Evolutionary Diversity Optimization Using Multi-Objective Indicators

Markus Wagner

Recent work with Carola Doerr, Wanru Gao,
Aneta Neumann, Frank Neumann
(original slides by Frank Neumann)

GECCO 2019: Best Paper Nominated

Motivation

- Diversity plays a crucial role in evolutionary computation, where we evolve sets (“populations”) of solutions

- Diversity
  – prevents premature convergence (”getting stuck early”)
  – enables successful recombination/crossover
  – allows to compute set of Pareto optimal solutions for multi-objective problems

Diversity

- Majority of approaches consider diversity in the objective space.
- Ulrich/Thiele considered diversity in the search space (Tamara Ulrich's PhD thesis, ~2011).
- Diversity with respect to other properties (features) is useful in various domains.
- Potential source for confusion: connections to subset selection problems, facility location problems (Operations Research), multi-modal optimisation, …

Goal:

- Compute a set of good solutions that differ in terms of interesting properties/features.
  – Think of (good) designs that vary with respect to important properties. → The objective space is not of immediate interest!

Application Areas

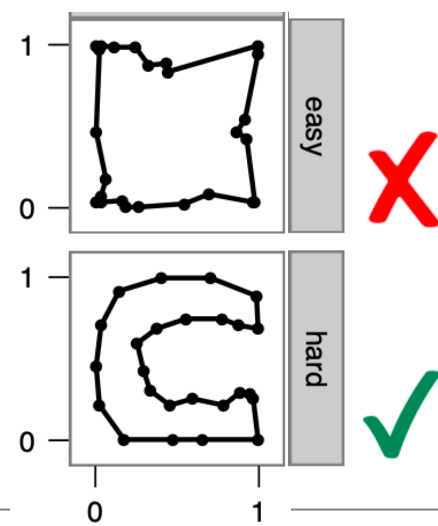
- Present a set of diverse high-quality solutions (instead of single one) to enable discussion for further refinement.
- See how good solutions distribute with respect to important features of solutions
- Understanding of algorithm performance with respect to important features through diverse problem instances
- Construction of diverse set of problem instances for algorithm selection.

Diversity of instances for TSP

- We want to construct a diverse set of TSP instances

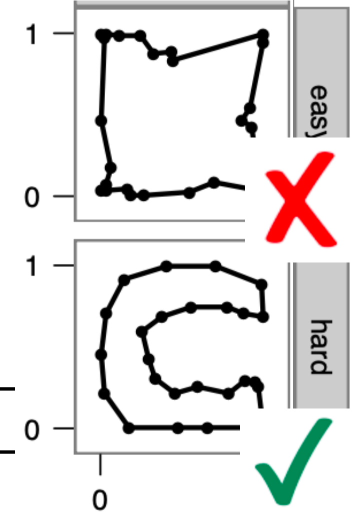
Examples:

- Diverse set where a certain algorithm is performing badly (high approximation ratio) $\alpha_A(I) = A(I)/OPT(I)$
- Diverse set where two solvers are performing differently.



EA for evolving diverse instances for the Traveling Salesperson Problem

(Gao, Nallaperuma, Neumann (PPSN 16))



Algorithm 1. $(\mu + \lambda)$-EA_D

---

1 Initialize the population P with μ TSP instances of approximation ratio at least α_h.
2 Let $C \subseteq P$ where $|C| = \lambda$.
3 For each $I \in C$, produce an offspring I' of I by mutation. If $\alpha_A(I') \geqslant \alpha_h$, add I' to P.
4 While $|P| > \mu$, remove an individual $I = \arg\min_{J \in P} d(J, P)$ uniformly at random.
5 Repeat step 2 to 4 until termination criterion is reached.

---

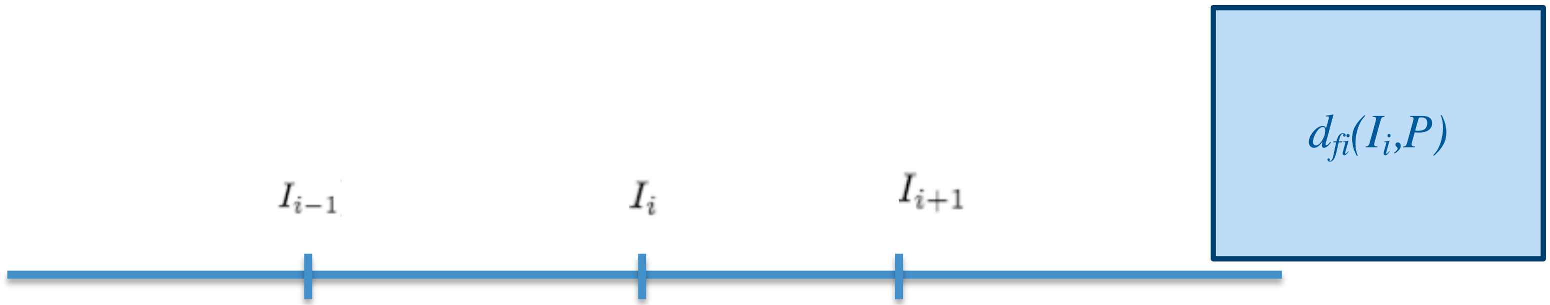
d(I,P) is the diversity contribution of instance I to the population P.
Let I be an individual (tour) and f(I) be its feature value.

We assume that $f(I_1) \leq f(I_2) \leq \ldots \leq f(I_k)$ holds.

Reminder: **α** is used here just as a quality constraint,
and survivor selection does not consider it, but only the diversity (contribution)

(Single-) feature diversity measure

Feature f:



$$f(I_1) \leq f(I_2) \leq \ldots \leq f(I_k),\ f(I_i) \neq f(I_1) \neq f(I_k)$$

"Diversity" of a single solution:

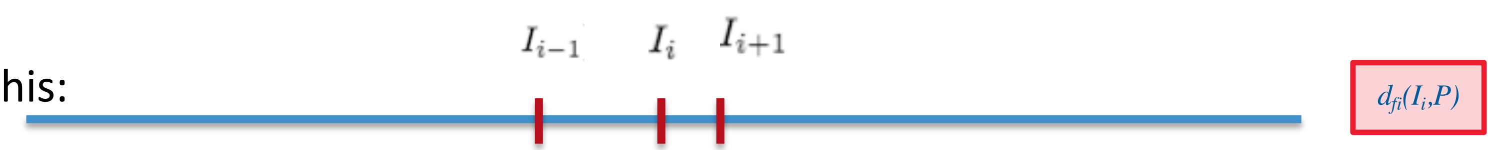
$$d_{f_i}(I_i, P) = (f(I_i) - f(I_{i-1})) \times (f(I_{i+1}) - f(I_i))$$

Diversity of a population:

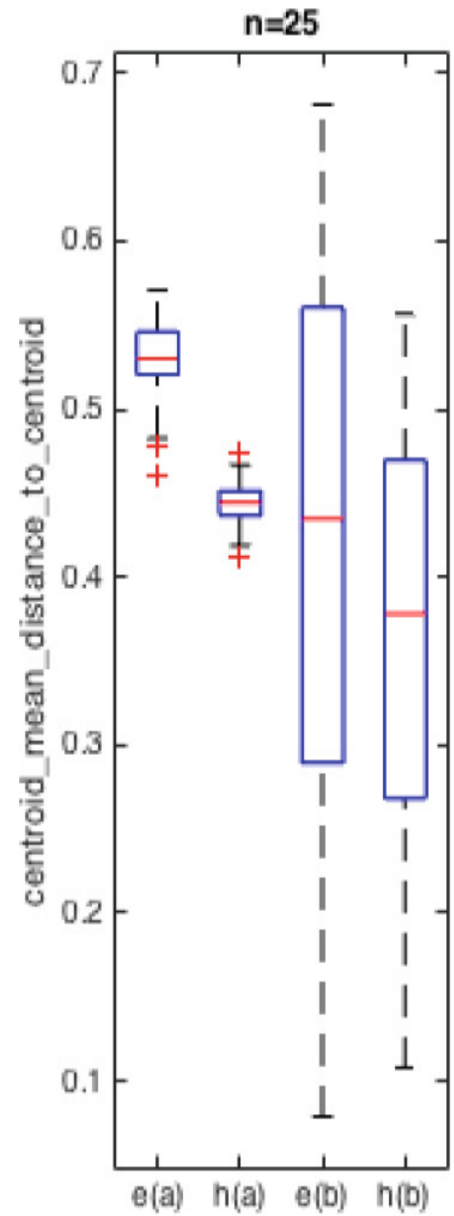
$$d'(I, P) = \sum_{i=1}^{k} (w_i \times d_{f_i}(I, P))$$

