Advanced Parallel Copula Based EDA

Martin Hyrš
Brno University of Technology
Faculty of Information Technology
Božetěchova 1/2, Brno, Czech Republic
ihyrs@fit.vutbr.cz

Josef Schwarz
Brno University of Technology
Faculty of Information Technology
Božetěchova 1/2, Brno, Czech Republic
schwarz@fit.vutbr.cz

***Abstract*—Estimation of distribution algorithms (EDAs) are stochastic optimization techniques that are based on building and sampling a probability model. Copula theory provides methods that simplify the estimation of the probability model. To improve the efficiency of current copula based EDAs (CEDAs) new modifications of parallel CEDA were proposed. We investigated eight variants of island-based algorithms utilizing the capability of promising copula families, inter-island migration and additional adaptation of marginal parameters using CT-AVS technique. The proposed algorithms were tested on two sets of well-known standard optimization benchmarks in the continuous domain. The results of the experiments validate the efficiency of our algorithms.**

I. Introduction

Estimation of distribution algorithms (EDAs) belong to a new class of evolutionary optimization methods that explore the search space by estimating and sampling an explicit probabilistic model of promising solutions. EDAs applied to discrete problems are described in the well-known papers UMDA [1], BMDA [2], MIMIC [3], and BOA [4]. Solutions of the optimization problems in the real value domain can be found in [5]. A very modern and accessible survey of the EDAs algorithm is presented in [6].

The main advantage of EDAs is their capacity to discover those variable linkages that yield a solution to a complex optimization problem. On the one hand this probability model-based approach has allowed EDAs to be applied to large and complex problems. On the other hand, explicit probabilistic models are very time consuming. That was the reason for implementing various advanced EDAs to solve this problem.

In the last ten years a new approach to building an efficient probabilistic model based on copula theory has appeared [7]. Copulas are special probability distribution functions. Due to their properties it is possible to use them effectively to model correlations within multivariate problems – the joint distribution is separated into the univariate marginal distributions and into the correlation structure that is expressed by the copula function. Copula theory has very often been used in finance and statistics works [8], [9], [10].

Recently copulas have been utilized in the field of the machine learning [11], [12]. More recently the copula theory has been applied to EDA probability models. The simplest case is the application with bivariate (2D) copulas, e.g.: [13] – 2D Gaussian copula EDA, [14] – 2D Clayton copula EDA, [15] – 2D Gumbel copula EDA.

In the case of multivariate models bivariate copulas are used as local building blocks in various graph dependence structures: [16] – MIMIC with Frank and Gaussian copula), [17] – Bayesian network with Archimedean copulas, [18] – D-vine copulas, [19] – C-vine, D-vine copulas.

Detailed overview of different copula-based EDAs is provided in [20].

This paper deals with an extension of the original concept of copula based EDA with the probability model migration [21]. The enhancement includes the additional migration of individuals and the adaptation of the margins parameters using CT-AVS technique [22] preventing the premature convergence.

The paper is organized as follows. In Section II, the basics of copula theory is given. In Section III, the utilization of copulas in EDA is described. Two variants of the island-based EDAs are described in Section IV – EDA with the migration of probability model and EDA with the migration of individuals. Next, CT-AVS adaptation of margins is briefly described in Section V. Experimental results are discussed in Section VI. The conclusions are given in Section VII.

II. Copula theory

The copula concept was introduced by [23] in order to separate the effect of dependence of variables from the effect of marginal distributions in a joint distribution. A copula is a function which joins the univariate distribution function and creates multivariate distribution functions. This approach allows us to transform multivariate statistic problems into the univariate problems with the relation represented by just the copula.

A copula C is a multivariate probability distribution function for which the marginal probability distribution of each variable is uniform in $[0; 1]$.

Sklar's theorem: *Let F be a d-dimensional distribution function with margins $F_1, \ldots, F_d$. Then there exists a d-dimensional copula C such that for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$ it holds that*

$$F(x_1, \ldots, x_d) = C\left(F_1(x_1), \ldots, F_d(x_d)\right) \quad (1)$$

If $F_1, \ldots, F_d$ are continuous, then C is unique. Conversely, if C is a d-dimensional copula and $F_1, \ldots, F_d$ are univariate distribution functions, then the function F defined via (1) is a d-dimensional distribution function.

We investigated two copula families – Archimedean and elliptic:

Archimedean copulas are quite popular because they model different patterns of dependence and have a relatively simple functional form $C(u_1, \dots, u_d) = \varphi_\theta(\varphi_\theta^{-1}(u_1) + \dots + \varphi_\theta^{-1}(u_d))$. Their definition is based on the generator function φ with typically one dependency parameter θ.

