M_i \right)^2} \right)^2 \leq n \sum_{j=1}^{n} \sum_{i=1}^{m} \left(x_{ij} - M_i \right)^2 = n \sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_{ij} - M_i \right)^2 \equiv \overline{D_v^2}(\mathbf{P}(X, n)) = n \sum_{i=1}^{m} \overline{D_i}, \tag{4}$$

where $\overline{D_k} = \sum_{j=1}^{n} \left(x_{kj} - M_k \right)^2$, $k = 1, \ldots, m$. Since $\overline{D_v^2}(\mathbf{P}(X, n))$ is additively separable in terms of the $\overline{D_k}$, with $1 \leq k \leq m$, we can derive the upper bounds on each $\overline{D_k}$ separately.

$$\begin{aligned} &\overline{D_k} = \sum_{j=1}^{n} \left(x_{kj} - M_k \right)^2 = \sum_{j=1}^{n} x_{kj}^2 - 2M_k \sum_{j=1}^{n} x_{kj} + nM_k^2 = \sum_{j=1}^{n} x_{kj}^2 - nM_k^2 = \sum_{j=1}^{n} x_{kj}^2 - \frac{1}{n} \left(\sum_{j=1}^{n} x_{kj} \right)^2 \\ &\equiv f(x_{k1}, \ x_{k2}, \ \ldots, \ x_{kn}) \geq 0. \end{aligned} \tag{5}$$

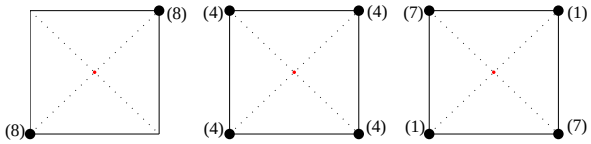

Figure 2: Some point arrangements that maximize the diversity D_v (the numbers in bracket indicate the cluster size).

It can easily be shown that the function f is convex and hence reaches its maximum at a vertex of the (convex) polytope $[a_k,\ b_k]^n$ [60]. To maximize the function f, we observe the permutation $x_{s1}, x_{s2}, \ldots, x_{sn}$ of $x_{k1},\ x_{k2},\ \ldots,\ x_{kn}$ with $x_{s1} = \ \ldots\ = x_{sj} = a_k$, $x_{sj+1} = \ \ldots\ = x_{sn} = b_k$. We have:

$$\begin{aligned}&\overline{D_k} = \textstyle\sum_{i=1}^{j} x_{si}^2 + \sum_{i=j+1}^{n} x_{si}^2 - \frac{1}{n}\left(\sum_{i=1}^{j} x_{si} + \sum_{i=j+1}^{n} x_{si}\right)^2 = \left[j a_k^2 + (n-j) b_k^2\right]\\ &-\frac{1}{n}\left[j a_k + (n-j) b_k\right]^2 = (a_k - b_k)^2\,\frac{j(n-j)}{n} = (a_k - b_k)^2\, g(n, j).\end{aligned} \tag{6}$$

The function $g(n, j) = j(n - j)$ achieves its maximum value when $j = \lfloor \frac{n}{2} \rfloor$, i.e., $\overline{D_k} \le \frac{1}{n}\left[\lfloor \frac{n}{2} \rfloor \left(n - \lfloor \frac{n}{2} \rfloor\right)\right](a_k - b_k)^2 \equiv \overline{D}_{k\max}$. Finally, from (4) and (6) we have:

$$\overline{D_v^2}(\mathbf{P}(X,n)) = n\sum_{i=1}^{m}\overline{D}_i \le n\sum_{i=1}^{m}\overline{D}_{i\max} = n\sum_{i=1}^{m}\frac{1}{n}(a_i - b_i)^2\left[\lfloor \frac{n}{2} \rfloor\left(n - \lfloor \frac{n}{2} \rfloor\right)\right] = \left[\lfloor \frac{n}{2} \rfloor\left(n - \lfloor \frac{n}{2} \rfloor\right)\right]\sum_{i=1}^{m}(a_i - b_i)^2\,. \tag{7}$$

In (4) the equality holds if and only if $\left\|\mathbf{x}_1 - \overline{\mathbf{x}}\right\| = \left\|\mathbf{x}_2 - \overline{\mathbf{x}}\right\| = \ \ldots\ = \left\|\mathbf{x}_n - \overline{\mathbf{x}}\right\|$. However, $\overline{D}_k$ reaches $\overline{D}_{k\max}$ if and only if $j = \lfloor \frac{n}{2} \rfloor$. The above equality conditions are easily satisfied when n is even and $n \ge 4$. For example, let $\frac{n}{2}$ points have coordinates $[a_1\ a_2\ \ldots\ a_m]^T$ and other $\frac{n}{2}$ points have coordinates $[b_1\ b_2\ \ldots\ b_m]^T$. Thus $D_v(\mathbf{P}(X,n))$ achieves its maximum value for a population consisting of two clusters at the opposite ends of the main diagonal of the hyperrectangle. Now, the theorem's claim follows directly from the Definition 2.1. □

Fig. 2 depicts three cases where $m = 2$ and $D_v(\mathbf{P}(X, 16))$ reaches its maximum value. For n odd, it is more difficult to create an intuitive arrangement that would obviously satisfy both equalities at once, but it is natural to expect similar behavior of $D_v(\mathbf{P}(X,n))$.

To investigate the degree of ectropy of the diversity measure $D_d(\mathbf{P}(X,n))$, we need the following lemma.

Lemma 3.1. *Consider a set of s points (foci) $A = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s\}$ in $\mathbb{R}^m$ and a given positive real constant r. The set $C = \left\{x \in \mathbb{R}^m \mid \sum_{i=1}^{s} \|\boldsymbol{x} - \boldsymbol{a}_i\| \le r\right\}$ is strictly convex.*

The simple proof can be found in [35].

Theorem 3.2. *The diversity measure $D_d(\boldsymbol{P}(X, 2^m + p))$ is $(\rho, 0)$-ectropic on a hyperrectangle $X = \bigotimes_{i=1}^{m}[a_i, b_i]$, where $a_i, b_i \in \mathbb{R},\ \ a_i < b_i,\ (i = 1, 2, \ldots, m),\ p \in \mathbb{N},\ \rho \le \frac{2^m}{2^m + p}$.*

Proof. It is sufficient to prove that, if the population $\mathbf{P}(X,n) \in E^*_{D_d}(X,n)$, all the points must be on the vertices of X. Then, for any $n = 2^m + p > 2^m$, some points must be collocated and according to Definition 2.1, the theorem's claim holds. Assume the opposite, i.e., let $\mathbf{P}_1(X,n) = (\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n) \in E^*_{D_d}(X,n)$ and that there exists at least one point $\mathbf{x}_k,\ k \in \{1, 2, \ldots, n\}$ from $\mathbf{P}_1(X,n)$ that differs from any of the vertices of X. The diversity measure $D_d(\mathbf{P}_1(X,n)$ has a maximum value $D_{d\max}(X,n)$ and may be written as:

$$D_d(\mathbf{P}_1(X,n) = \sum_{\substack{i=1\\ i\neq k}}^{n-1}\sum_{\substack{j=i+1\\ j\neq k}}^{n}\|\mathbf{x}_i - \mathbf{x}_j\| + \sum_{\substack{i=1\\ i\neq k}}^{n}\|\mathbf{x}_k - \mathbf{x}_i\| = D_{d\max}(X,n). \tag{8}$$

If we "fix" all points except $\mathbf{x}_k$, then the first term of the right hand side of (16) can be treated as a constant, while the second one as a function $h(\mathbf{x}_k)$ of $\mathbf{x}_k$:

$$h(\mathbf{x}_k) \equiv D_{d\max}(X,n) - \sum_{\substack{i=1\\ i\neq k}}^{n-1}\sum_{\substack{j=i+1\\ j\neq k}}^{n}\|\mathbf{x}_k - \mathbf{x}_i\| \equiv r_k. \tag{9}$$

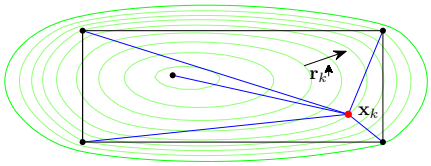

Figure 3: Contour lines of $h(\mathbf{x}_k)$ (boundaries of $C(r_k)$) : $h(\mathbf{x}_k)$ increases as $\mathbf{x}_k$ moves toward a vertex

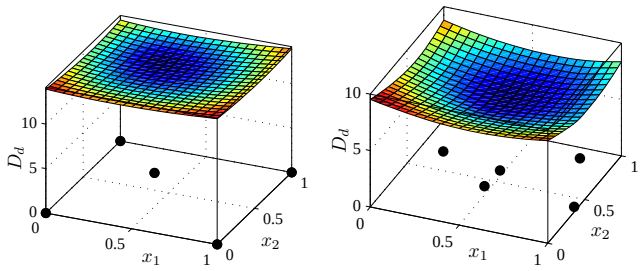

Figure 4: Profile of D_d diversity - simple $2D$ examples

According to Lemma 2.1, the set $C = \left\{\mathbf{x} \in X \mid \sum_{i=1}^{s} \|\mathbf{x} - \mathbf{x}_i\| \leq r_k\right\}$ is convex and $\mathbf{x}_k$ lies on its boundary. The boundary of C represents the set of all the points $\mathbf{x}_k$ for which the function h from (9) achieves its maximum value. However, by increasing r_k, the Lebesgue measure of the set C also increases while its boundary ∂C certainly does not intersect with X. If it still does, there would be a point $\mathbf{y}_k \in X$ such that $h(\mathbf{x}_k) < h(\mathbf{y}_k)$ and the population $\mathbf{P}_1(X, n)$ would not belong to $E^*_{D_d}(X, n)$. On the other hand, since C is strictly convex, and X is a polytope, the maximum value r_k, for which the boundary ∂C still intersects X is such that this intersection occurs at some vertex of X (see Fig. 3 for a $2D$ illustration). This is analogous with the fact that the convex function defined on a polytope reaches its maximum at the polytope's vertex (or vertices). This means that $\mathbf{x}_k$ has to be a vertex of X. □

Fig. 4 illustrates the profile of $D_d(\mathbf{P}(X, n))$ for $m = 2$ and $n = 6$, where five points are fixed (dots in x_1x_2 plane) and the sixth point takes all possible values in $[0, 1]^2$. Obviously, all local maxima of the measure coincide with the domain vertices.

4. Alternative measures

In the previous section, some drawbacks of commonly used measures have been pointed out. Although only two measures were considered, our approach has a much broader scope. For instance, expressions like the moment of inertia based measure or column-based variance implicitly appear in (4). Moreover, according to analysis presented in [72], all of these measures have very similar behavior.

In the sequel of this section we propose novel diversity measures that are designed in order to overcome the shortcomings of classical ones. Several measures, recently proposed in [37, 36], are also considered and their ectropic property is studied. For the sake of generality, three typical diversity measures for binary-coded populations are also described.

4.1. Volume-based diversity measure (L-diversity)

The rationale of this diversity measure is to consider for each point a compact cell and compute the volume of the union of these cells when they are centered in the corresponding points [37]. A $3D$ example is illustrated in Fig. 5. Clearly, the larger the volume of the union of the cells the more scattered the points of the population. Closely located points imply large overlapping of the corresponding cells and hence a smaller contribution to the overall measure. Informally, the L-diversity provides unambiguous information about fraction of the domain X that is populated by $\mathbf{P}(X, n)$.

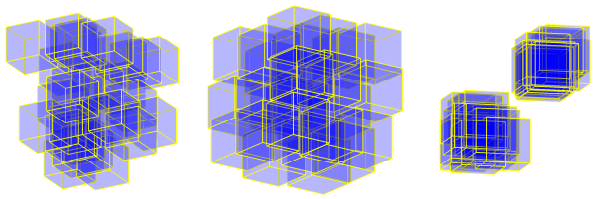

Figure 5: Diversity as the volume of the union of cells assigned to the points. Uniform random distribution of points (left), slightly perturbed lattice point set (middle), and two non-overlapping clusters (right).

To define L-diversity for any dimension, we assign to each point $\mathbf{x}_k = (x_{1k} \;\; x_{2k} \;\; \ldots \;\; x_{mk})^T$ the hypercube $S(\mathbf{x}_k, d) = \left\{(a_1 \;\ldots\; a_m)^T \in \mathbb{R}^m \;:\; |x_{ik} - a_i| \leq \frac{d}{2},\; i = 1, \ldots, m\right\},\; d \geq 0$ and consider the mapping $L : \mathbb{R}^{m \times n} \to \mathbb{R}$ with:

$$L(\mathbf{P}(X, n), d) = \mu_L\left(\bigcup_{i=1}^{n} S(\mathbf{x}_i, d)\right), \tag{10}$$

where $\mu_L(A)$ denotes the Lebesgue measure of a set A. The variable d represents the dimension of the hypercube $S(\mathbf{x}_k, d)$. A reasonable value can be derived from:

$$nd^m = \mu_L(X). \tag{11}$$

Similar definitions can be obtained by considering hyperectangles or any compact set instead of hypercubes. A naive algorithm for calculating the L-diversity is the inclusion-exclusion principle based on formula [22]:

$$\mu_L\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu_L(A_i) - \sum_{1 \leq i < j \leq n}^{n} \mu_L(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n}^{n} \mu_L(A_i \cap A_j \cap A_k) - \ldots + (-1)^n \mu_L\left(A_1 \cap A_2 \cap \ldots A_n\right). \tag{12}$$

An obvious drawback of this algorithm is its computational complexity, which is $O(m2^n)$. Fortunately, the problem of determining the volume of the union of axis aligned hyperrectangles (often referred to as Klee's measure problem [31]) has been extensively investigated in the literature. In [56] algorithms running in $O(n \log n)$ time, for dimensions $m = 1$ and $m = 2$, are described. In dimensions $m \geq 3$, Bentley's algorithm runs in $O(n^{m-1} \log n)$ time [7], van Leeuwen and Wood's algorithm in $O(n^{m-1})$ time [41], and Overmars and Yap's algorithm in $O(n^{\frac{m}{2}} \log n)$ time [54], which is the fastest known. For the special case where the assigned hyperrectangles are actually identical hypercubes (which is of particular interest for the proposed measure), an algorithm that runs in $O(n^{\lfloor \frac{m}{2} \rfloor})$ is described in [10].
The principle of L-diversity is somewhat motivated by the concept of S-metric (dominated hypervolume or just hypervolume): a popular method of performance assessment in evolutionary multiobjective optimization [78, 39, 22, 32, 29, 70, 12]. Usual techniques of calculating S-metric include "LebMeasure" [22], [32] and HSO (hypervolume by slicing objectives) [70], but not before the work of Beume and Rudolph [8] has the S-metric been connected to its obvious counterpart - Klee's measure problem. Despite its polynomial time computability (for fixed m), it is clear that the algorithm becomes very expensive for high dimensions. Therefore, one of the goals of this work is to find the measure that highly correlates with L-diversity, but has smaller computational complexity. The following analysis shows some interesting features of L-diversity and somewhat justifies considering it as a reference measure.