Maximum: if solutions are equally spaced out, as this is then the sum of squares

Compare with this:



Box plots features of “easy/hard” TSP instances for 2-opt (with and without diversity optimization)



Feature values of evolved instances:
From left to right:

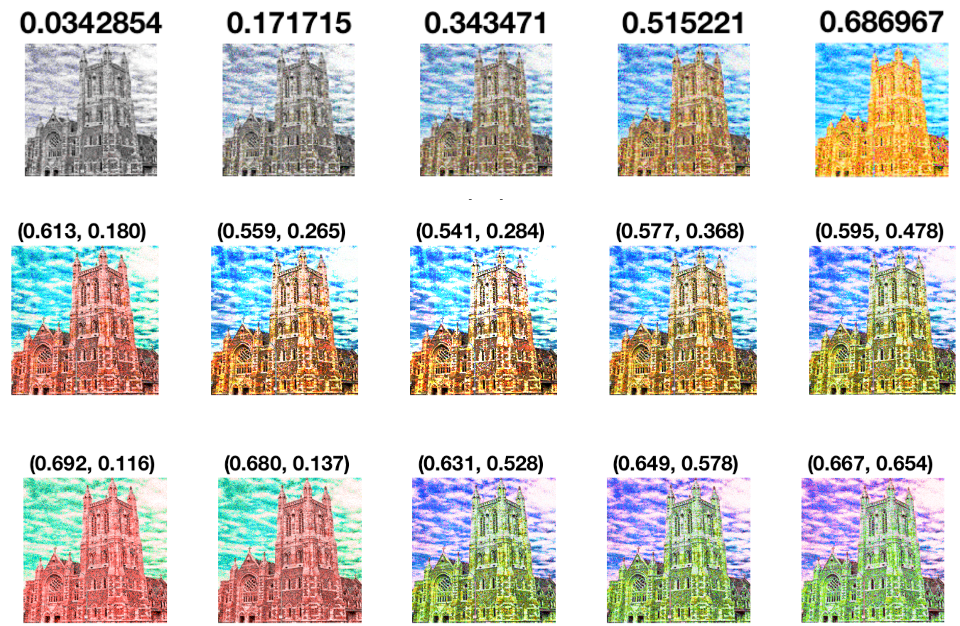
1. Easy instances / only using **α**
2. Hard instances / only using **α**
3. Easy instances / feature diversity (**α** as quality constraint)
4. Hard instances / feature diversity (**α** as quality constraint)

Works for other features, too… but not for all.

Diversity of Images

- Evolve a diverse set of images that are close to a given image.
- Close means:

RMSE to given image is less than 10. On the right: either 1 feature or a linear combination of two features as targets
[Alexander, Kortman, A. Neumann, GECCO’17]



Multiple features

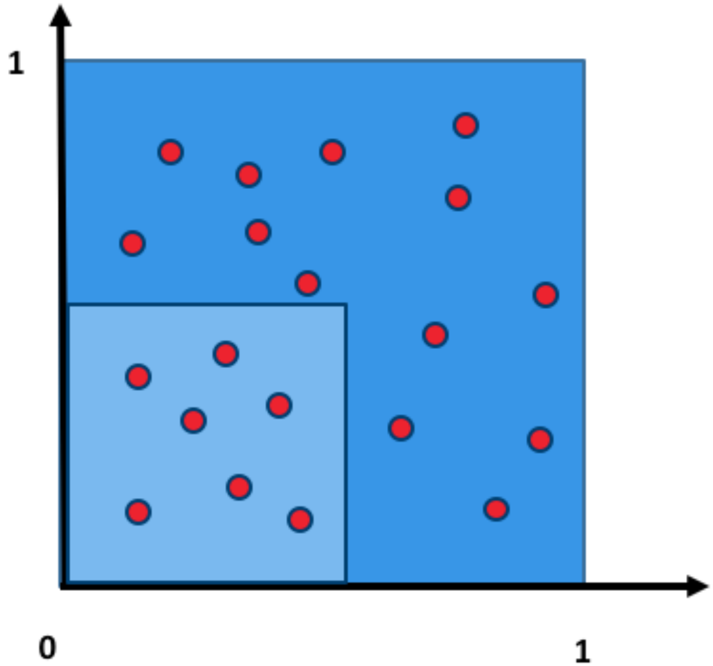
- For 2 or more features, weightening of diversity contributions might not lead to good diversity.
- Results depend on chosen weightening.

Questions:

- What is a good diversity measure?
- What is the diversity optimisation goal?

Discrepancy ("the number of points in a volume should be proportionate to the volume")

For further investigations we assume feature values are in [0,1]
(can be achieved through scaling)



Given a set of points $X := \{s^1, \ldots, s^n\}$
with $S = [0, 1]^d$, $s^1, \ldots, s^n \in S$

$$[a, b] := [a_1, b_1] \times \ldots \times [a_d, b_d]$$

$$\mathrm{Vol}([a, b]) - |X \cap [a, b]|/n$$

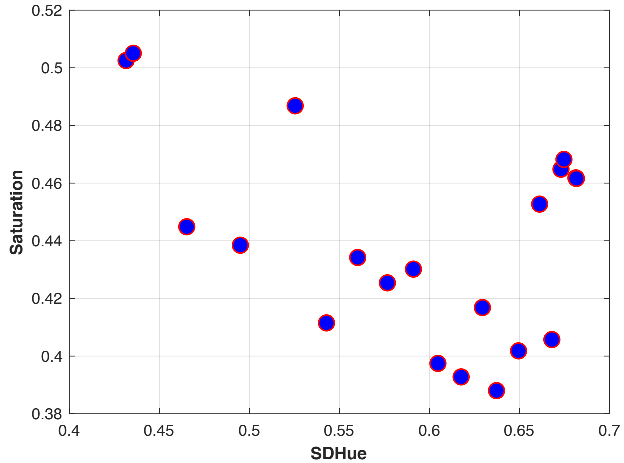
$$D(X, \mathcal{B}) := \sup\{\mathrm{Vol}([a, b]) - |X \cap [a, b]|/n \mid a \leq b \in [0, 1]^d\}$$

We consider special case of star discrepancy $a=0^d$

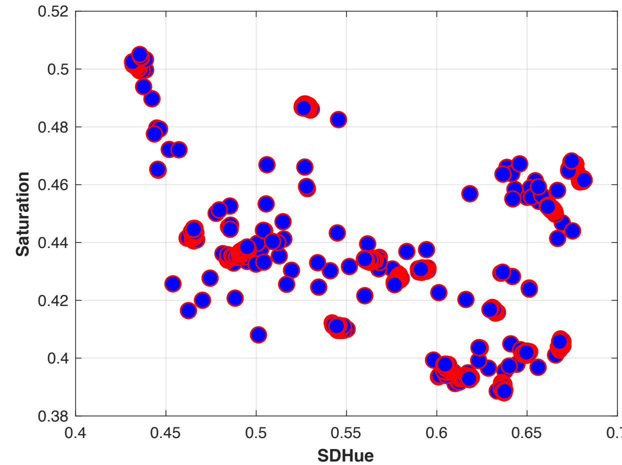
Example Runs Discrepancy

each dot is one solution (read: 1 image) in the 2D feature space

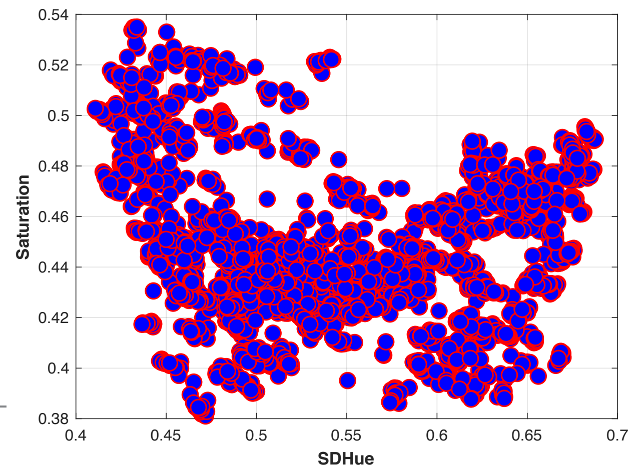
Final population of one run (1000 iterations)



All solutions during these 1000 iterations



All solutions during 10 independent runs



Images

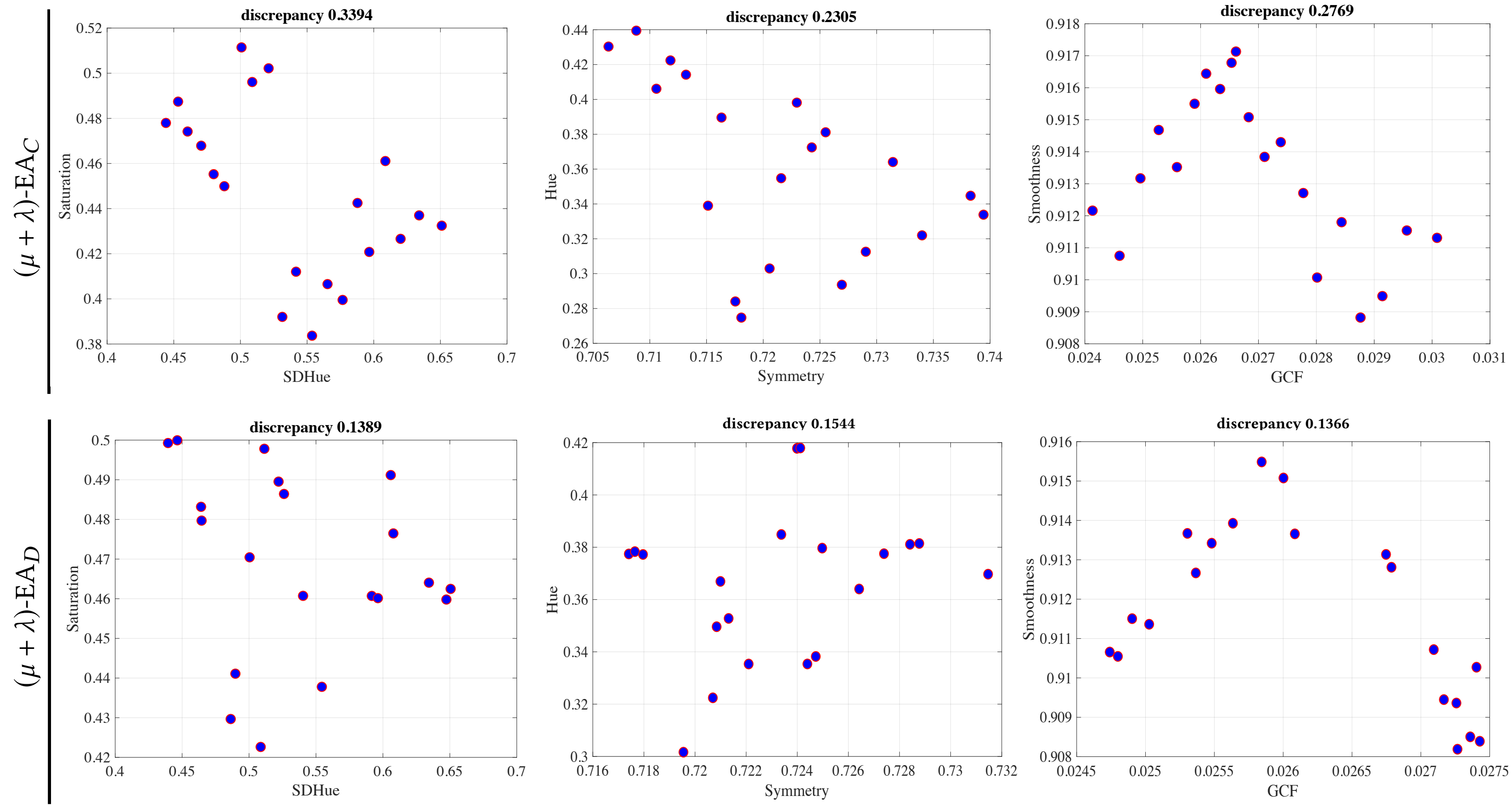


Figure 4: Feature vectors for final population of $(\mu + \lambda)$-EA$_C$ (top) and $(\mu + \lambda)$-EA$_D$ (bottom) for images based on two features from left to right: (SDHue, Saturation), (Symmetry, Hue), (GCF, Smoothness).

Images - Results

	$(\mu+\lambda)$-$EA_C(1)$				$(\mu+\lambda)$-EA_D (2)				$(\mu+\lambda)$-$EA_T(3)$			
	min	mean	std	stat	min	mean	std	stat	min	mean	std	stat
(f1, f2)	0.2014	0.3234	0.0595	$2^{(-)}$,$3^{(-)}$	0.1272	0.2038	0.1157	$1^{(+)}$	0.1119	0.1530	0.0269	$1^{(+)}$
(f3, f4)	0.1964	0.2945	0.0497	$2^{(-)}$,$3^{(-)}$	0.1574	0.2280	0.0592	$1^{(+)}$,$3^{(-)}$	0.1051	0.1417	0.0179	$1^{(+)}$,$2^{(+)}$
(f5, f6)	0.1997	0.2769	0.0344	$2^{(-)}$,$3^{(-)}$	0.1363	0.2025	0.0538	$1^{(+)}$	0.1457	0.1800	0.0234	$1^{(+)}$
(f1, f2, f3)	0.3389	0.4327	0.0613	$2^{(-)}$,$3^{(-)}$	0.1513	0.3335	0.1062	$1^{(+)}$	0.2253	0.2814	0.0422	$1^{(+)}$
(f1, f4, f3)	0.2754	0.3395	0.0483	$2^{(-)}$,$3^{(-)}$	0.2100	0.3118	0.1309	$1^{(+)}$	0.2224	0.2600	0.0123	$1^{(+)}$
(f5, f4, f2)	0.4775	0.6488	0.0841	$2^{(-)}$,$3^{(-)}$	0.2021	0.3007	0.1467	$1^{(+)}$	0.1983	0.2229	0.0125	$1^{(+)}$

Table 1: Statistics of discrepancy values for images. f1, f2, f3, f4, f5, f6 denote features SD-hue, Saturation, Symmetry, Hue, GCF and Smoothness, respectively.

TSP

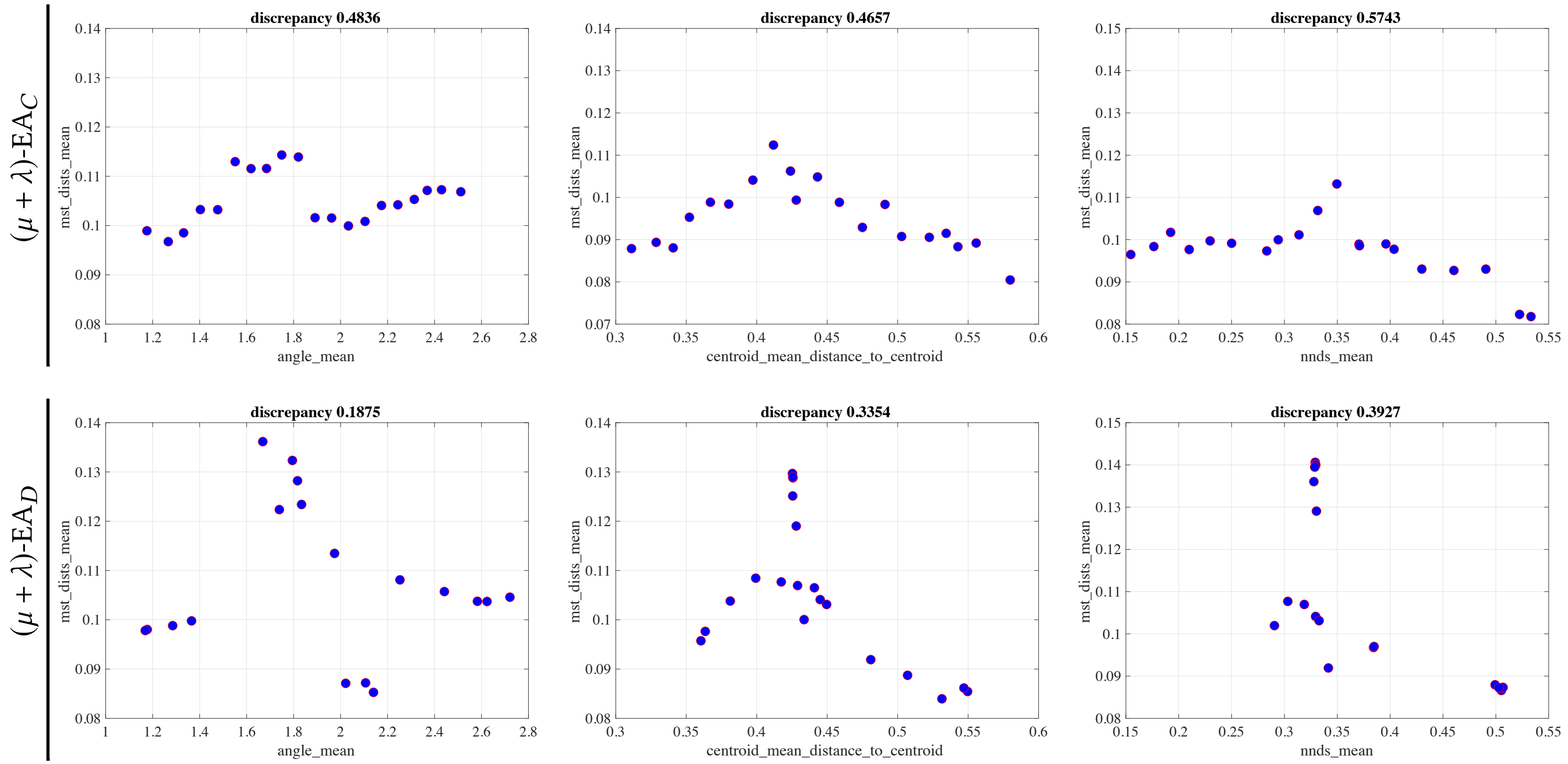


Figure 5: Feature vectors for final population of $(\mu + \lambda)$-EA$_C$ (top) and $(\mu + \lambda)$-EA$_D$ (bottom) for TSP based on two feature from left to right: (angle_mean, mst_dists_mean), (centroid_mean_distance_to_centroid, mst_dists_mean), (nnds_mean, mst_dists_mean)