Theorem 4.1. *If X is a hypercube $[a, b]^m$, then for arbitrary $n \in \mathbb{N}$, there exists a population $\mathbf{P}^*(X, n)$, such that $L(\mathbf{P}^*(X, n), d) = (b - a)^m = \sup_{\mathbf{P}(X,n)} L(\mathbf{P}(X, n), d)$, where $nd^m = (b - a)^m$.*

Proof. For any $n \in \mathbb{N}$, we exhibit a point arrangement $\mathbf{P}^*(X, n)$ such that no intersection occurs when the hypercubes of side d are assigned to the points. Without the loss of generality, assume $a = 0$ and $b = 1$. No intersection (with non-zero m-volume) will occur iff for any pair of points $\mathbf{x}_j, \mathbf{x}_k \in \mathbf{P}^*(X, n)$ we have that $\max_{i=1,\ldots,m} |x_{ij} - x_{ik}| \geq d = \frac{1}{\sqrt[m]{n}}$. For $p^m + 1 \leq n \leq (p+1)^m,\; p \geq 1$, we can pick the population of full cardinality from the lattice point set $\left\{0, \frac{1}{p}, \frac{2}{p}, \ldots, 1\right\}^m$. In this case $\max_{i=1,2,\ldots,m} |x_{ij} - x_{ik}| \geq \frac{1}{p} > \frac{1}{\sqrt[m]{n}}$ because $n \geq p^m + 1$. □

Theorem 4.2. *If the diversity measure $L(\mathbf{P}(X,n),d)$ is (ρ,ε)-ectropic on a hypercube $X=[a,b]^m$, where $nd^m=(b-a)^m$, then $\rho+\varepsilon=1$.*

Proof. Since $L(\mathbf{P}(X,n),d)$ is (ρ,ε)-ectropic on a hypercube $X=[a,b]^m$, the following holds:

$$\sup_{|\mathbf{P}(X,n)|=\rho n} L(\mathbf{P}(X,n),d)=(1-\varepsilon)\sup_{\mathbf{P}(X,n)} L(\mathbf{P}(X,n),d).$$

According to Theorem 4.1, the expression on the right hand side equals $(1-\varepsilon)(b-a)^m$. Moreover, the expression on the left hand side is a supremum on the diversity measure $L(\mathbf{P}(X,\rho n),d)$, but without adjusting (increasing) d to the "new" population size ρn (see (11)). Clearly the value of this supremum is $\rho n\cdot d^m=\rho\cdot(b-a)^m$. Comparing these values we have that $\rho=1-\varepsilon$. □

This convenient linear relationship between ρ and ε implies the following. The maximum value of $L(\mathbf{P}(X,n),d)$ with the cardinality constraint $|\mathbf{P}(X,n)|=\rho n$, is linearly increasing with respect to ρ. Loosely speaking, this means that the maximum achievable diversity is proportional to the available resources within the population (in terms of cardinality). Measures that have this feature surely promote populations that cover the domain uniformly.

A drawback of L-diversity is its slight sensitivity to rotations due to the finite number of symmetry axes of the hypercube. Although this problem could be avoided by choosing hyperspheres instead, computing the volume of the union of hyperspheres would become practically intractable. Deriving an upper bound on the sensitivity of L-diversity to rotations remains an open problem, but we conjecture that this bound is rather tight. Indeed, in the computational study, we observed that L-diversity rarely changes by more than 1% when subjected to rotations. The maximum relative change could reach up to 8% only for some pathological populations with low cardinality. For some evidence regarding the low sensitivity of L-diversity to rotations, the reader is referred to [35].

4.2. Quasi-entropic diversity measure (E-diversity)

A natural question is whether it is possible to use some kind of entropy measure of the population and thus evaluate the diversity. This way, estimation of population diversity would be equivalent to answering the question "how much information does a population contain about the space in which it lies?". However, exact formulations, like Shannon entropy, do not fit easily when analyzing population diversity in $\mathbb{R}^m$, since we have fixed positions of the points, rather than probabilities of finding a certain point in a specific region. On the other hand, it is possible to consider the entropy as a quantity, which depends on the number of "occupied" fixed regions of X. There is a clear connection with statistical thermodynamics where the entropy is defined as a quantity proportional to the number of microscopic configurations that result in the observed macroscopic description of the thermodynamic system. A simple illustration is given for a $2D$ problem (Fig. 6). For the given 16 points and the rectangular search space divided into 16 equal rectangles, the diversity measure would be 9, since 9 cells are occupied. It is obvious that this approach does not reward the collocation of points. However, shortcomings of this approach are obvious. Two or more points can be very close to each other, but can still belong to different cells. Further, it is difficult to choose the number of cells when the population size is not the power of 2. The problem becomes more difficult when the dimension of X is very large. For example, if the dimension of X is q, the smallest number of cells (if every coordinate axis is divided in 2 segments) is 2^q. One way to avoid these difficulties is to allow for arbitrary cell locations. In other words, the objective is to compute the minimum number of "floating cells" necessary to completely cover the population. For this purpose, each cell is represented as a hypersphere with a certain radius, although hypercubes of a given side would also do. An example is illustrated in Fig. 6.

Definition 4.1. *Population $\mathbf{P}(X,n)$ belongs to a class $C(k,\delta)$, if there exists the set of k hyperspheres with the radius δ that cover the population.*

We propose E-diversity measure as the mapping $E:\mathbb{R}^{m\times n}\to\{1,2,\ldots,n\}$ such that

$$E(\mathbf{P}(X,n),\delta)=\begin{cases}1 & \text{if } \mathbf{P}(X,n)\in C(1,\delta)\\ k & \text{if } \mathbf{P}(X,n)\in C(k,\delta) \quad\wedge\quad \mathbf{P}(X,n)\notin C(k-1,\delta)\end{cases} \tag{13}$$

The problem of covering a given set of points with a minimal number of hyperspheres of the predefined radius r is reported as $\mathcal{NP}$-complete (recognition version of the problem) [74, 20, 27, 26]. One way to estimate E-diversity is

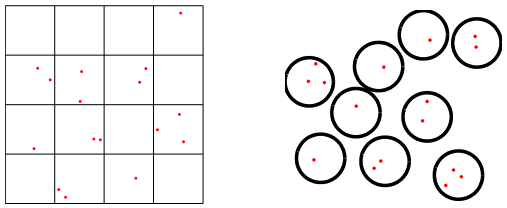

Figure 6: Diversity: a number of occupied cells vs. the number of cells needed for covering the population

to consecutively apply 2-approximation algorithm for the robust k-center problem (RKCP), [74] $(n-1)$ times (in the worst case). This algorithm allows to establish whether it is possible to cover a given set of n points with a set of k hyperspheres of predefined radius r, and its computational complexity is $O(n \log k)$ ([74, 20]). Thus, the overall worst-case computational complexity is $O\left(\sum_{k=1}^{n-1} n \log k\right) = O\left(n \log n!\right) = O\left(n^2 \log n\right)$.

There are two intuitive ways to choose the radius r of hyperspheres. One is to ensure that the volume of one hypershere is equal to the volume of the whole search space X divided by the number of points n. Thus, the radius r can easily be obtained knowing $\frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})} r^m = \frac{\mu_L(X)}{n}$. Left hand side of the last expression represents the volume of an m-dimensional hypersphere of radius r. Another way is to ensure that the diameter of the hypersphere equals the linear dimension of the hypercube whose volume is equal to the volume of the whole search space X divided by the number of points, i.e.:

$$n(2r)^m = \mu_L(X). \tag{14}$$

Of course, the value of r could also depend on the proximity of points that E-diversity measure may "tolerate". In the experimental part of this work, (14) is used.

Theorem 4.3. *If the diversity measure $E(\mathbf{P}(X,n), r)$ is (ρ, ε)-ectropic on a hypercube $X = [a,b]^m$, where $n(2r)^m = (a-b)^m$, then $\rho + \varepsilon = 1$.*

The proof is analogous to the one of Theorem 3.2.

In spite of this nice feature of the E-diversity measure, its obvious drawback is the discrete nature of its range, i.e. it may take only values $1, 2, \ldots, n$. An important shortcoming is also the lack of efficient algorithm for the exact calculation of E-diversity.

4.3. Power mean based diversity measure (H-diversity)

A commonly used diversity measure is the sum (or average) of all mutual distances between points. As shown in Section 3, the problem with this measure is inadequate treatment of the points' collocation at the vertices of X (if X is a polytope). For that purpose, we propose generalized mean-based measure (H-diversity) that can better deal with these difficulties. It is defined as the mapping $H : \mathbb{R}^{m \times n} \to \mathbb{R}$

$$H(\mathbf{P}(X,n), \alpha, \beta) = \sqrt[\beta]{\frac{1}{n} \sum_{i=1}^{n} d_i^{\alpha}}, \tag{15}$$

where $d_k^{\alpha} = \frac{1}{n-1} \sum_{i=1}^{n} \|\mathbf{x}_k - \mathbf{x}_i\|^{\alpha}$, with $k = 1, \ldots, n$ and $\alpha, \beta \neq 0$. If $\alpha = \beta = 1$, the function $H(\mathbf{P}(X,n), \alpha, \beta)$ becomes the average of all mutual distances (a scaled $D_d(\mathbf{P}(X,n))$ measure). Variable d_k represents the power mean of distances between point $\mathbf{x}_k$ and all other $n-1$ points. When α tends to zero, d_k tends to geometric mean of those distances that is in general, smaller than the arithmetic mean. When α further tends to $-\infty$, d_k tends to the minimum value of mentioned distances. One can draw similar conclusions about the behavior of $H(\mathbf{P}(X,n), \alpha, \beta)$ in terms of parameter β. It is natural to presume that smaller values of α and β would penalize the collocation of the points more heavily. A special case when $\alpha = -\infty$ and $\beta = 1$ yields the measure that is proposed in [2, 44]. The choice of proper values for α and β will be later elaborated in Section 5. The computational complexity of this measure is $O(mn^2)$. Although it seems very hard to perform a rigorous classification of this measure in terms of ectropic property, some conclusions will be drawn from experimental results based on how $H(\mathbf{P}(X,n), \alpha, \beta)$ correlates with other measures.

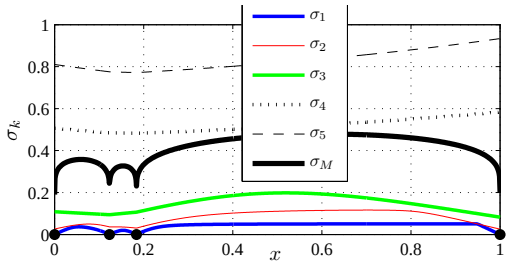

Figure 7: Profiles of the complete spectrum of σ_k measures

4.4. Singular values based diversity measure (σ-diversity)

An uncommon diversity measure has been presented in [37]. It is defined as the mapping $\sigma_k : \mathbb{R}^{m\times n} \to \mathbb{R}$:

$$\sigma_k(\mathbf{P}(X,n)) = \frac{1}{k}\sum_{j=1}^{k} s_{n-j+1}, \tag{16}$$

where $s_1 \geq s_2 \geq \ldots \geq s_n$ are the singular values of the distance matrix $\mathbf{D} = \left\{d_{ij}\right\}_{n\times n}$ whose entries are $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$. It can be shown that smaller k means higher penalty towards the populations with collocated points. In [40], it is proven that the number of nonzero singular values of the distance matrix is the number of distinct clusters that are extremely tight, meaning that they are consisted of collocated points. Here, a single isolated point is also considered a cluster. This result could also be expressed with the following theorem.

Theorem 4.4. *Diversity measure $\sigma_k(\mathbf{P}(X,n))$ is $(\rho,1)$-ectropic for all $\rho \leq 1 - \frac{k}{n}$.*

In other words, if $|\mathbf{P}(X,n)| = q < n$, then $\sigma_k = s_{n-k+1} = 0, \quad \forall k = 1,2,\ldots,n-q$. For example, if only two points within $\mathbf{P}(X,n)$ are collocated, then $\sigma_1 = 0$, no matter how the remaining points are arranged. If there are exactly two collocations, i.e. $|\mathbf{P}(X,n)| = n-2$, then $\sigma_1 = \sigma_2 = 0$, again invariant to all the arrangements allowed. Obvious drawback of $\sigma_k(\mathbf{P}(X,n))$ measure is that the choice of small k implies excessive rigor towards populations with some collocated points. On the other hand, it will be demonstrated that if k approaches n, then $\sigma_k(\mathbf{P}(X,n))$ highly correlates with $D_d(\mathbf{P}(X,n))$ measure that is shown to exert significant ectropic behavior. In [37] it is suggested that the values around $\frac{n}{2}$ are a convenient choice for k, ensuring high correlation with $L(\mathbf{P}(X,n),d)$ diversity. Finally, we propose one additional measure, based on the complete spectrum of distance matrix singular values:

$$\sigma_M(\mathbf{P}(X,n),\alpha) = \sqrt[\alpha]{\frac{1}{n}\sum_{i=1}^{n} s_i^{\alpha}}, \ \alpha \neq 0. \tag{17}$$

This measure is an attempt to avoid the above difficulties of $\sigma_k(\mathbf{P}(X,n))$ when k is either too small or too large. Parameter α should be smaller than 1 to avoid the domination of $\sigma_M(\mathbf{P}(X,n))$ over $\sigma_n(\mathbf{P}(X,n))$. Note that the measure σ_M represents the scaled Schatten norm of the distance matrix $\mathbf{D}$. Fig. 7 illustrates a $1D$ example for $n = 5$, where four points are fixed (dots on the abscissa) and all $\sigma_k(\mathbf{P}(X,n))$ measures are evaluated with the respect to the fifth point that takes all possible values from $[0,1]$. Clearly, σ_1 becomes zero whenever collocation occurs. Measures σ_2, σ_3 and σ_M with $\alpha = \frac{1}{5}$ also penalize the collocation, but not as strongly as σ_1. Functions σ_4 and σ_5 however increase when the fifth point approaches the edge of the interval, regardless of already existing points there. The diversity measures $D_d(\mathbf{P}(X,n))$ and $D_v(\mathbf{P}(X,n))$ have similar properties. To obtain σ_k measures we need two steps: to compute the distance matrix that can be done in $O(mn^2)$ time and to compute the singular values that requires $O(n^3)$ time [25]. Hence the measure can be obtained in $O(mn^2 + n^3)$ time.