Okay, so there is some diversity... but can we do better?

TSP - Results

	$(\mu+\lambda)$-EA_C				$(\mu+\lambda)$-EA_D				$(\mu+\lambda)$-EA_T			
	min	mean	std	stat	min	mean	std	stat	min	mean	std	stat
(f1,f4)	0.4836	0.5535	0.0362	$2^{(-)}$,$3^{(-)}$	0.2229	0.2942	0.0512	$1^{(+)}$,$3^{(-)}$	0.2013	0.2354	0.0252	$1^{(+)}$,$2^{(+)}$
(f2, f4)	0.4657	0.5192	0.0256	$2^{(-)}$,$3^{(-)}$	0.3229	0.3708	0.0414	$1^{(+)}$,$3^{(-)}$	0.2816	0.3363	0.0435	$1^{(+)}$,$2^{(+)}$
(f3, f4)	0.5743	0.6296	0.0219	$2^{(-)}$,$3^{(-)}$	0.3590	0.4422	0.0534	$1^{(+)}$	0.3831	0.4113	0.0175	$1^{(+)}$
(f1, f3, f4)	0.7765	0.7997	0.0204	$2^{(-)}$,$3^{(-)}$	0.4303	0.4585	0.0183	$1^{(+)}$	0.4372	0.4604	0.0422	$1^{(+)}$
(f2, f3, f4)	0.7641	0.7962	0.0198	$2^{(-)}$,$3^{(-)}$	0.4197	0.4563	0.0215	$1^{(+)}$	0.3730	0.4514	0.0327	$1^{(+)}$
(f1, f2, f3)	0.7593	0.7836	0.0111	$2^{(-)}$,$3^{(-)}$	0.3900	0.4095	0.0160	$1^{(+)}$	0.3547	0.3988	0.0217	$1^{(+)}$

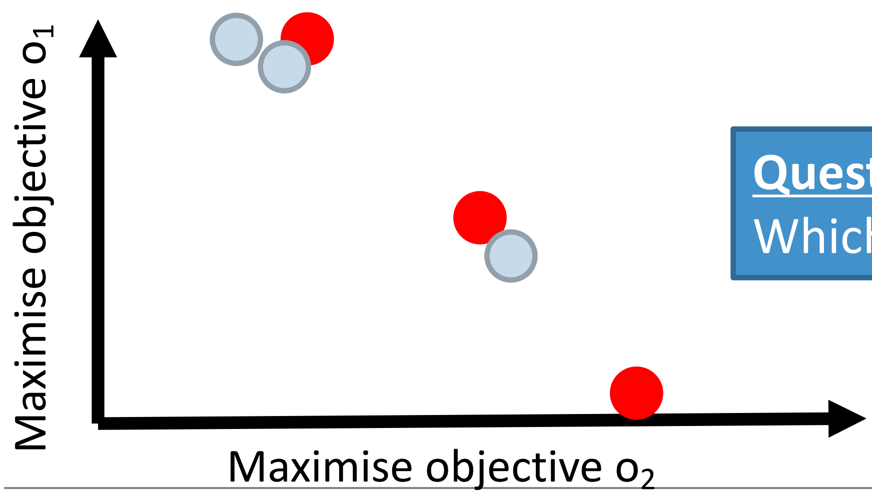
Table 2: Statistics of discrepancy values for TSP. f1, f2, f3, f4 denote the feature angle_mean, centroid_mean_dist_centroid, nnds_mean, mst_dists_mean respectively.

Evolutionary diversity optimization using multi-objective indicators

(A. Neumann, W. Gao, M. Wagner, F. Neumann, GECCO 2019)

Indicator-based Multi-Objective Optimization

- Let I be a search point
  - $f: X \rightarrow R^d$ a function that assigns to each search point I an objective vector
  - $q: X \rightarrow R^e$ be a function measures constraint violations
- An indicator Ind: $2^X \rightarrow R$ measures the quality of a given set of search points



Question
Which algorithm performed best?

Indicator-Based Diversity Optimisation

- Let I be a search point
  - $f: X \rightarrow R^d$ a function that assigns to each search point a feature vector
  - $q: X \rightarrow R$ be a function assigning a quality score to each $I \in X$ e.g.: require $q(I) \geq \alpha$ for all "good" solutions (constraint)
- Define Ind: $2^X \rightarrow R$ which measures the diversity of a given set of search points.

Goal:

Compute set $P=\{I_1, \ldots, I_\mu\}$ of μ solutions maximizing (minimizing) Ind among all sets of μ solutions under the condition that $q(I) \geq \alpha$ holds for all $I \in P$, where α is a given quality threshold.

Multi-Objective Indicators

Popular indicators in multi-objective optimization:

- Hypervolume (HYP)

$$HYP(S, r) = VOL\left(\cup_{(s_1,\ldots,s_d)\in S}\left[r_1, s_1\right] \times \cdots \left[r_d, s_d\right]\right)$$

- Inverted generational distance (IGD) (with respect to reference set R)

$$IGD(R, S) = \frac{1}{|R|}\sum_{r\in R}\min_{s\in S} d(r, s),$$

- Additive epsilon approximation (EPS) (with respect to reference set R)

$$\alpha(R, S) := \max_{r\in R}\min_{s\in S}\max_{1\leq i\leq d}(s_i - r_i).$$

How to use Multi-Objective Indicators

- Diversity Optimisation aims to compute a diverse set of solutions for a given single-objective problem
- Multi-Objective indicators guide the search towards a diverse set of Pareto optimal solutions.
  (read: some vectors are better than others, but we do not want this in diversity optimisation)

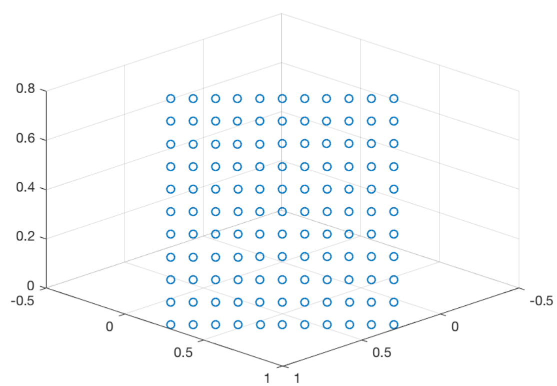
Use of multi-objective indicators:

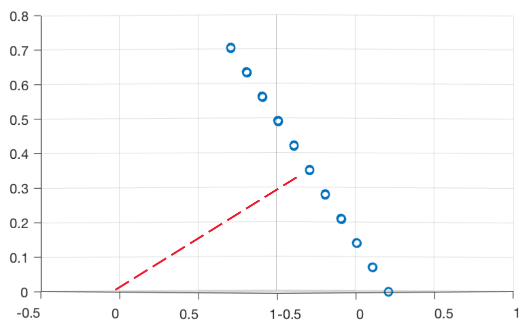
- Transform feature vectors of search points to make them incomparable.
- Apply multi-objective indicators after transformation has occurred.