4.5. Diversity measures based on discrepancy

In a number of papers, spanning the areas from numerical integration to sampling based motion planning, different kinds of discrepancies have been used to measure the quality of uniform distribution of a given point set [28, 38, 66,

18, 47, 19, 42, 43, 13, 24]. In Koksma-Hlawka inequality [28], a quadrature error bound is represented as the product of variation of the integrand and the measure of non-uniformity of the domain sample. The domain set X, in which the points are distributed, is assumed $[0,1]^m$. Let $\mathbf{P}(X,n)$ be the set of points in X. For any subset $B \subseteq X$, let

$$R(B,\mathbf{P}(X,n)) = \frac{A(B,\mathbf{P}(X,n))}{n} - \mu_L(B), \tag{18}$$

where $A(B,\mathbf{P}(X,n))$ is the function that counts the number of points from $(P(X,n))$ that fall inside B. The most common type of discrepancy is so-called *star discrepancy* $D^*_\infty(\mathbf{P}(X,n))$ (the argument $\mathbf{P}(X,n)$ omitted for brevity):

$$D^*_\infty = \sup_{\substack{z_k\in[0,1]\\ k=1,\ldots,m}} \left| R\left(\bigotimes_{i=1}^{m}[0,z_k)\mathbf{P}(X,n)\right)\right|. \tag{19}$$

Smaller values of $D^*_\infty(\mathbf{P}(X,n))$ should indicate that the points from $\mathbf{P}(X,n)$ are scattered more uniformly in $[0,1]^m$. Thus, the diversity measure might be formulated as a complement of discrepancy. Discrepancy $D^*_\infty(\mathbf{P}(X,n))$ is not invariant (in general) with respect to isometric transformations of the population $\mathbf{P}(X,n)$ since the set $\bigotimes_{i=1}^{m}[0,z_k)$ is "anchored" to the origin and coordinate axes. In order to mitigate this problem, so-called *extreme discrepancy* is defined:

$$D^{ex}_\infty = \sup_{\substack{u_k,v_k\in[0,1]\\ u_k\le v_k\\ k=1,\ldots,m}} \left| R\left(\bigotimes_{i=1}^{m}[u_k,v_k),\mathbf{P}(X,n)\right)\right|. \tag{20}$$

However, both star and extreme discrepancy are very hard to compute and hence have small practical significance [66]. A frequently used substitution is based on the replacement of the L_∞ norms in (19) and (20) with L_2 norms. In (21) and (22), *L_2-star discrepancy* and *L_2-extreme discrepancy* are given respectively:

$$D^*_2 = \sqrt{\int_{[0,1]^m} R^2\left(\bigotimes_{i=1}^{m}[0,z_k),\mathbf{P}(X,n)\right) d\mathbf{z}}, \tag{21}$$

where $d\mathbf{z} \equiv dz_1 dz_2 \ldots dz_m$.

$$D^{ex}_2 = \sqrt{\int_{[0,1]^m}\int_{[0,1]^m} R^2\left(\bigotimes_{i=1}^{m}[u_k,v_k),\mathbf{P}(X,n)\right) d\mathbf{u}d\mathbf{v}}, \tag{22}$$

where $d\mathbf{u} \equiv du_1 du_2 \ldots du_m$ and $d\mathbf{v} \equiv dv_1 dv_2 \ldots dv_m$.
Another interesting type of discrepancy is *wrap-around L_2-discrepancy* [28], [18], [19] defined as:

$$WD_2 = \sqrt{\sum_{t\neq\emptyset}\int_{C^t}\int_{C^t} R^2\left(J_w(\mathbf{u}_t,\mathbf{v}_t)\mathbf{P}(X,n)\right) d\mathbf{u}_t d\mathbf{v}_t}, \tag{23}$$

where t is a non-empty subset of the set of coordinate indices $M = \{1,2,\ldots,m\}$, C^t is $|t|$-dimensional unit cube involving the coordinates in t; $\mathbf{u}_t$ and $\mathbf{v}_t$ are projections of $\mathbf{u} = (u_1\ u_2\ \ldots\ u_m)$ and $\mathbf{v} = (v_1\ v_2\ \ldots\ v_m)$ on C^t respectively. Moreover, $J_w(\mathbf{u}_t,\mathbf{v}_t) = \bigotimes_{i=1}^{|t|} J_w(u_j,v_j)$ and

$$J_w(\alpha,\beta) = \begin{cases} [\alpha,\beta) & \text{if } \alpha\le\beta \\ [0,\beta)\cup[\alpha,1) & \text{if } \alpha>\beta. \end{cases}$$

$WD_2(\mathbf{P}(X,n))$ captures uniformity of points over both the unit cube X, and over all projections C^t. The advantage of L_2 based discrepancies is that there exist algebraic expressions for their exact calculation [28, 18, 47, 19].

$$(D^*_2)^2 = \frac{1}{3^m} - \frac{2}{n2^m}\sum_{j=1}^{n}\prod_{k=1}^{m}(1-x_{kj}^2) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\left[1-\max(x_{ki},x_{kj})\right] \tag{24}$$

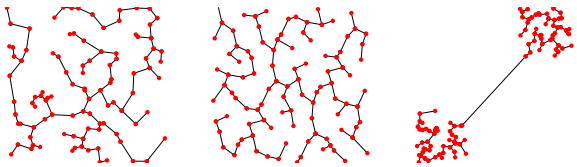

Figure 8: Diversity as the length of the EMST. Uniform random distribution of points (left), low-discrepancy sequence of points (middle), and two non-overlapping clusters (right).

$$\left(D_2^{ex}\right)^2 = \frac{1}{12^m} - \frac{2}{n2^m}\sum_{j=1}^{n}\prod_{k=1}^{m}(1-x_{kj})x_{kj} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\left[1-\max(x_{ki},x_{kj})\right]\min(x_{ki},x_{kj}) \tag{25}$$

$$(WD_2)^2 = -\left(\frac{4}{3}\right)^m + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\left[\frac{3}{2} - \left|x_{ki}-x_{kj}\right|\left(1-\left|x_{ki}-x_{kj}\right|\right)\right]. \tag{26}$$

If $D(\mathbf{P}(X,n))$ represents any of the discrepancies given, we define the corresponding diversity measure as:

$$\overline{D}(\mathbf{P}(X,n)) = UB(D(\mathbf{P}(X,n))) - D(\mathbf{P}(X,n)), \tag{27}$$

where $UB(D(\mathbf{P}(X,n)))$ represents any upper bound on discrepancy $D(\mathbf{P}(X,n))$. For instance, by observing (26), it clearly implies that the expression $\left(\frac{3}{2}\right)^m - \left(\frac{4}{3}\right)^m$ represents a conservative upper bound on $(WD_2)^2$. Likewise, the expressions $1 + \frac{1}{3^m}$ and $1 + \frac{1}{12^m}$ represent the upper bounds on $\left(D_2^*\right)^2$ and $\left(D_2^{ex}\right)^2$ respectively. The computational cost of the discrepancy based measures is $O(mn^2)$.

4.6. Diversity measure based on Euclidean minimum spanning tree

Let $G(V,C)$ be the complete undirected graph where the set of vertices V is the set of points of the population $\mathbf{P}(X,n) \in E(X,n) \subseteq \mathbb{R}^{m\times n}$, and the set of edges C is the set of all $\frac{n(n-1)}{2}$ pairwise connections between the points. The weight of each edge equals the corresponding Euclidean distance between the points. The problem of finding a minimum spanning tree (MST) in such a graph is known as the Euclidean minimum spanning tree (EMST) problem. We propose a diversity measure, described as:

$$D_{MST}(\mathbf{P}(X,n)) = \mu(MST(G(V,C))), \tag{28}$$

where $MST(G(V,C))$ denotes the MST subgraph of $G(V,C)$ and μ denotes the total length of the subgraph, i.e., the sum of the weights of all of its edges. Fig. 8 illustrates some populations in $\mathbb{R}^2$ and their corresponding EMSTs. The motivation behind this measure lies in the attempt to extract the "principal" connections (distances) out of the complete set of $\frac{n(n-1)}{2}$ ones. If the population consists of several distinct clusters, the diversity measure does not consider inter-cluster distances many times. As a matter of fact, multiple considerations of inter-cluster distances seem to be a core problem of the measures that reduce to the sum of all pairwise distances. A simple approach for computing $D_{MST}(\mathbf{P}(X,n))$ is to construct a complete graph $G(V,C)$ and then apply some of the classical algorithms for MST, e.g., Kruskal's or Prim's algorithm. This can be achieved in $O(n^2 \log n)$ and $O(n^2)$ time respectively [5]. An alternative is to use the method in [16] that has a worst-case complexity of $O(n^2\alpha(n^2,n))$, where $\alpha(\cdot,\cdot)$ is the classical functional inverse of Ackermann's function. Some algorithms are also available for the EMST problem: an $O(n \log n)$ algorithm for $\mathbb{R}^2$ [56], an $O((n \log n)^{1.8})$ algorithm for $\mathbb{R}^3$ can be found in [75], as well as an algorithm that runs in $O(n^{2-a(m)}(\log n)^{1-a(m)})$ for $\mathbb{R}^m$ problem, where $a(m) = 2^{-(m+1)}$.

4.7. Diversity measures of binary coded population

Consider a population $\mathbf{P}_b(X,n))$ of n binary strings (chromosomes) c_k, $(k = 1,2,\ldots,n)$ of length $m \cdot l$, where l is the number of bits used to encode each coordinate of a single point). As for $D_d(\mathbf{P}_b(X,n))$, the most frequently used diversity measure is [72, 53, 71]:

$$D_H(\mathbf{P}_b(X,n)) = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} HD(c_i,c_j), \tag{29}$$

where $HD(c_i, c_j)$ denotes the Hamming distance between two strings c_i and c_j. If the population is inherently binary, (29) is directly applicable. If the population $\mathbf{P}_b(X, n) \in E(X, n) \subseteq \mathbb{R}^{m \times n}$, then the points must be binary encoded before evaluating the diversity using (29). A possible encoding scheme is to map each component $x_{ij} \in [a_i, b_i]$ from the matrix $\mathbf{P}(X, n)$ into a binary string $\beta_l \beta_{l-1} \ldots \beta_1$, where β_l is the most significant bit. Here, X is assumed as a hyperrectangle $\bigotimes_{i=1}^{m} [a_i, b_i]$. Further conversion to Gray code is performed (in the experiments each real variable is encoded using 8 bits) in order to mitigate the effect of different significance for different bit locations.

Another popular diversity measure for populations consisting of binary/symbolic chromosomes is based on Shannon information entropy:

$$H(S) = \sum_{x \in A} p(x) \log \frac{1}{p(x)}, \tag{30}$$

where S is a discrete random variable with values taken from an alphabet A and a probability mass function $p(x) = \Pr(S = x),\ x \in A$. If we consider the above probabilities as the relative frequencies of the symbols from alphabet A at a specific locus, the diversity measure can be expressed as [72, 53, 71, 48]:

$$D_E(\mathbf{P}(X, n)) = \frac{1}{ml} \sum_{k=1}^{ml} \sum_{\alpha \in A} f_k(\alpha) log \frac{1}{f_k(\alpha)}, \tag{31}$$

where $f_k(\alpha)$ represents the number of times symbol α appears on the k-th location, divided by the population size n. For binary coded chromosomes $A = \{0, 1\}$. The computational complexity for both measures (29) and (31) can be reduced to $O(mln)$ [72]. An interesting diversity measure, based on the number of substrings, is proposed in [48]. For the population $\mathbf{P}_b(X, n)$ of n strings $c_k,\ (k = 1, 2, \ldots, n)$ its diversity is defined as:

$$D_{ss}(\mathbf{P}_b(X, n)) = n \frac{\left| S_{\{c_1, c_2, \ldots, c_n\}} \right|}{\sum_{j=1}^{n} \left| S_{\{c_i\}} \right|}, \tag{32}$$

where S_Y denotes the set of all substrings in the set Y. Clearly $S_{\{c_1, c_2, \ldots, c_n\}} = \bigcup_{j=1}^{n} S_{\{c_j\}}$. Computational complexity of diversity measure (32) is reducible to $O(mln)$, using so-called *suffix trees* [48].

5. Computational study

This section provides a computational analysis for the measures considered in the paper. As previously stated, the objective of the study is to establish the mutual dependencies among the considered measures. The first goal of the study is to support the theoretical analysis performed in Sections 3 and 4. On the other hand, the extensive empirical results provide additional insight into the behavior of those measures for which the rigorous analysis could not be conducted (e.g., for H-diversity or D_{MST} measure).

We first describe how we carried out the experiments and interpreted the data. Later on, we discuss the obtained results.

5.1. Description of the experiments

All the above measures are evaluated and tested on a large number of random and structured populations in $[0, 1]^m$. Separate experiments are performed for three different space dimensions $m = 2, 3, 4$, nine different population sizes $n = 9, 16, 25, 50, 100, 200, 500, 750, 1000$ and three distinctive "qualitative" types of population for a total of $3 \times 9 \times 3 = 81$ experiments. The first type of population is created using random uniform distribution. The second type of populations consist of Gaussian type clusters. The number of clusters is a random number (between 2 and $\lfloor \frac{n}{3} \rfloor$), the centers of clusters are random vectors, and the variance of each cluster is a random number. A cluster may consists of a single point. The third type of population is similar to the second one, with the variance of clusters set to zero. Thus the population is composed of groups of collocated points. Moreover, the number of clusters is a random number between 2 and $\lfloor \frac{n}{2} \rfloor$. Examples of the three types of populations with $n = 100$ points ($m = 2$) are shown in Fig. 9. The instances of the populations of type 1, 2, and 3 clearly cover the wide spectrum of conceivable point distributions in the domain set.

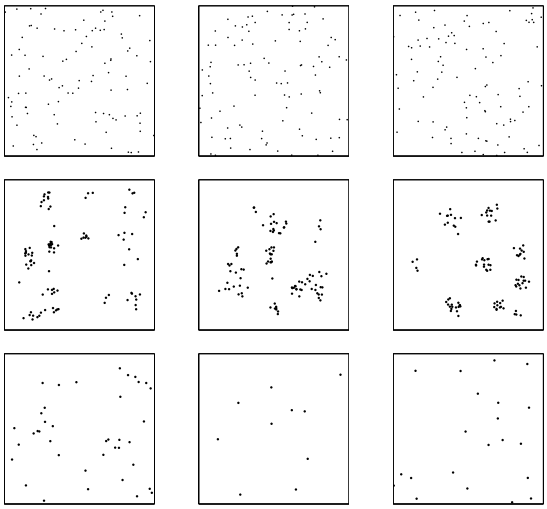

Figure 9: Population instances for $m = 2$, $n = 100$: uniform distribution (type 1 - top), Gaussian clusters (type 2) and tight clusters (type 3 - bottom)

We report results for the following measures: D_v, D_d, L, $H(-\infty, 1)$, $H(-\infty, m)$, $H(\frac{1}{n}, 1)$, $H(\frac{1}{n}, \frac{1}{n})$, E, all σ_k measures $k = 1, \ldots, n$, σ_M (with $\alpha = \frac{1}{n}$), $\overline{D}_2^*$, $\overline{D}_2^{ex}$, $\overline{WD}_2$, D_{MST}, D_H, D_E and D_{ss}. The other measures have also been considered but were discarded as candidates for complete test, due to the inconsistent behavior. We also considered the intuitive ideas of binning the set of all $\frac{n(n-1)}{2}$ pairwise distances or all the n singular values σ_k, with $k = 1, 2, \ldots, n$, and calculating entropy over them. In each one of the 81 experiments we have generated 1000 distinct populations and evaluated all the measures.

5.2. Comparison of the measures

In order to assess the mutual dependencies among different diversity measures, the Pearson's correlation coefficient is used:

$$CC(\mathbf{M}_1, \mathbf{M}_2, N) = \frac{\sum_{i=1}^{N} \left(\mathbf{M}_{1,i} - \overline{\mathbf{M}}_1\right)\left(\mathbf{M}_{2,i} - \overline{\mathbf{M}}_2\right)}{\sqrt{\sum_{i=1}^{N} \left(\mathbf{M}_{1,i} - \overline{\mathbf{M}}_1\right)^2 \sum_{i=1}^{N} \left(\mathbf{M}_{2,i} - \overline{\mathbf{M}}_2\right)^2}}, \tag{33}$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are N-dimensional vectors of measure values for generated N populations. Clearly, $|CC| \leq 1$ and equality holds if and only if the vectors are linearly dependent.

A natural question is what is the good choice of N. A large N means that many populations have to be generated, which leads to long lasting experiments. On the other hand, for N too small, the resulting CC may be "unreliable", meaning that it could take noticeably different value for another set of N populations. We have observed that the CC computed with a set of $N = 1000$ different populations is reliable enough, namely it does not fluctuate significantly for repeated experiments. This is supported by Fig. 10 that shows results for 30 repeated experiments for the case when $m = 2$, $n = 100$. Each repeated experiment consists in recomputing the CC for another set of $N = 1000$ randomly generated populations of "type 1". The CC (presented on the ordinate axis) is computed for L-diversity and D_{MST} measure for each $N \in \{2, 3, \ldots, 1000\}$ (abscissa). When N increases, the CC tends to a value close to 0.9, for each of the 30 experiments. Each experiment is represented by a grey line. The averaged curve is highlighted (black). This convergence-like behavior appeared to be even stronger for the populations of type 2 and 3.

The Pearson's correlation coefficient may not appear to be a suitable criterion for comparing the measures since it captures only the linear dependence. When two measures have the corresponding CC close to one it means that there is a strong linear correlation among them and hence there is not much ambiguity left. On the other hand, a low value of CC does not necessarily imply the absence of nonlinear dependence.

Although in general it may call for an alternative criterion to be used, the classical Pearson's correlation coefficient turns out to be good choice for pointing out the dependencies among the measures. Indeed, for the considerable number of pairs of measures that have relatively low CC, we observed that there is rarely any underlying nonlinear

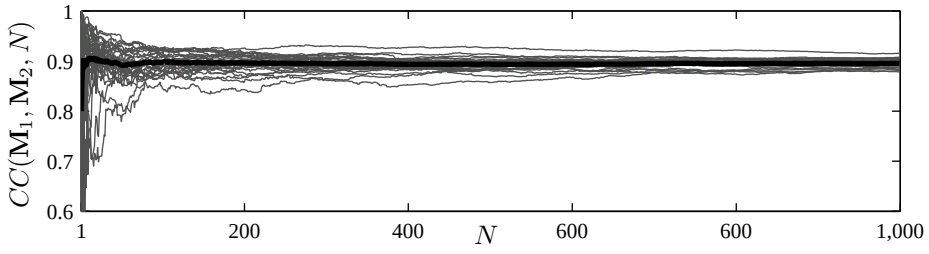

Figure 10: Correlation coefficients versus the number of experiments

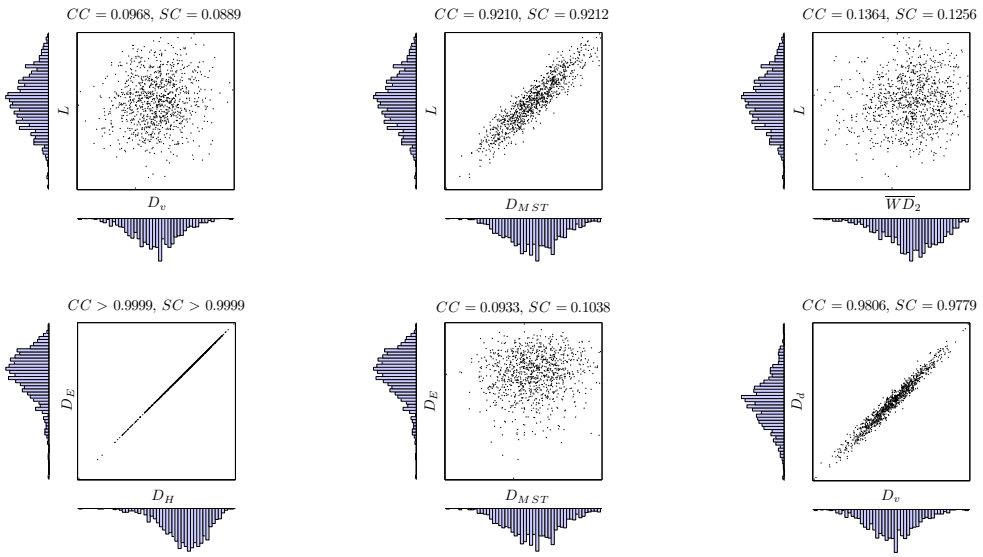

Figure 11: The scatter plots for some pairs of diversity measures (type 1 population, $n = 1000$, $m = 2$). When the CC for the pair of measures is low, the corresponding data set usually forms a cloud without the underlying functional dependence. Spearman's rank correlation coefficients are also indicated (SC).

dependence among them. The observation is based on both scatter plots for the measures of interest and the Spearman's rank correlation coefficient that can successfully capture any kind of monotonic functional dependence among the tested measures (see Fig.11 for some examples). On the other hand, addressing the assessment of the functional dependence between the measures solely by the Spearman's coefficient (or some other criterion like Kendall's coefficent or mutual information) does not seem a good choice because it cannot help make a clear distinction between the nonlinear and linear dependence. For instance, if the two measures have the corresponding Spearman's coefficient close to one, it implies that there is a strong monotonic dependence among them, but we cannot say anything about the nature of such dependence.

5.3. Numerical results

In this subsection, we present and discuss a representative selection of the numerical results. Each experiment is assigned a table (see Appendix), consisting of two parts. The first part (the leftmost four columns) contains the minimum, maximum, average and standard deviation values of all the measures in the experiment while the corresponding CCs are in the second part of the table. Table 2 shows the results for the case with $m = 2$, $n = 100$, and type 1 population. It contains the numerical data for all of the aforementioned diversity measures. The correlation data from the second part of Table 2 is based on the Pearson's coefficient. Table 3 however, reports the Spearman's coefficients for the same experiment. Since there is a negligible deviation from Pearson's coefficients reported in

Table 2, this further justifies the choice of Pearson's correlation coefficient CC as an indicator of dependency among diversity measures. For space constraints, the Tables 4-14 show the results (minimum, maximum, average, standard deviation, and Pearson's correlation coefficients) for the reduced set of measures: D_v, D_d, L, $H(-\infty, 1)$, E, selected σ_k measures, σ_M, $\overline{WD}_2$, D_{MST}, and D_E. The selection is mostly based on the consistency of CCs with other measures through experiments. Sometimes, a representative measure is picked from a set of related ones (e.g., $\overline{WD}_2$ out of the set of discrepancy based measures). For the complete set of numerical results, the reader is referred to [35].

In the sequel of this section, we first discuss the numerical results separately for the measures of interest, in the similar fashion as the measures are introduced in Sections 3 and 4. Special attention is devoted to how the measures correlate with L-diversity in order to identify a good candidate for substitution with lower computational complexity. To this end, the bar charts in Figs. 13-16 allow to visualize the selected data from the corresponding tables to emphasize the correlation of L-diversity with the measures of interest.

5.3.1. L-diversity

Not surprisingly, the numerical results obtained for L-diversity are quite coherent with the intuition behind it and its ectropic properties proven in Section 4. Observing the data in the left part of the Table 2 and Tables 4-14, it is clear that L-diversity decreases as the degree of collocation increases regardless of the population size n or the space dimension m. Indeed, L-diversity has the largest maximum, minimum, and average value for the populations of type 1. The corresponding values are smaller for the populations of type 2 and the smallest when the population consists of collocated points (type 3). This is an empirical evidence for the fact that the maximum value that the L-diversity can achieve for the population $\mathbf{P}(X, n)$ increases with the cardinality $|\mathbf{P}(X, n)|$ (see Theorem 4.2). Moreover, the numerical results for the populations of type 2 suggest that L-diversity also captures the emphasized proximity of points (Gaussian clusters), even if the corresponding populations have the full cardinality n. Therefore, the L-diversity provides a meaningful estimate of how uniformly the population $\mathbf{P}(X, n)$ fills the domain set X.

5.3.2. Cumulated distance-based measures (D_v and D_d)

As proven in Section 2, the measures D_v and D_d exhibit rather strong ectropic behavior since they can reach their maximum values when the population $\mathbf{P}(X, n)$ has a cardinality much smaller than n. It may appear that the specific populations used in proofs of Theorems 3.1 and 3.2 are too pathological. However, the numerical data from Tables 2-14 indicate that D_v and D_d do not penalize the collocation/proximity of points for a wide range of populations. For the experiments with $n = 100$ and $m = 2$ (Tables 2, 4, and 5), the average value of D_v does not substantially change with respect to the collocation degree, e.g, for the populations of type 1, it is 0.38, while the corresponding values for the populations of type 2 and 3 are 0.29 and 0.35 respectively. The maximum value of 0.57 is captured within the experiment for the populations of type 3 that supports the claims of Theorem 3.1. The same conclusions can be derived from the data corresponding to measure D_d. Furthermore, this kind of behavior does not seem to change for other values of n and m considered in the paper. The measures D_v and D_d are consistently poorly correlated with L-diversity (see Figs. 13-16). The corresponding CC rarely exceeds 0.5, while sometimes becomes smaller than 0.1. On the other hand their mutual CC is very high for all the instances (usually around 0.98 or 0.99 - see also Fig.11) and thus empirically supports the predictions from [72]. According to numerical results and their ectropic property, measures D_v and D_d turn out not to be the meaningful estimates of the population uniformity.

5.3.3. E-diversity

The E-diversity measure behaves in a similar way as the L-diversity in terms of numerical data from the left part of the Table 2 and Tables 4-14. For all the instances of n and m, the E-diversity decreases as the degree of collocation increases. Since this holds for the maximum, minimum, and the average value of the measure, it is a clear evidence that E-diversity captures the uniformity of points distribution to some extent. As expected, the correlation with the measures D_v and D_d is quite low for a wide range of instances. Somewhat surprisingly, the CC with L-diversity is not so consistent with respect to different experiments. In particular, it appears that the corresponding CC significantly increases when the degree of collocation increases (see Figs. 13-16). For the populations of type 3, the CC is nearly 1 regardless of the values of n and m considered in the paper (see Tables 5, 8, 11, and 14). For the populations of type 2, the CC takes values from 0.92 up to 0.97 (Tables 4, 7, 10, and 13), while for the populations of type 1 the values go from 0.53 to 0.71 (Tables 2, 6, 9, and 12). Notice that for the populations of type 1, the CC with L-diversity slightly decreases with the increase of the space dimensionality m, regardless of the population size n. Further investigation is

needed to establish whether this is the inherent property of the E-diversity defined by (13), or it is the consequence of the specific algorithm used to approximate the E-diversity.

5.3.4. Power mean based measures

The four measures based on power means were investigated: $H(-\infty, 1)$, $H(-\infty, m)$, $H(\frac{1}{n}, 1)$, and $H(\frac{1}{n}, \frac{1}{n})$. The numerical results in Table 2 as well as in [35] suggest that $H(\frac{1}{n}, 1)$ and $H(\frac{1}{n}, \frac{1}{n})$ significantly correlate with D_v and D_d, although this may appear counterintuitive because of the small values of α and β. Nevertheless, the same conclusions for D_v and D_d also apply for the measures $H(\frac{1}{n}, 1)$ and $H(\frac{1}{n}, \frac{1}{n})$. The measures $H(-\infty, 1)$, $H(-\infty, m)$ on the other hand have a very low CCs with D_v and D_d and a considerably high CC with L-diversity and E-diversity, particularly the measure $H(-\infty, 1)$. The value of CCs among $H(-\infty, 1)$ and L-diversity is usually above 0.7 and tends to increase with the increase of the collocation degree. For the populations of type 3, the CC is usually around 0.9, for all of n and m considered. Note that out of the four power mean based measures, only $H(-\infty, 1)$ is present in the Tables 4-14 and Figs. 13-16 (denoted H for brevity).

5.3.5. Singular values based measures

For all the experiments, a complete spectrum of σ_k measures, $k = 1, 2, \ldots, n$, is computed, as well as σ_M. We focus on the results for: σ_1, $\sigma_{\lfloor \frac{n}{2} \rfloor}$, σ_{best}, σ_n and σ_M. The measure σ_{best} stands for such σ_j, $j \in \{1, 2, \ldots, n\}$ that has the highest CC with L-diversity for a specific experiment. According to Table 2, the measure σ_n correlates very well with D_d (and hence D_v). This is not surprising since both σ_n and D_d represent the norms of the distance matrix $\mathbf{D}$. Since this is observed in all the experiments (see [35]), the conclusions about D_d also hold for σ_n and the results for σ_n are hence omitted in Tables 4-14 and Figs. 13-16. We also omit the results for σ_1 since it does not correlate with any of the measures of interest.

Fig. 12 shows the bar charts for the CC of L-diversity and all the σ_k measures, $k = 1, 2, \ldots, n$, for the case with $n = 100$, $m = 2$. Similar sets of charts are obtained for other values of n and m considered in the paper. For the population of type 1 (the top part), the values of k between 50 and 80 ensure a correlation above 0.9. When k exceeds 80 and approaches $n = 100$, the CC decreases rapidly. The highest CC is achieved for $\sigma_{best} = \sigma_{69}$. For the population of type 2 (the middle part of Fig.12), the CC with L-diversity becomes higher than 0.9 for the values of k between 77 and 94. This interval is much narrower than the one for type 1 populations. The measure $\sigma_{best} = \sigma_{90}$ has the highest CC with L-diversity. The bottom part of Fig.12 shows the bar chart for the population of type 3. The values of k between 82 and 96 ensure the CC with L-diversity higher than 0.9, while $\sigma_{best} = \sigma_{91}$. Note that the CCs for σ_k, $k = 1, 2, \ldots, 57$ and L-diversity are not defined since all of 57 σ_k vectors (1×1000) consist of all zeroes (see Theorem 4.4).