Transformations (1/2)

For 2 features (transform into 3D) as follows:

- Place the unit square with its original x/y-coordinates in the three- dimensional space using z = 0.
- We rotate it around the x and y axis by 45 degrees each time.
- Translate it such that the centre point of the transformed unit square is at (sqrt(2)/4)



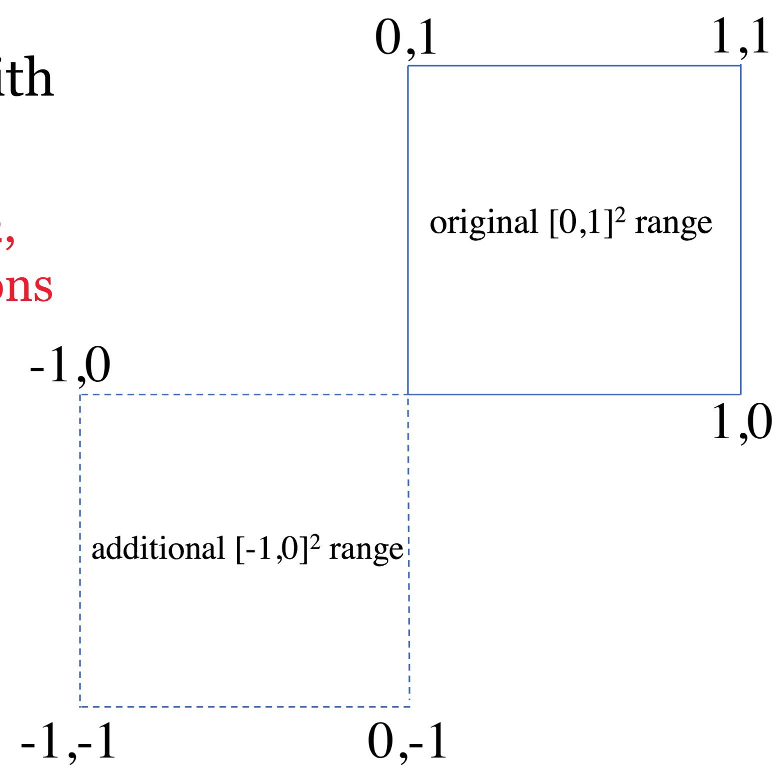


Transformations (2/2)

For d features:

- Double the number of dimensions to make vectors incomparable:
  - For feature value p_i, use p_i and $-p_i$
  - Instead of $p = (p_1, p_2, \ldots, p_d)$ work with

    $p' = (p_1, p_2, \ldots, p_d, -p_1, -p_2, \ldots, -p_d)$
- Effect: if the original vector v1 dominates v2, then v2 dominates v1 in the added dimensions → v1 and v2 are now incomparable



Which reference point in the (2*d)-dimensional space?
$(1^d, 0^d)$ would be based on the ranges' extreme values, and
$(2^d, 1^d)$ would put an increased focus on extreme points.

Algorithm

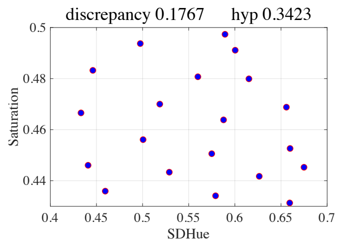
Algorithm 1: $(\mu + \lambda)$-EA_D

1 Initialize the population P with μ instances of quality at least α.
2 Let $C \subseteq P$ where $|C| = \lambda$.
3 For each $I \in C$, produce an offspring I' of I by mutation. If $q(I') \geqslant \alpha$, add I' to P.
4 While $|P| > \mu$, remove an individual with the smallest loss to the diversity indicator D.
5 Repeat step 2 to 4 until termination criterion is reached.

In plain English: it's a population-based EA that
(1) mutates lambda individuals in each generation, and
(2) considers diversity to select the survivors.

Multi-Objective Indicators (Images)

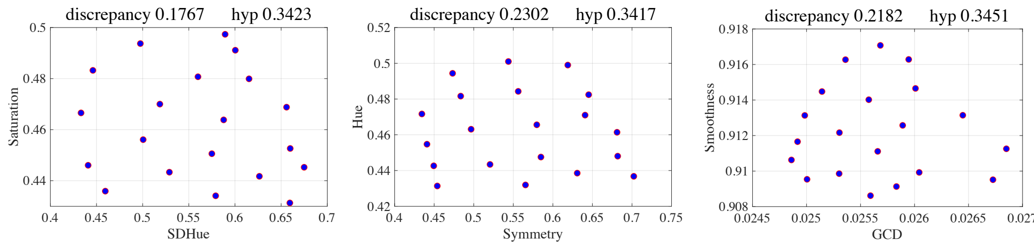
$EA_{HYP\text{-}2D}$



Multi-Objective Indicators (Images)

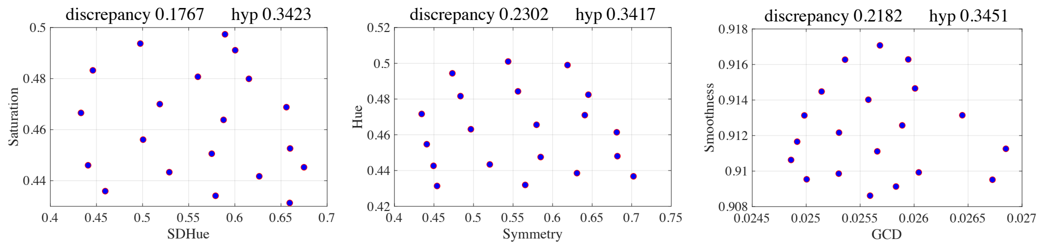
✓

$EA_{HYP\text{-}2D}$

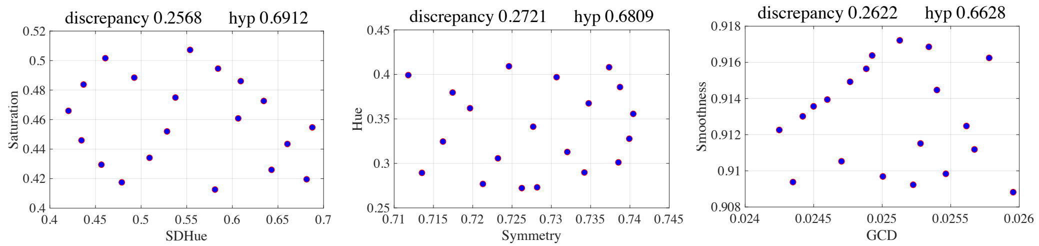


Multi-Objective Indicators (Images)

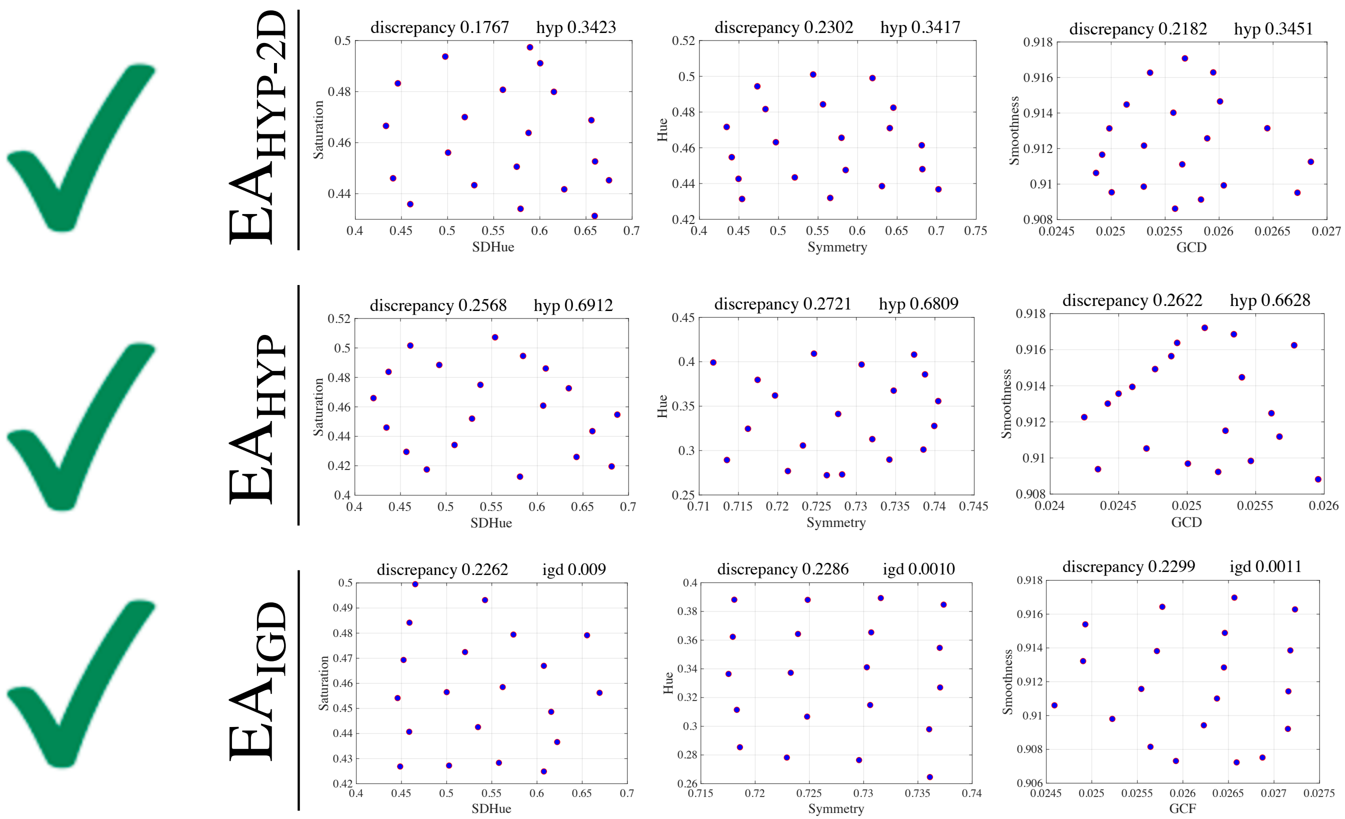
$EA_{HYP\text{-}2D}$



EA_{HYP}

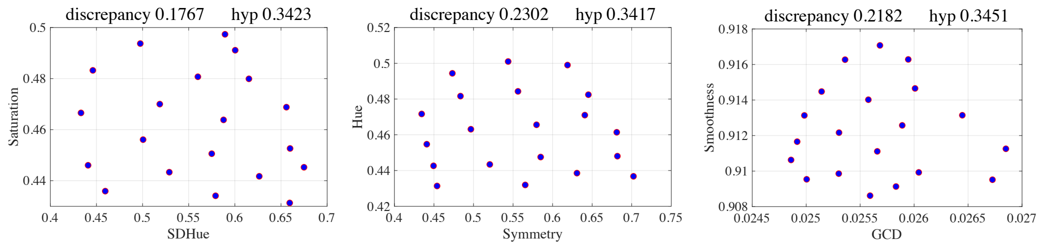


Multi-Objective Indicators (Images)

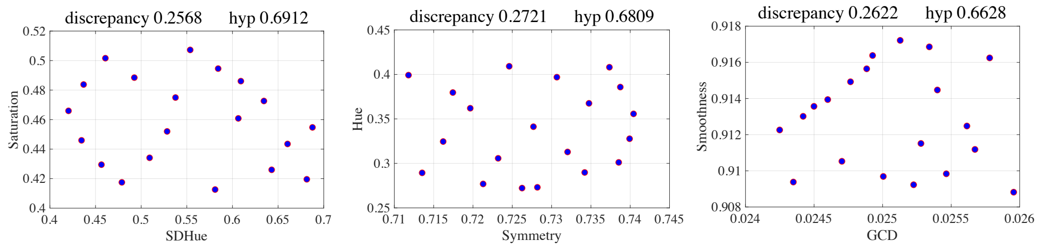


Multi-Objective Indicators (Images)

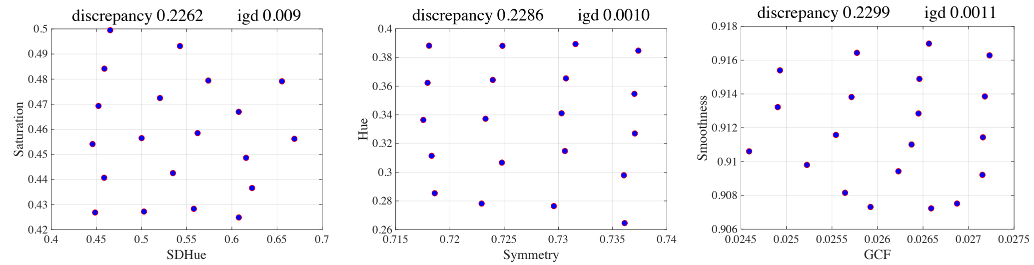
✓ $EA_{HYP\text{-}2D}$



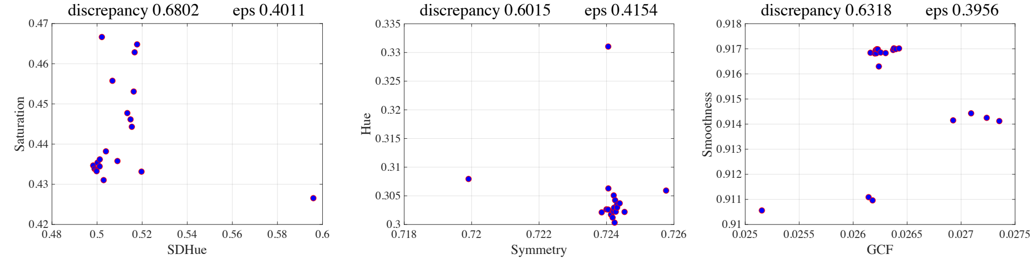
✓ EA_{HYP}



✓ EA_{IGD}



✗ EA_{EPS}



Not locally sensitive, even when using the vector of all ref.-grid approx.

Results Images

2-feature combinations

		$EA_{HYP\text{-}2D}$ (1)			EA_{HYP} (2)			EA_{IGD} (3)			EA_{EPS} (4)			EA_{DIS} (5)		
		mean	st	stat	mean	st	stat	mean	st	stat	mean	st	stat	mean	st	stat
HYP-2D	f_1,f_2	0.347	0.004	$4^{(+)},5^{(+)}$	0.382	0.007	$3^{(+)},4^{(+)},5^{(+)}$	0.335	0.003	$2^{(-)},5^{(+)}$	0.198	0.019	$1^{(-)},2^{(-)}$	0.112	0.030	$1^{(-)},2^{(-)},3^{(-)}$
	f_3,f_4	0.344	0.004	$2^{(+)},4^{(+)},5^{(+)}$	0.268	0.014	$1^{(-)},3^{(-)},4^{(+)},5^{(+)}$	0.339	0.004	$2^{(+)},4^{(+)},5^{(+)}$	0.221	0.015	$1^{(-)},2^{(-)},3^{(-)}$	0.105	0.025	$1^{(-)},2^{(-)},3^{(-)}$
	f_5,f_6	0.350	0.007	$2^{(+)},3^{(+)},4^{(+)},5^{(+)}$	0.342	0.004	$1^{(-)},4^{(+)},5^{(+)}$	0.332	0.004	$1^{(-)},4^{(+)},5^{(+)}$	0.220	0.045	$1^{(-)},2^{(-)},3^{(-)}$	0.134	0.016	$1^{(-)},2^{(-)},3^{(-)}$
HYP	f_1,f_2	0.525	0.012	$3^{(+)},4^{(+)},5^{(+)}$	0.693	0.013	$3^{(+)},4^{(+)},5^{(+)}$	0.374	0.006	$1^{(-)},2^{(-)},4^{(+)}$	0.344	0.003	$1^{(-)},2^{(-)},3^{(-)}$	0.363	0.014	$1^{(-)},2^{(-)}$
	f_3,f_4	0.500	0.007	$3^{(+)},4^{(+)},5^{(+)}$	0.681	0.010	$3^{(+)},4^{(+)},5^{(+)}$	0.268	0.072	$1^{(-)},2^{(-)},4^{(+)},5^{(+)}$	0.280	0.010	$1^{(-)},2^{(-)},3^{(-)}$	0.267	0.014	$1^{(-)},2^{(-)},3^{(-)}$
	f_5,f_6	0.518	0.012	$2^{(-)},4^{(+)},5^{(+)}$	0.663	0.010	$1^{(+)},3^{(+)},4^{(+)},5^{(+)}$	0.335	0.004	$2^{(-)},4^{(+)}$	0.317	0.006	$1^{(-)},2^{(-)},3^{(-)}$	0.327	0.008	$1^{(-)},2^{(-)}$
IGD	f_1,f_2	0.001	0.335	$2^{(+)},4^{(+)},5^{(+)}$	0.003	0.000	$1^{(-)},3^{(-)}$	0.001	0.000	$2^{(+)},4^{(+)},5^{(+)}$	0.003	0.000	$1^{(-)},3^{(-)},5^{(+)}$	0.005	0.001	$1^{(-)},3^{(-)},4^{(-)}$
	f_3,f_4	0.001	0.339	$2^{(+)},4^{(+)},5^{(+)}$	0.004	0.000	$1^{(-)},3^{(-)},5^{(+)}$	0.001	0.000	$2^{(+)},4^{(+)},5^{(+)}$	0.003	0.000	$1^{(-)},3^{(-)},5^{(+)}$	0.005	0.001	$1^{(-)},2^{(-)},3^{(-)},4^{(-)}$
	f_5,f_6	0.002	0.332	$2^{(+)},5^{(+)}$	0.007	0.000	$1^{(-)},3^{(-)},4^{(-)},5^{(-)}$	0.001	0.000	$2^{(+)},4^{(+)},5^{(+)}$	0.003	0.001	$2^{(+)},3^{(-)}$	0.004	0.001	$1^{(-)},2^{(+)},3^{(-)}$
EPS	f_1,f_2	0.190	0.198	$2^{(+)},4^{(+)},5^{(+)}$	0.498	0.011	$1^{(-)}, 3^{(-)}$	0.194	0.032	$2^{(+)},4^{(+)},5^{(+)}$	0.402	0.039	$1^{(-)},3^{(-)},5^{(+)}$	0.600	0.106	$1^{(-)},3^{(-)},4^{(-)}$
	f_3,f_4	0.198	0.221	$2^{(+)},4^{(+)},5^{(+)}$	0.569	0.016	$1^{(-)},3^{(-)}$	0.208	0.035	$2^{(+)},4^{(+)},5^{(+)}$	0.418	0.036	$1^{(-)},3^{(-)},5^{(+)}$	0.615	0.069	$1^{(-)},3^{(-)},4^{(-)}$
	f_5,f_6	0.125	0.220	$2^{(+)},4^{(+)},5^{(+)}$	0.946	0.001	$1^{(-)},3^{(-)},4^{(-)}$	0.225	0.064	$2^{(+)},4^{(+)},5^{(+)}$	0.397	0.110	$1^{(-)},2^{(+)},3^{(-)}$	0.587	0.063	$1^{(-)},3^{(-)}$
DIS	f_1,f_2	0.171	0.018	$2^{(+)},4^{(+)},5^{(+)}$	0.257	0.010	$1^{(-)},4^{(+)}$	0.201	0.031	$4^{(+)},5^{(+)}$	0.686	0.064	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.204	0.116	$1^{(-)},3^{(-)},4^{(+)}$
	f_3,f_4	0.234	0.031	$4^{(+)}$	0.273	0.041	$3^{(-)},4^{(+)},5^{(-)}$	0.198	0.017	$2^{(+)},4^{(+)}$	0.606	0.054	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.228	0.059	$2^{(+)},4^{(+)}$
	f_5,f_6	0.221	0.026	$4^{(+)}$	0.263	0.070	$3^{(-)},4^{(+)},5^{(-)}$	0.205	0.055	$2^{(+)},4^{(+)}$	0.633	0.158	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.203	0.054	$2^{(+)},4^{(+)}$