5.3.6. Discrepancy based measures

The discrepancy based measures $\overline{D}_2^*$, $\overline{D}_2^{ex}$ and $\overline{WD}_2$ poorly correlate with other measures (see e.g, Table 2). The corresponding CCs with L-diversity and E-diversity are low as well. This is somewhat surprising since the discrepancy is by definition an assessment of uniform distribution. The correlation becomes significant only for the populations of type 3 (see Tables 5, 8, 11, 14 and Figs. 13-16). This may be due to the fact that it is hard to extract any information about the spatial distribution of points from a specific value of discrepancy-based measures, although these measures have very intuitive foundation. For instance, an L-diversity value that is in the middle of its range ($[\frac{1}{n}, 1]$) suggests that approximately 50% of the domain is "populated", meaning that around 50% of the population "resources" are exploited to cover the domain set. Specific values of a discrepancy-based measure clearly provide much less information. Another reason may be that L_2 norms are used instead of L_∞ norms that appear in the original definitions of discrepancies. For brevity, Tables 4-14 and Figs. 13-16 contain the results only for the $\overline{WD}_2$ measure.

5.3.7. Minimum spanning tree based measure

Clearly, the Euclidean minimum spanning tree based measure D_{MST} in general decreases when the collocation degree increases (see the minimum, average, and the maximum values in Table 2 and Tables 4-14). Furthermore, it significantly correlates with other measures that promote uniformity. In particular the CC with L-diversity is consistently high and does not fluctuate by more than 10%. In most cases it is higher than 0.9 and it usually approaches 1 as the collocation degree increases. This is true for all the considered values of n and m (see Figs. 13-16). The

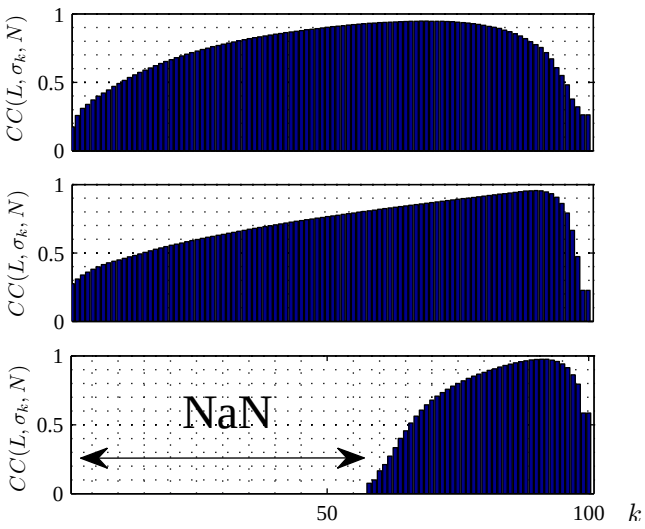

Figure 12: CCs of σ_k and L-diversity, $m = 2$, $n = 100$. Top - "type 1" population, middle - "type 2" population, bottom - "type 3" population

smallest CC with L-diversity is 0.84 in the experiment with $n = 25$, $m = 2$, and the population of type 2 (see [35]). Consistently large values and low sensitivity of the CC with respect to a wide range of instances qualify the D_{MST} measure as a reliable alternative for the L-diversity.

5.3.8. Measures for binary coded populations

None of the measures for binary coded populations has shown a consistent correlation with neither L-diversity nor any of the other measures for real-coded populations. For the populations of type 1, there seems to be hardly any correlation at all. Only the CCs of both D_H and D_E with $\overline{WD}_2$ stick out, but they are still below 0.5 (0.42). With the increase of the collocation degree, the correlation becomes slightly higher, particularly with $\overline{WD}_2$ measure (see e.g., Table 5). Note that D_H and D_E have very high CC (nearly 1 - see Table 2 and Fig. 11) that supports the analysis in [72]. This holds for all the experiments (see [35]). Tables 4-14 and Figs. 13-16 only report the results for D_E as the representative measure.

Although based on our experiments, none of D_H, D_E, or D_{ss} provides a meaningful information about the uniformity in the distribution, further research is needed to investigate whether a different encoding scheme may affect the results.

5.4. Summary and general discussion of results

In our experiments, the pairwise correlations of the tested measures do not vary substantially with respect to population size n nor the space dimension m. However, the type of population (1, 2 or 3) significantly affects the CCs for some pairs of measures. For instance, the CC of L-diversity and E-diversity reaches almost 1 in the experiments with heavily clustered populations. The D_E and $\overline{WD}_2$ measures also have a higher correlation with L-diversity for type 2 and 3 populations. On the other hand, measures based on singular values exhibit weaker correlation with L-measure for clustered populations. For type 3 populations, the measure $\sigma_{\frac{n}{2}}$ always returns a zero value that may provide a deceptive information. This does not occur for σ_{best} that is consistently well correlated with L-diversity. However, one should keep in mind that this measure is not defined a priori, but optimized a posteriori (with respect to k and CC with L-diversity) for each experiment. The measure D_{MST} exhibits both very consistent and very high correlation with L-diversity throughout all of the experiments. The value of the corresponding CC was mostly above 0.9 and often close to 1. Conversely, D_v and D_d correlate very poorly with L-diversity, the corresponding CCs rarely exceed 0.5. Their mutual correlation is very high for all instances (CC of 0.98 or 0.99). The measure $H(-\infty, 1)$ correlates significantly with L-diversity. For almost all of the experiments, the corresponding CC is higher then 0.7.

Table 1 provides insights about some inherent properties of the considered diversity measures. The second column indicates the worst-case computational complexity of the algorithms to compute the measures for general dimension

Table 1: General properties of tested diversity measures

Measure	Complexity $O(\cdot)$	Ect. behaviour
L	$n^{\lfloor \frac{m}{2} \rfloor}$	++
D_v	mn	- -
D_d	mn^2	- -
$H(-\infty, 1)$	mn^2	+
E	$n^2 \log n$	++
$\sigma_{\frac{n}{2}}$	$mn^2 + n^3$	+
σ_k^*	$mn^2 + n^3$	+
σ_M	$mn^2 + n^3$	+
$\overline{WD}_2$	mn^2	+
D_{MST}	$n^{2-a(m)}(\log n)^{1-a(m)}$	+
D_E	lmn	/

(to the best of our knowledge). In the ectropic behavior column, "++" indicates that the measures are proved not to be (ρ, ε)-ectropic where both ρ and ε can simultaneously take values close to zero. The label "- -" indicates the measures that behave just in the opposite way. Measures that are assigned the label "+" are likely not to exhibit a behavior similar to those with the label "- -", based on experimental results. Because of the low CCs with most of the other diversity measures, it is difficult to draw conclusions about the ectropic behavior of D_E. On the other hand, a variety of encoding possibilities prevent a clear classification.

Although extensive computational experiments for populations in spaces of dimension higher than $m = 4$ would be very welcome, the time needed to compute the L-diversity (which grows exponentially with m) prevented us to do so. Nevertheless, our preliminary results indicate that no substantial changes of pairwise CCs occur for other measures with the increase of m. For smaller values of n, this also appears to hold for L-diversity.

A question naturally arises from the above results and discussion: what is the best diversity measure? There is no clear general answer since it depends on the precise features we focus on. Do we wish to promote uniformity or do we prefer measures based on the aggregated distances between all pairs of points? What is the maximum complexity we can handle? If there is no concern about the ectropic behavior, i.e., populations consisting of several distant clusters can be considered as diverse, there is no reason for avoiding classical measures such as D_v, D_d and other related ones. However, if we want to "discourage" point collocation/proximity, we should use one of the measures that promote uniform distribution of points in the domain. This includes L-diversity, E-diversity, σ_k, $H(-\infty, 1)$ and D_{MST} measures. For larger dimensions, L-diversity becomes practically intractable because the time needed to compute it grows exponentially with m. In terms of CCs with L-diversity measure, D_{MST} and some σ_k-based measures exhibit the best performance. For σ_k diversity, the best choice for k is between $\frac{n}{2}$ and $\frac{9n}{10}$, and it usually ensures a CC with L-diversity above 0.9. However, if we deal with populations of type 2 or 3, the choice of k might strongly influence the corresponding CC (see Fig. 12). In any case, the best insight about the population is obtained when the complete spectrum of σ_k values ($k = 1, 2, \ldots, n$) is observed. This way, the exact number of possible collocations within the population can be estimated. Some analysis about this issue can be found in [37]. On the other hand, an $O(n^3)$ complexity may be too high for some large-scale applications. Finally, the D_{MST} measure appears to be the most interesting alternative to L-diversity. Its natural and intuitive definition, consistent CC with L-diversity and sub-quadratic complexity make it a very convenient measure for a wide range of applications.

6. Concluding remarks

The concept of diversity arises in many scientific and technical fields. Phrases like "promoting" or "preserving population diversity" are ubiquitous. But often "of the shelf" formulas for evaluating population diversity are used without precisely knowing what they really measure.

Since classical measures based on cumulated distances between all pairs of points do not properly account for point collocation/proximity, we have introduced the degree of ectropy of a diversity measure to estimate to what extent the point collocation is rewarded/penalized. To better account for collocation/proximity of points, we have

proposed and investigated several novel diversity measures based on quasi-entropy, discrepancy, power mean and Euclidean minimum spanning trees. Those based on Klee's problem (L-diversity) and on singular values have also been considered. Although L-diversity is the most attractive measure from the ectropy point of view, it has a high computational complexity. All measures have been compared using the correlation coefficients on populations of different types, sizes and dimensions. The measures that strive for large cumulative distances between the points are very poorly correlated with those that promote populations which cover the domain uniformly. Among the measures with moderate computational cost, the one based on Euclidean minimum spanning trees (D_{MST}) turns out to be the best alternative to L-diversity.

It is worth pointing out that, although the main focus of the paper is on Euclidean space, the concept of ectropy, the novel measures and related study can easily be extended to an arbitrary metric space, provided that the diversity measure is based on a corresponding distance function (metric).

There are several directions for future work. The first one is to investigate variants of the measures based on singular values and on general power means. Upper bounds on σ_k and D_{MST}, which are still an open issue, would allow to scale a priori the measures (for any population size or space dimension) and to classify them with respect to their ectropic property. Convex combinations of existing measures could also be considered. Another direction is to evaluate the potential impact of the various measures on the performance of diversity-guided algorithms and those that implicitly pursue the enhancement of diversity. Finally, from a theoretical point of view, it would be interesting to see whether variants or extensions of our concept of ectropy can help make further steps towards a unified framework for defining sound and suitable diversity measures.

Table 2: $m = 2$, $n = 100$, uniform distribution

min	max	avg	std	measure	L	D_v	D_d	$H(-\infty,1)$	$H(-\infty,m)$	$H(\frac{1}{n},1)$	$H(\frac{1}{n},\frac{1}{n})$	E	σ_1	$\sigma_{n/2}$	σ_{69}	σ_n	σ_M	$\overline{D}_2^*$	$\overline{D}_2^{ex}$	$\overline{WD}_2$	D_{MST}	D_H	D_E	D_{ss}
0.58	0.71	0.65	0.02	L	1.00	0.21	0.25	0.80	0.66	0.35	0.33	0.71	0.17	0.90	0.95	0.26	0.88	0.24	0.36	0.32	0.90	0.15	0.15	0.02
0.33	0.42	0.38	0.01	D_v	0.21	1.00	0.98	0.04	0.02	0.94	0.97	0.06	0.02	0.14	0.20	0.96	0.24	0.09	0.18	0.07	0.20	0.02	0.02	0.02
0.47	0.58	0.52	0.02	D_d	0.25	0.98	1.00	0.06	0.03	0.99	0.99	0.06	0.02	0.17	0.23	1.00	0.28	0.11	0.19	0.08	0.25	0.02	0.02	0.01
0.04	0.06	0.05	0.00	$H(-\infty,1)$	0.80	0.04	0.06	1.00	0.92	0.12	0.11	0.70	0.23	0.89	0.87	0.06	0.81	0.10	0.12	0.14	0.80	0.06	0.06	0.02
0.00	0.00	0.00	0.00	$H(-\infty,m)$	0.66	0.02	0.03	0.92	1.00	0.07	0.06	0.53	0.11	0.69	0.75	0.04	0.61	0.02	0.04	0.05	0.77	0.03	0.03	0.04
16.94	25.18	21.19	1.20	$H(\frac{1}{n},1)$	0.35	0.94	0.99	0.12	0.07	1.00	0.99	0.12	0.04	0.25	0.32	0.99	0.37	0.13	0.20	0.11	0.35	0.04	0.04	0.01
99.09	99.29	99.20	0.03	$H(\frac{1}{n},\frac{1}{n})$	0.33	0.97	0.99	0.11	0.06	0.99	1.00	0.12	0.04	0.24	0.30	0.99	0.35	0.17	0.28	0.16	0.31	0.06	0.06	0.01
57.00	82.00	70.38	3.83	E	0.71	0.06	0.06	0.70	0.53	0.12	0.12	1.00	0.14	0.78	0.72	0.06	0.67	0.17	0.21	0.22	0.60	0.09	0.09	-0.04
0.00	0.02	0.01	0.00	σ_1	0.17	0.02	0.02	0.23	0.11	0.04	0.04	0.14	1.00	0.29	0.21	0.02	0.44	0.01	0.05	0.06	0.14	0.02	0.02	0.11
0.03	0.04	0.04	0.00	$\sigma_{n/2}$	0.90	0.14	0.17	0.89	0.69	0.25	0.24	0.78	0.29	1.00	0.93	0.17	0.94	0.20	0.26	0.25	0.80	0.10	0.10	0.00
0.04	0.06	0.05	0.00	σ_{69}	0.95	0.20	0.23	0.87	0.75	0.32	0.30	0.72	0.21	0.93	1.00	0.24	0.92	0.21	0.27	0.27	0.90	0.13	0.13	0.02
0.95	1.16	1.06	0.03	σ_n	0.26	0.96	1.00	0.06	0.04	0.99	0.99	0.06	0.02	0.17	0.24	1.00	0.29	0.10	0.17	0.06	0.27	0.02	0.02	0.01
0.09	0.11	0.10	0.00	σ_M	0.88	0.24	0.28	0.81	0.61	0.37	0.35	0.67	0.44	0.94	0.92	0.29	1.00	0.21	0.28	0.27	0.81	0.13	0.13	0.03
0.96	1.04	1.02	0.01	$\overline{D}_2^*$	0.24	0.09	0.11	0.10	0.02	0.13	0.17	0.17	0.01	0.20	0.21	0.10	0.21	1.00	0.50	0.55	0.17	0.27	0.27	0.01
0.98	0.99	0.99	0.00	$\overline{D}_2^{ex}$	0.36	0.18	0.19	0.12	0.04	0.20	0.28	0.21	0.05	0.26	0.27	0.17	0.28	0.50	1.00	0.77	0.23	0.31	0.31	0.01
0.57	0.65	0.62	0.01	$\overline{WD}_2$	0.32	0.07	0.08	0.14	0.05	0.11	0.16	0.22	0.06	0.25	0.27	0.06	0.27	0.55	0.77	1.00	0.22	0.42	0.42	0.04
5.82	7.43	6.74	0.23	D_{MST}	0.90	0.20	0.25	0.80	0.77	0.35	0.31	0.60	0.14	0.80	0.90	0.27	0.81	0.17	0.23	0.22	1.00	0.10	0.10	0.05
39110	39889	39604	132	D_H	0.15	0.02	0.02	0.06	0.03	0.04	0.06	0.09	0.02	0.10	0.13	0.02	0.13	0.27	0.31	0.42	0.10	1.00	1.00	0.10
15.74	15.97	15.89	0.04	D_E	0.15	0.02	0.02	0.06	0.03	0.04	0.06	0.09	0.02	0.10	0.13	0.02	0.13	0.27	0.31	0.42	0.10	1.00	1.00	0.10
32.06	36.17	34.17	0.67	D_{ss}	0.02	0.02	0.01	0.02	0.04	0.01	0.01	-0.04	0.11	0.00	0.02	0.01	0.03	0.01	0.01	0.04	0.05	0.10	0.10	1.00

Table 3: $m = 2$, $n = 100$, uniform distribution, Spearman's rank correlation coefficients

measure	L	D_v	D_d	$H(-\infty,1)$	$H(-\infty,m)$	$H(\frac{1}{n},1)$	$H(\frac{1}{n},\frac{1}{n})$	E	σ_1	$\sigma_{n/2}$	σ_{69}	σ_n	σ_M	$\overline{D}_2^*$	$\overline{D}_2^{ex}$	$\overline{WD}_2$	D_{MST}	D_H	D_E	D_{ss}
L	1.00	0.20	0.24	0.78	0.64	0.34	0.32	0.70	0.17	0.89	0.94	0.25	0.88	0.23	0.35	0.31	0.89	0.15	0.15	0.03
D_v	0.20	1.00	0.98	0.04	0.02	0.94	0.97	0.04	0.02	0.13	0.19	0.96	0.22	0.08	0.14	0.07	0.19	0.02	0.02	0.01
D_d	0.24	0.98	1.00	0.05	0.03	0.98	0.99	0.05	0.02	0.16	0.22	0.99	0.26	0.10	0.15	0.08	0.24	0.03	0.03	0.01
$H(-\infty,1)$	0.78	0.04	0.05	1.00	0.91	0.12	0.10	0.69	0.23	0.89	0.85	0.06	0.80	0.08	0.12	0.12	0.78	0.05	0.05	0.01
$H(-\infty,m)$	0.64	0.02	0.03	0.91	1.00	0.08	0.06	0.52	0.12	0.68	0.73	0.04	0.60	0.01	0.04	0.05	0.75	0.03	0.03	0.04
$H(\frac{1}{n},1)$	0.34	0.94	0.98	0.12	0.08	1.00	0.99	0.11	0.04	0.24	0.31	0.99	0.36	0.13	0.18	0.11	0.33	0.05	0.05	0.01
$H(\frac{1}{n},\frac{1}{n})$	0.32	0.97	0.99	0.10	0.06	0.99	1.00	0.10	0.04	0.23	0.29	0.98	0.34	0.15	0.23	0.14	0.30	0.06	0.06	0.01
E	0.70	0.04	0.05	0.69	0.52	0.11	0.10	1.00	0.14	0.77	0.71	0.05	0.66	0.16	0.22	0.22	0.58	0.10	0.10	-0.04
σ_1	0.17	0.02	0.02	0.23	0.12	0.04	0.04	0.14	1.00	0.28	0.20	0.02	0.42	0.01	0.06	0.06	0.13	0.02	0.02	0.10
$\sigma_{n/2}$	0.89	0.13	0.16	0.89	0.68	0.24	0.23	0.77	0.28	1.00	0.92	0.16	0.93	0.18	0.25	0.23	0.79	0.11	0.11	0.01
σ_{69}	0.94	0.19	0.22	0.85	0.73	0.31	0.29	0.71	0.20	0.92	1.00	0.23	0.91	0.21	0.27	0.26	0.89	0.14	0.14	0.03
σ_n	0.25	0.96	0.99	0.06	0.04	0.99	0.98	0.05	0.02	0.16	0.23	1.00	0.27	0.09	0.13	0.06	0.25	0.02	0.02	0.01
σ_M	0.88	0.22	0.26	0.80	0.60	0.36	0.34	0.66	0.42	0.93	0.91	0.27	1.00	0.21	0.28	0.26	0.80	0.14	0.14	0.04
$\overline{D}_2^*$	0.23	0.08	0.10	0.08	0.01	0.13	0.15	0.16	0.01	0.18	0.21	0.09	0.21	1.00	0.54	0.57	0.17	0.29	0.29	0.00
$\overline{D}_2^{ex}$	0.35	0.14	0.15	0.12	0.04	0.18	0.23	0.22	0.06	0.25	0.27	0.13	0.28	0.54	1.00	0.75	0.24	0.32	0.32	0.01
$\overline{WD}_2$	0.31	0.07	0.08	0.12	0.05	0.11	0.14	0.22	0.06	0.23	0.26	0.06	0.26	0.57	0.75	1.00	0.21	0.41	0.41	0.03
D_{MST}	0.89	0.19	0.24	0.78	0.75	0.33	0.30	0.58	0.13	0.79	0.89	0.25	0.80	0.17	0.24	0.21	1.00	0.11	0.11	0.06
D_H	0.15	0.02	0.03	0.05	0.03	0.05	0.06	0.10	0.02	0.11	0.14	0.02	0.14	0.29	0.32	0.41	0.11	1.00	1.00	0.08
D_E	0.15	0.02	0.03	0.05	0.03	0.05	0.06	0.10	0.02	0.11	0.14	0.02	0.14	0.29	0.32	0.41	0.11	1.00	1.00	0.08
D_{ss}	0.03	0.01	0.01	0.01	0.04	0.01	0.01	-0.04	0.10	0.01	0.03	0.01	0.04	0.00	0.01	0.03	0.06	0.08	0.08	1.00

Table 4: $m = 2$, $n = 100$, 2 to 33 natural clusters with Gaussian distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{90}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.13	0.41	0.29	0.05	L	1.00	0.16	0.24	0.85	0.92	0.76	0.96	0.82	0.49	0.91	0.56
0.11	0.40	0.29	0.04	D_v	0.16	1.00	0.98	-0.07	0.08	-0.12	0.09	-0.05	0.80	0.30	0.49
0.16	0.53	0.38	0.05	D_d	0.24	0.98	1.00	-0.02	0.15	-0.07	0.16	0.01	0.85	0.38	0.56
0.01	0.03	0.02	0.00	H	0.85	-0.07	-0.02	1.00	0.82	0.83	0.83	0.83	0.21	0.79	0.36
10.00	41.00	25.33	5.72	E	0.92	0.08	0.15	0.82	1.00	0.71	0.86	0.74	0.37	0.82	0.45
0.00	0.02	0.01	0.00	$\sigma_{n/2}$	0.76	-0.12	-0.07	0.83	0.71	1.00	0.84	0.98	0.19	0.68	0.39
0.02	0.07	0.05	0.01	σ_{90}	0.96	0.09	0.16	0.83	0.86	0.84	1.00	0.89	0.42	0.90	0.55
0.01	0.06	0.03	0.01	σ_M	0.82	-0.05	0.01	0.83	0.74	0.98	0.89	1.00	0.27	0.75	0.45
0.21	0.58	0.45	0.06	$\overline{WD}_2$	0.49	0.80	0.85	0.21	0.37	0.19	0.42	0.27	1.00	0.56	0.78
1.72	4.38	3.35	0.43	D_{MST}	0.91	0.30	0.38	0.79	0.82	0.68	0.90	0.75	0.56	1.00	0.60
12.75	15.81	15.20	0.38	D_E	0.56	0.49	0.56	0.36	0.45	0.39	0.55	0.45	0.78	0.60	1.00

Table 5: $m = 2$, $n = 100$, 2 to 50 tight clusters

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{91}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.02	0.37	0.20	0.09	L	1.00	0.43	0.58	0.86	1.00	NaN	0.98	0.78	0.82	0.96	0.71
0.02	0.57	0.35	0.06	D_v	0.43	1.00	0.97	0.31	0.42	NaN	0.39	0.23	0.64	0.56	0.63
0.02	0.63	0.46	0.09	D_d	0.58	0.97	1.00	0.43	0.58	NaN	0.54	0.32	0.78	0.71	0.77
0.00	0.02	0.01	0.00	H	0.86	0.31	0.43	1.00	0.86	NaN	0.80	0.74	0.62	0.83	0.53
2.00	39.00	20.05	9.66	E	1.00	0.42	0.58	0.86	1.00	NaN	0.97	0.78	0.81	0.96	0.71
0.00	0.00	0.00	0.00	$\sigma_{n/2}$	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
0.00	0.09	0.04	0.03	σ_{91}	0.98	0.39	0.54	0.80	0.97	NaN	1.00	0.74	0.78	0.94	0.65
0.00	0.00	0.00	0.00	σ_M	0.78	0.23	0.32	0.74	0.78	NaN	0.74	1.00	0.50	0.69	0.38
0.03	0.61	0.47	0.10	$\overline{WD}_2$	0.82	0.64	0.78	0.62	0.81	NaN	0.78	0.50	1.00	0.86	0.90
0.09	4.74	3.04	1.02	D_{MST}	0.96	0.56	0.71	0.83	0.96	NaN	0.94	0.69	0.86	1.00	0.78
3.49	15.86	14.37	1.97	D_E	0.71	0.63	0.77	0.53	0.71	NaN	0.65	0.38	0.90	0.78	1.00

Table 6: $m = 2$, $n = 1000$, uniform distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{741}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.62	0.66	0.64	0.01	L	1.00	0.10	0.10	0.78	0.70	0.88	0.94	0.83	0.14	0.92	0.12
0.37	0.40	0.38	0.00	D_v	0.10	1.00	0.98	0.02	0.03	0.07	0.11	0.09	0.05	0.05	0.10
0.50	0.54	0.52	0.01	D_d	0.10	0.98	1.00	0.02	0.03	0.07	0.11	0.10	0.06	0.05	0.10
0.02	0.02	0.02	0.00	H	0.78	0.02	0.02	1.00	0.71	0.90	0.81	0.80	0.06	0.81	0.06
658.00	737.00	696.10	11.68	E	0.70	0.03	0.03	0.71	1.00	0.77	0.71	0.67	0.08	0.63	0.06
0.01	0.01	0.01	0.00	$\sigma_{n/2}$	0.88	0.07	0.07	0.90	0.77	1.00	0.90	0.94	0.10	0.82	0.08
0.02	0.02	0.02	0.00	σ_{741}	0.94	0.11	0.11	0.81	0.71	0.90	1.00	0.88	0.11	0.91	0.10
0.03	0.03	0.03	0.00	σ_M	0.83	0.09	0.10	0.80	0.67	0.94	0.88	1.00	0.10	0.77	0.07
0.65	0.67	0.67	0.00	$\overline{WD}_2$	0.14	0.05	0.06	0.06	0.08	0.10	0.11	0.10	1.00	0.07	0.45
20.15	21.30	20.80	0.19	D_{MST}	0.92	0.05	0.05	0.81	0.63	0.82	0.91	0.77	0.07	1.00	0.09
15.97	16.00	15.99	0.00	D_E	0.12	0.10	0.10	0.06	0.06	0.08	0.10	0.07	0.45	0.09	1.00

Table 7: $m = 2$, $n = 1000$, 2 to 333 natural clusters with Gaussian distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{912}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.30	0.46	0.40	0.03	L	1.00	0.22	0.25	0.95	0.97	0.94	0.98	0.93	0.37	0.98	0.43
0.26	0.35	0.30	0.01	D_v	0.22	1.00	0.98	0.14	0.14	0.08	0.18	0.10	0.93	0.18	0.60
0.36	0.48	0.42	0.02	D_d	0.25	0.98	1.00	0.16	0.17	0.10	0.21	0.13	0.96	0.22	0.67
0.01	0.01	0.01	0.00	H	0.95	0.14	0.16	1.00	0.94	0.93	0.94	0.92	0.27	0.96	0.34
317.00	502.00	427.07	35.12	E	0.97	0.14	0.17	0.94	1.00	0.94	0.95	0.92	0.28	0.95	0.36
0.00	0.01	0.01	0.00	$\sigma_{n/2}$	0.94	0.08	0.10	0.93	0.94	1.00	0.97	0.99	0.22	0.93	0.32
0.02	0.02	0.02	0.00	σ_{912}	0.98	0.18	0.21	0.94	0.95	0.97	1.00	0.97	0.33	0.98	0.40
0.01	0.02	0.02	0.00	σ_M	0.93	0.10	0.13	0.92	0.92	0.99	0.97	1.00	0.25	0.93	0.34
0.46	0.61	0.54	0.02	$\overline{WD}_2$	0.37	0.93	0.96	0.27	0.28	0.22	0.33	0.25	1.00	0.33	0.75
11.42	15.77	14.09	0.82	D_{MST}	0.98	0.18	0.22	0.96	0.95	0.93	0.98	0.93	0.33	1.00	0.40
15.49	15.92	15.76	0.07	D_E	0.43	0.60	0.67	0.34	0.36	0.32	0.40	0.34	0.75	0.40	1.00

Table 8: $m = 2$, $n = 1000$, 2 to 500 tight clusters

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{926}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.00	0.34	0.19	0.09	L	1.00	0.15	0.31	0.97	1.00	NaN	0.99	0.12	0.74	0.99	0.45
0.20	0.45	0.38	0.02	D_v	0.15	1.00	0.94	0.13	0.15	NaN	0.13	0.02	0.31	0.19	0.40
0.20	0.59	0.51	0.03	D_d	0.31	0.94	1.00	0.26	0.30	NaN	0.27	0.02	0.55	0.37	0.64
0.00	0.00	0.00	0.00	H	0.97	0.13	0.26	1.00	0.97	NaN	0.94	0.15	0.64	0.93	0.36
2.00	359.00	192.72	96.23	E	1.00	0.15	0.30	0.97	1.00	NaN	0.99	0.12	0.74	0.98	0.45
0.00	0.00	0.00	0.00	$\sigma_{n/2}$	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
0.00	0.03	0.01	0.01	σ_{926}	0.99	0.13	0.27	0.94	0.99	NaN	1.00	0.10	0.70	0.97	0.40
0.00	0.00	0.00	0.00	σ_M	0.12	0.02	0.02	0.15	0.12	NaN	0.10	1.00	0.05	0.10	0.02
0.20	0.66	0.61	0.05	$\overline{WD}_2$	0.74	0.31	0.55	0.64	0.74	NaN	0.70	0.05	1.00	0.82	0.82
0.39	13.57	9.31	2.92	D_{MST}	0.99	0.19	0.37	0.93	0.98	NaN	0.97	0.10	0.82	1.00	0.54
5.99	15.99	15.76	0.62	D_E	0.45	0.40	0.64	0.36	0.45	NaN	0.40	0.02	0.82	0.54	1.00