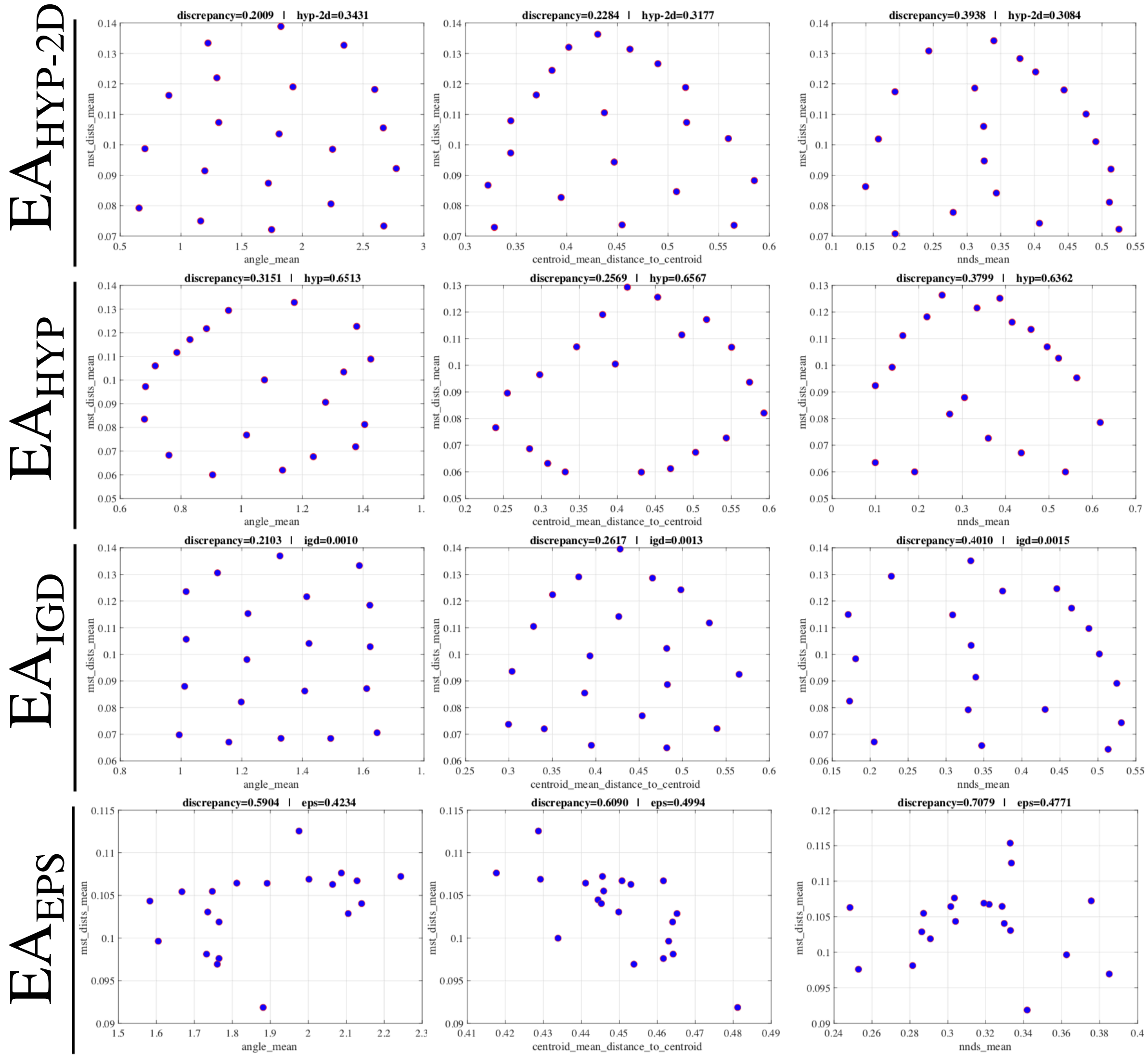
3-feature combinations

		EA_{HYP} (1)			EA_{IGD} (2)			EA_{DIS} (3)		
		mean	st	stat	mean	st	stat	mean	st	stat
HYP	f_1,f_2,f_3	0.5251	0.0122	$2^{(+)},3^{(+)}$	0.2096	0.0018	$1^{(-)},3^{(-)}$	0.2196	0.0110	$1^{(-)},2^{(+)}$
	f_1,f_4,f_3	0.4998	0.0071	$2^{(+)},3^{(+)}$	0.2142	0.0036	$1^{(-)},3^{(-)}$	0.2286	0.0034	$1^{(-)},2^{(+)}$
	f_5,f_4,f_2	0.5181	0.0122	$2^{(+)},3^{(+)}$	0.1785	0.0017	$1^{(-)},3^{(-)}$	0.1961	0.0023	$1^{(-)},2^{(+)}$
IGD	f_1,f_2,f_3	0.0146	0.0001	$2^{(-)},3^{(+)}$	0.0067	0.0003	$1^{(+)},3^{(+)}$	0.0148	0.0003	$1^{(-)},2^{(-)}$
	f_1,f_4,f_3	0.0150	0.0001	$2^{(-)}$	0.0074	0.0002	$1^{(+)},3^{(+)}$	0.0151	0.0001	$2^{(-)}$
	f_5,f_4,f_2	0.0193	0.0001	$2^{(-)},3^{(+)}$	0.0062	0.0002	$1^{(+)},3^{(+)}$	0.0199	0.0007	$1^{(-)},2^{(-)}$
DIS	f_1,f_2,f_3	0.3554	0.0458	$2^{(+)},3^{(-)}$	0.3809	0.0522	$1^{(-)},3^{(-)}$	0.3350	0.1002	$1^{(+)},2^{(+)}$
	f_1,f_4,f_3	0.3493	0.0532	$2^{(-)}$	0.2860	0.0342	$1^{(+)},3^{(+)}$	0.3118	0.1309	$2^{(-)}$
	f_5,f_4,f_2	0.4237	0.0643	$2^{(-)},3^{(-)}$	0.3227	0.0557	$1^{(+)},3^{(-)}$	0.3007	0.1467	$1^{(+)},2^{(+)}$

30 independent runs per setup

Kruskal-Wallis-Test: like a multi-set Mann-Whitney-U-Test

Multi-Objective Indicators (TSP)



Not locally sensitive, even when using the vector of all ref.-grid approx.