Table 9: $m = 4$, $n = 100$, uniform distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{69}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.68	0.80	0.75	0.02	L	1.00	0.62	0.65	0.76	0.57	0.91	0.94	0.94	0.28	0.91	0.11
0.50	0.60	0.56	0.01	D_v	0.62	1.00	0.99	0.24	0.17	0.46	0.55	0.66	0.05	0.45	0.02
0.71	0.83	0.78	0.02	D_d	0.65	0.99	1.00	0.27	0.17	0.49	0.59	0.70	0.05	0.48	0.03
0.19	0.24	0.21	0.01	H	0.76	0.24	0.27	1.00	0.69	0.91	0.86	0.79	0.20	0.89	0.08
77.00	97.00	88.54	3.09	E	0.57	0.17	0.17	0.69	1.00	0.73	0.60	0.60	0.21	0.57	0.08
0.13	0.16	0.15	0.00	$\sigma_{n/2}$	0.91	0.46	0.49	0.91	0.73	1.00	0.94	0.93	0.26	0.91	0.10
0.17	0.19	0.18	0.00	σ_{69}	0.94	0.55	0.59	0.86	0.60	0.94	1.00	0.96	0.24	0.94	0.10
0.28	0.33	0.31	0.01	σ_M	0.94	0.66	0.70	0.79	0.60	0.93	0.96	1.00	0.23	0.89	0.10
1.16	1.28	1.24	0.02	$\overline{WD}_2$	0.28	0.05	0.05	0.20	0.21	0.26	0.24	0.23	1.00	0.22	0.47
21.67	25.93	23.72	0.64	D_{MST}	0.91	0.45	0.48	0.89	0.57	0.91	0.94	0.89	0.22	1.00	0.08
31.55	31.89	31.77	0.06	D_E	0.11	0.02	0.03	0.08	0.08	0.10	0.10	0.10	0.47	0.08	1.00

Table 10: $m = 4$, $n = 100$, 2 to 33 natural clusters with Gaussian distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{91}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.14	0.34	0.25	0.05	L	1.00	0.24	0.37	0.75	0.93	0.65	0.96	0.77	0.65	0.96	0.68
0.28	0.54	0.42	0.05	D_v	0.24	1.00	0.98	-0.17	0.23	-0.21	0.16	-0.11	0.79	0.19	0.57
0.39	0.72	0.57	0.06	D_d	0.37	0.98	1.00	-0.08	0.34	-0.12	0.29	0.00	0.86	0.31	0.67
0.03	0.08	0.05	0.01	H	0.75	-0.17	-0.08	1.00	0.66	0.88	0.74	0.88	0.22	0.82	0.36
8.00	29.00	18.07	4.88	E	0.93	0.23	0.34	0.66	1.00	0.50	0.89	0.64	0.60	0.89	0.64
0.01	0.05	0.03	0.01	$\sigma_{n/2}$	0.65	-0.21	-0.12	0.88	0.50	1.00	0.69	0.97	0.18	0.73	0.32
0.05	0.17	0.11	0.03	σ_{91}	0.96	0.16	0.29	0.74	0.89	0.69	1.00	0.80	0.60	0.96	0.66
0.03	0.13	0.08	0.02	σ_M	0.77	-0.11	0.00	0.88	0.64	0.97	0.80	1.00	0.32	0.83	0.45
0.64	1.07	0.92	0.09	$\overline{WD}_2$	0.65	0.79	0.86	0.22	0.60	0.18	0.60	0.32	1.00	0.60	0.86
5.12	10.68	8.37	1.24	D_{MST}	0.96	0.19	0.31	0.82	0.89	0.73	0.96	0.83	0.60	1.00	0.66
27.53	31.51	30.49	0.67	D_E	0.68	0.57	0.67	0.36	0.64	0.32	0.66	0.45	0.86	0.66	1.00

Table 11: $m = 4$, $n = 100$, 2 to 50 tight clusters

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{91}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.02	0.39	0.20	0.10	L	1.00	0.49	0.64	0.89	1.00	NaN	0.98	0.73	0.85	0.99	0.73
0.14	0.70	0.52	0.07	D_v	0.49	1.00	0.97	0.36	0.48	NaN	0.46	0.23	0.71	0.55	0.72
0.15	0.88	0.69	0.11	D_d	0.64	0.97	1.00	0.48	0.63	NaN	0.62	0.32	0.84	0.70	0.85
0.00	0.06	0.02	0.01	H	0.89	0.36	0.48	1.00	0.90	NaN	0.83	0.76	0.67	0.88	0.55
2.00	43.00	21.42	10.46	E	1.00	0.48	0.63	0.90	1.00	NaN	0.97	0.74	0.84	0.98	0.72
0.00	0.00	0.00	0.00	$\sigma_{n/2}$	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
0.00	0.23	0.10	0.07	σ_{91}	0.98	0.46	0.62	0.83	0.97	NaN	1.00	0.67	0.83	0.97	0.69
0.00	0.00	0.00	0.00	σ_M	0.73	0.23	0.32	0.76	0.74	NaN	0.67	1.00	0.46	0.69	0.34
0.13	1.17	0.94	0.19	$\overline{WD}_2$	0.85	0.71	0.84	0.67	0.84	NaN	0.83	0.46	1.00	0.87	0.93
0.31	13.73	7.56	3.13	D_{MST}	0.99	0.55	0.70	0.88	0.98	NaN	0.97	0.69	0.87	1.00	0.76
7.80	31.57	28.89	3.56	D_E	0.73	0.72	0.85	0.55	0.72	NaN	0.69	0.34	0.93	0.76	1.00

Table 12: $m = 4$, $n = 1000$, uniform distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{804}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.68	0.72	0.70	0.01	L	1.00	0.53	0.55	0.73	0.53	0.88	0.95	0.93	0.15	0.87	0.08
0.55	0.57	0.56	0.00	D_v	0.53	1.00	0.99	0.14	0.12	0.37	0.50	0.54	0.00	0.24	0.00
0.76	0.79	0.78	0.01	D_d	0.55	0.99	1.00	0.15	0.13	0.38	0.53	0.56	0.00	0.26	0.00
0.11	0.12	0.11	0.00	H	0.73	0.14	0.15	1.00	0.73	0.91	0.76	0.79	0.11	0.90	0.08
840.00	912.00	875.47	10.41	E	0.53	0.12	0.13	0.73	1.00	0.74	0.55	0.62	0.11	0.62	0.08
0.07	0.08	0.08	0.00	$\sigma_{n/2}$	0.88	0.37	0.38	0.91	0.74	1.00	0.89	0.94	0.13	0.90	0.09
0.10	0.11	0.10	0.00	σ_{804}	0.95	0.50	0.53	0.76	0.55	0.89	1.00	0.96	0.12	0.88	0.07
0.14	0.15	0.15	0.00	σ_M	0.93	0.54	0.56	0.79	0.62	0.94	0.96	1.00	0.12	0.86	0.07
1.32	1.35	1.34	0.01	$\overline{WD}_2$	0.15	0.00	0.00	0.11	0.11	0.13	0.12	0.12	1.00	0.11	0.47
124.63	130.56	127.76	1.05	D_{MST}	0.87	0.24	0.26	0.90	0.62	0.90	0.88	0.86	0.11	1.00	0.07
31.95	31.99	31.98	0.01	D_E	0.08	0.00	0.00	0.08	0.08	0.09	0.07	0.07	0.47	0.07	1.00

Table 13: $m = 4$, $n = 1000$, 2 to 333 natural clusters with Gaussian distribution

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{960}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.19	0.31	0.26	0.03	L	1.00	0.11	0.16	0.87	0.96	0.79	0.98	0.85	0.29	0.95	0.26
0.39	0.46	0.42	0.01	D_v	0.11	1.00	0.99	-0.24	-0.04	-0.32	0.01	-0.24	0.95	-0.13	0.74
0.54	0.64	0.59	0.02	D_d	0.16	0.99	1.00	-0.20	0.01	-0.28	0.07	-0.20	0.97	-0.07	0.79
0.04	0.05	0.05	0.00	H	0.87	-0.24	-0.20	1.00	0.91	0.96	0.87	0.95	-0.06	0.97	-0.03
171.00	356.00	265.73	43.11	E	0.96	-0.04	0.01	0.91	1.00	0.83	0.94	0.85	0.14	0.96	0.12
0.01	0.03	0.03	0.00	$\sigma_{n/2}$	0.79	-0.32	-0.28	0.96	0.83	1.00	0.83	0.98	-0.15	0.92	-0.09
0.07	0.12	0.10	0.01	σ_{960}	0.98	0.01	0.07	0.87	0.94	0.83	1.00	0.88	0.20	0.96	0.19
0.04	0.07	0.06	0.01	σ_M	0.85	-0.24	-0.20	0.95	0.85	0.98	0.88	1.00	-0.06	0.94	-0.01
0.97	1.14	1.05	0.03	$\overline{WD}_2$	0.29	0.95	0.97	-0.06	0.14	-0.15	0.20	-0.06	1.00	0.06	0.81
44.01	67.24	57.56	4.96	D_{MST}	0.95	-0.13	-0.07	0.97	0.96	0.92	0.96	0.94	0.06	1.00	0.09
31.06	31.69	31.41	0.12	D_E	0.26	0.74	0.79	-0.03	0.12	-0.09	0.19	-0.01	0.81	0.09	1.00

Table 14: $m = 4$, $n = 1000$, 2 to 500 tight clusters

min	max	avg	std	measure	L	D_v	D_d	H	E	$\sigma_{n/2}$	σ_{920}	σ_M	$\overline{WD}_2$	D_{MST}	D_E
0.00	0.35	0.20	0.10	L	1.00	0.30	0.42	0.97	1.00	NaN	0.99	0.79	0.76	1.00	0.51
0.24	0.63	0.55	0.02	D_v	0.30	1.00	0.97	0.23	0.29	NaN	0.26	0.14	0.57	0.33	0.64
0.24	0.84	0.76	0.04	D_d	0.42	0.97	1.00	0.32	0.41	NaN	0.37	0.19	0.73	0.45	0.79
0.00	0.03	0.01	0.01	H	0.97	0.23	0.32	1.00	0.97	NaN	0.95	0.88	0.65	0.95	0.40
2.00	388.00	213.44	105.78	E	1.00	0.29	0.41	0.97	1.00	NaN	0.99	0.80	0.75	1.00	0.50
0.00	0.00	0.00	0.00	$\sigma_{n/2}$	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
0.00	0.10	0.05	0.03	σ_{920}	0.99	0.26	0.37	0.95	0.99	NaN	1.00	0.75	0.72	0.98	0.45
0.00	0.00	0.00	0.00	σ_M	0.79	0.14	0.19	0.88	0.80	NaN	0.75	1.00	0.44	0.77	0.25
0.25	1.31	1.23	0.10	$\overline{WD}_2$	0.76	0.57	0.73	0.65	0.75	NaN	0.72	0.44	1.00	0.80	0.89
0.68	65.85	41.03	16.68	D_{MST}	1.00	0.33	0.45	0.95	1.00	NaN	0.98	0.77	0.80	1.00	0.55
15.77	31.96	31.58	0.93	D_E	0.51	0.64	0.79	0.40	0.50	NaN	0.45	0.25	0.89	0.55	1.00

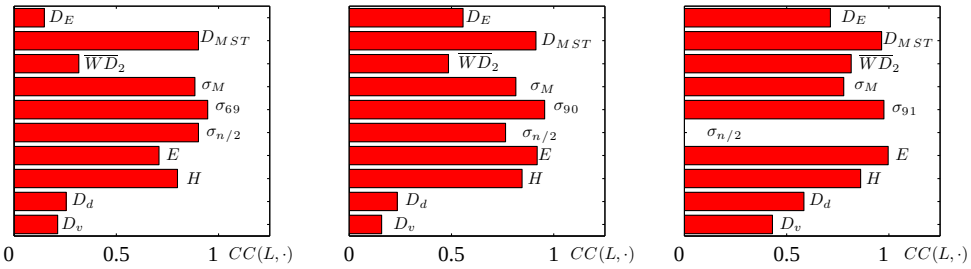

Figure 13: Bar charts of *CC*s of *L*-diversity and some of the measures - $m = 2$, $n = 100$. Left—uniform distribution, middle—Gausian clusters, right—tight clusters.

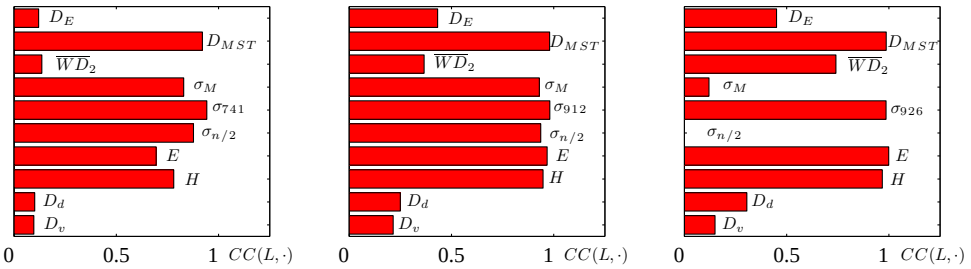

Figure 14: Bar charts of *CC*s of *L*-diversity and some of the measures - $m = 2$, $n = 1000$. Left—uniform distribution, middle—Gausian clusters, right—tight clusters.

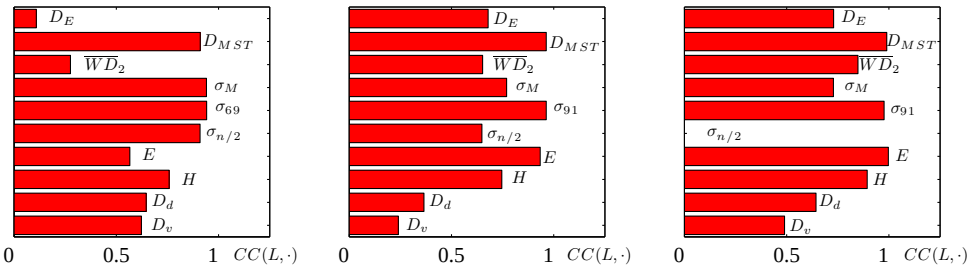

Figure 15: Bar charts of *CC*s of *L*-diversity and some of the measures - $m = 4$, $n = 100$. Left—uniform distribution, middle—Gausian clusters, right—tight clusters.

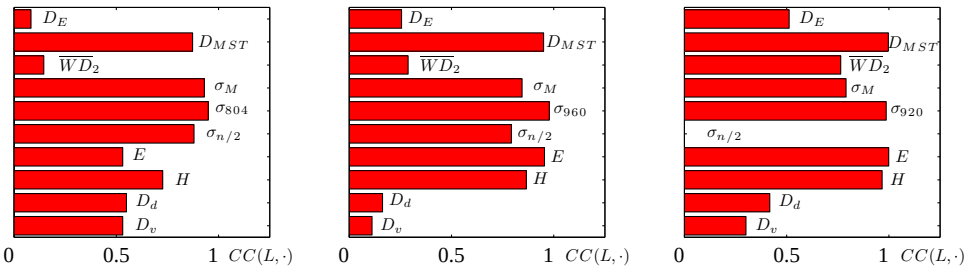

Figure 16: Bar charts of *CC*s of *L*-diversity and some of the measures - $m = 4$, $n = 1000$. Left—uniform distribution, middle—Gausian clusters, right—tight clusters.

References

[1] H.B. Amor, A. Rettinger, Intelligent exploration for genetic algorithms: using self-organizing maps in evolutionary computation, in: Proceedings of the Genetic and Evolutionary Computation Conference - GECCO, pp. 1531–1538.

[2] Z. Avdagic, S. Konjicija, B. Lacevic, Genetic algorithm with adaptive mutation probability in nonstationary environments, in: CD Proceedings of the International Convention MIPRO, Opatija, Croatia.

[3] A. Barker, W.N. Martin, Dynamics of a distance-based population diversity measure, in: Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, 2000, pp. 1002–1009.

[4] A.L. Barker, W.N. Martin, Population Diversity and Fitness Measures Based on Genomic Distances, Technical Report, Charlottesville, VA, USA, 1999.

[5] C.F. Bazlamaçci, K.S. Hindi, Minimum-weight spanning tree algorithms a survey and empirical study, Comput. Oper. Res. 28 (2001) 767–785.

[6] M. Bedau, M. Zwick, Variance and uncertainty measures of population diversity dynamics, Advances in Systems Science and Applications Special Issue I (1995) 7–12.

[7] J. Bentley, Algorithms for Klee's rectangle problem, Unpublished notes, Computer Science Department, Carnegie Mellon University, 1977.

[8] N. Beume, G. Rudolph, Faster s-metric calculation by considering dominated hypervolume as Klee's measure problem, in: Computational Intelligence, pp. 233–238.

[9] M. Blackwell, Particle swarms and population diversity, Soft Comput. 9 (2005) 793–802.

[10] J.D. Boissonnat, M. Sharir, B. Tagansky, M. Yvinec, Voronoi diagrams in higher dimensions under certain polyhedral distance functions, Discrete & Computational Geometry 19 (1998) 485–519.

[11] P. Bosman, D. Thierens, The balance between proximity and diversity in multiobjective evolutionary algorithms, Evolutionary Computation, IEEE Transactions on 7 (2003) 174 – 188.

[12] L. Bradstreet, R.L. While, L. Barone, Incrementally maximizing hypervolume for selection in multi-objective evolutionary algorithms, in: Proceedings of the IEEE Congress on Evolutionary Computation, pp. 3203–3210.

[13] M. Branicky, S. LaValle, K. Olson, L. Yang, Quasi-randomized path planning, in: Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on, volume 2, pp. 1481 – 1487.

[14] J. Buongiorno, S. Dahir, H.C. Lu, C.R. Lin, Tree size diversity and economic returns in uneven-aged forest stands, Forest Science 40 (1994) 83–103.

[15] E. Burke, S. Gustafson, G. Kendall, Diversity in genetic programming: an analysis of measures and correlation with fitness, Evolutionary Computation, IEEE Transactions on 8 (2004) 47 – 62.

[16] B. Chazelle, A minimum spanning tree algorithm with inverse-Ackermann type complexity, J. ACM 47 (2000) 1028–1047.

[17] J. Domingo-Ferrer, A. Solanas, A measure of variance for hierarchical nominal attributes, Inf. Sci. 178 (2008) 4644–4655.

[18] K.T. Fang, X. Lu, P. Winker, Lower bounds for centered and wrap-around L_2-discrepancies and construction of uniform designs by threshold accepting, J. Complex. 19 (2003) 692–711.

[19] K.T. Fang, C.X. Ma, Wrap-around L_2-discrepancy of random sampling, latin hypercube and uniform designs, J. Complex. 17 (2001) 608–624.

[20] T. Feder, D. Greene, Optimal algorithms for approximate clustering, in: STOC '88: Proceedings of the twentieth annual ACM symposium on Theory of computing, ACM, New York, NY, USA, 1988, pp. 434–444.

[21] C. Fernandes, A. Rosa, A study on non-random mating and varying population size in genetic algorithms using a royal road function, in: Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, 2001, pp. 60–66.

[22] M. Fleischer, M. Fleischer, The measure of pareto optima. applications to multi-objective metaheuristics, in: Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003, Springer, 2003, pp. 519–533.

[23] M. Gal-Or, J.H. May, W.E. Spangler, Assessing the predictive accuracy of diversity measures with domain-dependent, asymmetric misclassification costs, Information Fusion 6 (2005) 37–48.

[24] R. Geraerts, M.H. Overmars, A Comparative Study of Probabilistic Roadmap Planners, Technical Report UU-CS-2002-041, Department of Information and Computing Sciences, Utrecht University, 2002.

[25] G.H. Golub, C.F. Van Loan, Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition), The Johns Hopkins University Press, 3rd edition, 1996.

[26] T.F. Gonzalez, Clustering to minimize the maximum intercluster distance, Theor. Comput. Sci. 38 (1985) 293–306.

[27] T.F. Gonzalez, Covering a set of points in multidimensional space, Inf. Process. Lett. 40 (1991) 181–188.

[28] F.J. Hickernell, A generalized discrepancy and quadrature error bound, Math. Comput. 67 (1998) 299–322.

[29] S. Huband, P. Hingston, An evolution strategy with probabilistic mutation for multi-objective optimisation, in: Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, Piscataway NJ, 2003, pp. 2284–2291.

[30] J. Jie, J. Zeng, Particle swarm optimization with diversity-controlled acceleration coefficients, in: ICNC '07: Proceedings of the Third International Conference on Natural Computation, IEEE Computer Society, Washington, DC, USA, 2007, pp. 150–154.

[31] V. Klee, Can the measure of $\bigcup[a_i, b_i]$ be computed in less than $O(n \log n)$ steps?, American Mathematical Monthly 84 (1977) 284–285.

[32] J.D. Knowles, D.W. Corne, M. Fleischer, Bounded archiving using the Lebesgue measure, in: Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, 2003, pp. 2490–2497.

[33] P. Kouchakpour, A. Zaknich, T. Bräunl, Dynamic population variation in genetic programming, Inf. Sci. 179 (2009) 1078–1091.

[34] L.I. Kuncheva, C.J. Whitaker, Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy, Mach. Learn. 51 (2003) 181–207.

[35] B. Lacevic, E. Amaldi, On Population Diversity Measures, Technical Report, Politecnico di Milano, Dipartimento di Eletrronica e Informazione, http://home.dei.polimi.it/amaldi/ectropy-DEI-TReport-2009-17.pdf, 2009.

[36] B. Lacevic, E. Amaldi, On population diversity measures in Euclidean space, Evolutionary Computation (CEC), 2010 IEEE Congress on, pp. 1 –8.

[37] B. Lacevic, S. Konjicija, Z. Avdagic, Population diversity measure based on singular values of the distance matrix, in: Proceedings of the Congress on Evolutionary Computation, pp. 1863 –1869.
[38] G. Larcher, F. Pillichshammer, A note on optimal point distributions in [0,1)s, J. Comput. Appl. Math. 206 (2007) 977–985.
[39] M. Laumanns, E. Zitzler, L. Thiele, A unified model for multi-objective evolutionary algorithms with elitism, in: Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, 2000, pp. 46–53.
[40] S. Lee, Z.L. Zhang, S. Sahu, D. Saha, M. Srinivasan, Fundamental effects of clustering on the euclidean embedding of internet hosts, in: 6th International IFIP-TC6 Networking Conference, pp. 890–901.
[41] J. van Leeuwen, D. Wood, The measure problem for rectangular ranges in d-space, J. Algorithms 2 (1981) 282–300.
[42] Y.W. Leung, Y. Wang, Multiobjective programming using uniform design and genetic algorithm, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 30 (2000) 293 – 304.
[43] Y.W. Leung, Y. Wang, U-measure: a quality measure for multiobjective programming, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 33 (2003) 337 – 343.
[44] M. Lozano, F. Herrera, J.R. Cano, Replacement strategies to preserve useful diversity in steady-state genetic algorithms, Inf. Sci. 178 (2008) 4421–4433.
[45] A. Magurran, Ecological diversity and its measurement, Princeton University Press, 1988.
[46] R. Mallipeddi, P. Suganthan, B. Qu, Diversity enhanced adaptive evolutionary programming for solving single objective constrained problems, Evolutionary Computation, 2009. CEC '09. IEEE Congress on, pp. 2106 –2113.
[47] J. Matoušek, On the L_2-discrepancy for anchored boxes, J. Complex. 14 (1998) 527–556.
[48] C. Mattiussi, M. Waibel, D. Floreano, Measures of diversity for populations and distances between individuals with highly reorganizable genomes, Evol. Comput. 12 (2004) 495–515.
[49] M.M. Mayoral, Renyi's entropy as an index of diversity in simple-stage cluster sampling, Inf. Sci. 105 (1998) 101–114.
[50] R.W. Morrison, K.A.D. Jong, Measurement of population diversity, in: Selected Papers from the 5th European Conference on Artificial Evolution, Springer-Verlag, London, UK, 2002, pp. 31–41.
[51] R. Myers, E.R. Hancock, Genetic algorithms for ambiguous labelling problems, Pattern Recognition 33 (2000) 685–704.
[52] H. Onal, A computationally convenient diversity measure: Theory and application, Environmental & Resource Economics 9 (1997) 409–427.
[53] F. Oppacher, M. Wineberg, A linear time algorithm for determining population diversity in evolutionary computation, in: Proceedings of ISC - Intelligent Systems and Control.
[54] M.H. Overmars, C.K. Yap, New upper bounds in Klee's measure problem, SIAM J. Comput. 20 (1991) 1034–1045.
[55] R. Pasti, L.N. de Castro, Bio-inspired and gradient-based algorithms to train MLPs: The influence of diversity, Inf. Sci. 179 (2009) 1441–1453.
[56] F.P. Preparata, M.I. Shamos, Computational Geometry - An Introduction, Springer, 1985.
[57] B.Y. Qu, P.N. Suganthan, Multi-objective differential evolution with diversity enhancement, Journal of Zhejiang University-SCIENCE C (Computers & Electronics) 11 (2010) 538–543.
[58] B.Y. Qu, P.N. Suganthan, Multi-objective evolutionary algorithms based on the summation of normalized objectives and diversified selection, Inf. Sci. 180 (2010) 3170–3181.
[59] J. Riget, J. Vesterstrm, A Diversity-Guided Particle Swarm Optimizer - the ARPSO, Technical Report, Dept. of Computer Science, University of Aarhus, EVALife, 2002.
[60] B. Schölkopf, A.J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning), The MIT Press, 1st edition, 2001.
[61] Y. Shi, R. Eberhart, Population diversity of particle swarms, in: Proceedings of the IEEE Congress on Evolutionary Computation, pp. 1063–1067.
[62] H. Shimodaira, A diversity control oriented genetic algorithm (DCGA): Development and experimental results, in: Proceedings of the Genetic and Evolutionary Computation Conference - GECCO, pp. 603–611.
[63] A. Solow, S. Polasky, J. Broadus, On the measurement of biological diversity, Journal of Environmental Economics and Management 24 (1993) 60–68.
[64] E.K. Tang, P.N. Suganthan, X. Yao, An analysis of diversity measures, Mach. Learn. 65 (2006) 247–271.
[65] R.K. Ursem, Diversity-guided evolutionary algorithms, in: Proceedings of the Congress on Evolutionary Computation, IEEE Press, 2002, pp. 1633–1640.
[66] B. Vandewoestyne, R. Cools, Good permutations for deterministic scrambled Halton sequences in terms of L_2-discrepancy, J. Comput. Appl. Math. 189 (2006) 341–361.
[67] P. Varga, H.Y. Chen, K. Klinka, Tree-size diversity between single- and mixed-species stands in three forest types in western canada, Canadian Journal of Forest Research 35 (2005) 593–601.
[68] M. Ventresca, H.R. Tizhoosh, A diversity maintaining population-based incremental learning algorithm, Inf. Sci. 178 (2008) 4038–4056.
[69] M.L. Weitzman, On diversity, The Quarterly Journal of Economics 107 (1992) 363–405.
[70] L. While, P. Hingston, L. Barone, S. Huband, A faster algorithm for calculating hypervolume, Evolutionary Computation, IEEE Transactions on 10 (2006) 29 – 38.
[71] M. Wineberg, F. Oppacher, Distance between populations, in: Proceedings of the Genetic and Evolutionary Computation Conference - GECCO, pp. 1481–1492.
[72] M. Wineberg, F. Oppacher, The underlying similarity of diversity measures used in evolutionary computation, in: Proceedings of the Genetic and Evolutionary Computation Conference - GECCO, pp. 1493–1504.
[73] L. Xiangdong, L. Yuanchang, Information entropy measures for stand structural diversity: Joint entropy, Forestry Studies In China 6 (2004) 12–15.
[74] B. Xiao, J. Cao, Q. Zhuge, Y. He, E.H.M. Sha, Approximation algorithms design for disk partial covering problem, in: Parallel Architectures, Algorithms, and Networks, International Symposium on, pp. 104–109.
[75] A.C.C. Yao, On constructing minimum spanning trees in k-dimensional spaces and related problems, SIAM J. Comput. 11 (1982) 721–736.

[76] S. Zhao, P. Suganthan, Diversity enhanced particle swarm optimizer for global optimization of multimodal problems, Evolutionary Computation, 2009. CEC '09. IEEE Congress on, pp. 590 –597.
[77] K.Q. Zhu, Z. Liu, Population diversity in permutation-based genetic algorithm, in: European Conference on Machine Learning, pp. 537–547.
[78] E. Zitzler, Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications, Ph.D. thesis, ETH Zurich, Switzerland, 1999.
[79] H. Zouari, L. Heutte, Y. Lecourtier, Controlling the diversity in classifier ensembles through a measure of agreement., Pattern Recognition 38 (2005) 2195–2199.