Results TSP

2-feature combinations

		$EA_{HYP\text{-}2D}$ (1)			EA_{HYP} (2)			EA_{IGD} (3)			EA_{EPS} (4)			EA_{DIS} (5)		
		mean	st	stat	mean	st	stat	mean	st	stat	mean	st	stat	mean	st	stat
HYP-2D	f_1,f_4	0.338	2E-3	$2^{(+)},4^{(+)},5^{(+)}$	0.309	4E-3	$1^{(-)},4^{(+)}$	0.331	3E-3	$4^{(+)},5^{(+)}$	0.190	1E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.256	1E-2	$1^{(-)},3^{(-)}$
	f_2,f_4	0.317	3E-3	$2^{(+)},4^{(+)},5^{(+)}$	0.303	5E-3	$1^{(-)},3^{(-)},4^{(+)}$	0.316	3E-3	$2^{(+)},4^{(+)},5^{(+)}$	0.178	1E-7	$1^{(-)},2^{(-)},3^{(-)}$	0.252	1E-2	$1^{(-)},3^{(-)}$
	f_3,f_4	0.303	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.296	5E-3	$1^{(-)},3^{(-)},4^{(+)},5^{(+)}$	0.304	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.190	2E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.238	2E-2	$1^{(-)},2^{(-)},3^{(-)}$
HYP	f_1,f_4	0.645	5E-3	$4^{(+)},5^{(+)}$	0.638	7E-3	$4^{(+)},5^{(+)}$	0.639	6E-3	$4^{(+)},5^{(+)}$	0.424	2E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.529	3E-2	$1^{(-)},2^{(-)},3^{(-)}$
	f_2,f_4	0.609	7E-3	$2^{(-)},4^{(+)},5^{(+)}$	0.632	1E-2	$1^{(+)},4^{(+)},5^{(+)}$	0.621	6E-3	$4^{(+)},5^{(+)}$	0.398	1E-6	$1^{(-)},2^{(-)},3^{(-)}$	0.505	2E-2	$1^{(-)},2^{(-)},3^{(-)}$
	f_3,f_4	0.584	3E-2	$2^{(-)},4^{(+)}$	0.621	9E-3	$1^{(+)},3^{(+)},4^{(+)},5^{(+)}$	0.595	4E-2	$2^{(-)},4^{(+)},5^{(+)}$	0.410	2E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.485	3E-2	$2^{(-)},3^{(-)}$
IGD	f_1,f_4	0.001	2E-5	$4^{(+)},5^{(+)}$	0.001	6E-5	$3^{(-)},4^{(+)}$	0.001	4E-5	$2^{(+)},4^{(+)},5^{(+)}$	0.003	2E-5	$1^{(-)},2^{(-)},3^{(-)}$	0.002	2E-4	$1^{(-)},3^{(-)}$
	f_2,f_4	0.001	3E-5	$2^{(+)},4^{(+)},5^{(+)}$	0.002	6E-5	$1^{(-)},3^{(-)},4^{(+)}$	0.001	3E-5	$2^{(+)},4^{(+)},5^{(+)}$	0.003	2E-10	$1^{(-)},2^{(-)},3^{(-)}$	0.002	2E-4	$1^{(-)},3^{(-)}$
	f_3,f_4	0.002	3E-4	$4^{(+)},5^{(+)}$	0.002	6E-5	$3^{(-)},4^{(+)},5^{(+)}$	0.002	3E-4	$2^{(+)},4^{(+)},5^{(+)}$	0.003	3E-5	$1^{(-)},2^{(-)},3^{(-)}$	0.003	3E-4	$1^{(-)},2^{(-)},3^{(-)}$
EPS	f_1,f_4	0.196	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.249	2E-2	$1^{(-)},3^{(-)},4^{(+)}$	0.189	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.423	1E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.345	4E-2	$1^{(-)},3^{(-)}$
	f_2,f_4	0.226	8E-3	$2^{(+)},4^{(+)},5^{(+)}$	0.256	2E-2	$1^{(-)},3^{(-)},4^{(+)},5^{(+)}$	0.228	1E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.499	2E-16	$1^{(-)},2^{(-)},3^{(-)}$	0.360	5E-2	$1^{(-)},2^{(-)},3^{(-)}$
	f_3,f_4	0.260	4E-2	$4^{(+)},5^{(+)}$	0.278	2E-2	$4^{(+)},5^{(+)}$	0.265	4E-2	$4^{(+)},5^{(+)}$	0.477	3E-3	$1^{(-)},2^{(-)},3^{(-)}$	0.368	5E-2	$1^{(-)},2^{(-)},3^{(-)}$
DIS	f_1,f_4	0.222	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.353	2E-2	$1^{(-)},3^{(-)},4^{(+)}$	0.249	2E-2	$2^{(+)},4^{(+)}$	0.589	4E-3	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.292	5E-2	$1^{(-)},4^{(+)}$
	f_2,f_4	0.230	2E-2	$2^{(+)},4^{(+)},5^{(+)}$	0.274	2E-2	$1^{(-)},4^{(+)},5^{(+)}$	0.252	1E-3	$4^{(+)},5^{(+)}$	0.609	1E-16	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.336	4E-2	$1^{(-)},2^{(-)},3^{(-)},4^{(+)}$
	f_3,f_4	0.418	6E-2	$4^{(+)}$	0.416	3E-2	$4^{(+)}$	0.401	7E-2	$4^{(+)},5^{(+)}$	0.719	6E-3	$1^{(-)},2^{(-)},3^{(-)},5^{(-)}$	0.448	9E-2	$3^{(-)},4^{(+)}$

3-feature combinations

		EA_{HYP} (1)			EA_{IGD} (2)			EA_{DIS} (3)		
		mean	st	stat	mean	st	stat	mean	st	stat
HYP	f_1,f_2,f_3	0.4511	1E-2	$2^{(+)},3^{(+)}$	0.4261	7E-3	$1^{(-)},3^{(+)}$	0.3385	6E-3	$1^{(-)},2^{(-)}$
	f_1,f_3,f_4	0.4579	8E-3	$2^{(+)},3^{(+)}$	0.4260	6E-3	$1^{(-)},3^{(+)}$	0.3430	6E-3	$1^{(-)},2^{(-)}$
	f_2,f_3,f_4	0.4478	8E-3	$2^{(+)},3^{(+)}$	0.4262	6E-3	$1^{(-)},3^{(+)}$	0.3430	6E-3	$1^{(-)},2^{(-)}$
IGD	f_1,f_2,f_3	0.0083	3E-4	$2^{(-)},3^{(+)}$	0.0075	2E-4	$1^{(+)},3^{(+)}$	0.0110	1E-4	$1^{(-)},2^{(-)}$
	f_1,f_3,f_4	0.0082	2E-4	$2^{(-)},3^{(+)}$	0.0077	1E-4	$2^{(+)},3^{(+)}$	0.0107	1E-4	$1^{(-)},2^{(-)}$
	f_2,f_3,f_4	0.0086	2E-4	$2^{(-)},3^{(+)}$	0.0080	2E-2	$2^{(+)},3^{(+)}$	0.0112	8E-5	$1^{(-)},2^{(-)}$
DIS	f_1,f_2,f_3	0.4115	3E-2	$2^{(+)},3^{(+)}$	0.4839	3E-2	$1^{(-)},3^{(-)}$	0.4399	2E-2	$1^{(-)},2^{(+)}$
	f_1,f_3,f_4	0.5220	4E-2	$3^{(-)}$	0.5474	3E-2	$3^{(-)}$	0.4757	2E-2	$1^{(+)},2^{(+)}$
	f_2,f_3,f_4	0.4669	3E-2	$2^{(+)}$	0.5111	3E-2	$1^{(-)},3^{(-)}$	0.4667	2E-2	$2^{(+)}$

In summary:

- EA_{HYP} and EA_{IGD} perform best
- Beats our GECCO'18 results (discrepancy theory)

30 independent runs per setup

Kruskal-Wallis-Test: like a multi-set Mann-Whitney-U-Test

Some of the next questions to answer:

- What type of features are good to characterize problem instances of a given problem (e.g. TSP) for a particular algorithm.
- What is a good diversity measure?
- What is the runtime behavior of EAs maximizing search-space/feature diversity?
- How do we compute diverse sets of high-quality solutions for important combinatorial optimization problems?
- How do we change state of the art solvers to compute diverse sets of solutions (instead of a single one)

- We provide code: https://tinyurl.com/geccoDiversity (Java code, Matlab wrapper provided)
- Email: markus.wagner@adelaide.edu.au