The elliptical copulas are derived from the related elliptical distribution, e.g. Gaussian copula $C(u_1, \dots, u_d) = \Phi_R\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)\right)$ where $\Phi_R(x_1, \dots, x_d)$ is the joint normal distribution with a positive-semidefinite correlation matrix R, Φ is standard normal distribution, Φ^{-1} is its quantile function.

III. Copula-based Estimation of Distribution Algorithm

Estimation of distribution algorithms belongs to the advanced evolutionary algorithms. Solving the numerical optimization problem, vector $\mathbf{x} = (x_1, \dots, x_d)$ of the optimal solution is searched out.

The core of the canonical EDA consists of three main steps:

1) Select promising solutions into subpopulation from the current population.
2) Create the probability model from the selected subpopulation.
3) Sample the probability model and generate the new population.

In the case of copula-based EDA it is necessary to choose the proper type of copula and derive the copula parameters and the marginal distribution parameters.

The principle of sampling schema for generating the new individuals using the copula model is described in two steps:

1) Obtain the random copula sample $(u_1, \dots, u_d) \sim C$, where $u_i \in [0; 1]$.
2) Derive the vector $\mathbf{x}$ of the searched solution using inverse marginal distributions, $x_i = F_i^{-1}(u_i)$.

A. Identification of copula probability model

The copula-based probability model includes two parts: univariate marginal distributions and the copula function. The marginal distributions can be identified separately for each variable and the copula includes the correlation between variables.

For the marginal distribution in each dimension $i = 1, \dots, d$ we used normal distribution, which is parameterized by the mean value μ_i and standard deviation σ_i.

For assessing the parameter of Archimedean copulas we used the Kendall τ correlation coefficient (in the case of d-variate copulas, $d \geq 3$, we use average $\bar{\tau}$; for $d = 2$ the standard pairwise τ is used). The following relation hold for the parameter θ for Clayton copula $\theta_{Clayton} = \frac{2\tau}{1-\tau}$ [7].

Elliptical copulas are parameterized by correlation matrix R. Elements R_{ij} are calculated as Spearman's ρ_S for each pair of dimensions.

1) Generate initial populations.
2) FOR each island DO IN PARALLEL:
3) WHILE (termination criteria is false):
4) IF (sending condition):
5) Send best $replace\_count$ individuals
6) WHILE (immigrant individuals received):
7) Replace worst $replace\_count$ individuals in resident population
8) Select promising individuals
9) Create probability model
10) Sample new population from probability model

Fig. 1. The pseudocode of EDA with migration of individuals.

1) Generate initial populations.
2) FOR each island DO IN PARALLEL:
3) WHILE (termination criteria is false):
4) Select promising individuals
5) Create probability model
6) IF (sending condition):
7) Send model
8) WHILE (immigrant model received):
9) Combine models
10) Sample new population from probability model

Fig. 2. The pseudocode of EDA with model migration.

IV. Island-based copula EDA

We implemented the parallel copula based EDA on the island based platform. This concept allows to study the efficiency of the migration of individuals and our new concept of migration of probabilistic models.

The evolution process on every island runs independently. When a sending condition is met the communication (transfer of individuals or transfer of model parameters) is activated.

A. EDA with migration of individuals

Migration of individuals is well-known approach used in the island-based algorithm [24], [25]. In comparison to the probabilistic model migration the transfer of individuals does not cause any troubles with the mixture of the various probability model of each island population.

The effect of migration is influenced by coefficient *replace count* – the number of individuals which are sent to neighboring island. When the sent individuals are received, the promising solution are selected and the new model is built, see Fig. 1.

B. EDA with migration of probability models

In the case of EDAs only a few papers deal with the probability model migration [26], [27], [28]. On the field of copula-based EDA we refer to [21] in which we tested first version of model migration.

According to the island-based topology we have decomposed the migration process into pairwise interactions of two islands, see Fig. 2 – one of them is the resident island specified by resident probabilistic model M_R and the other one is the immigrant island whose probabilistic model M_I is transferred to the new resident model.

The combination of the immigrant model with the model of the resident island is described in more details. In general, the modification of the resident model by the immigrant model can be formalized by [27]:

$$M_R^{new} = (1-\beta)M_R + \beta M_I \tag{2}$$

where the coefficient $\beta \in [0;1]$ specifies the influence of the immigrant model.

We have proposed the following model combination rules according to [29]:

- Learning the mean value μ_i of each univariate marginal distribution $F_i(x_i)$

$$\mu_i^{new} = (1-\beta)\mu_i^R + \beta\mu_i^I \tag{3}$$

- Learning the standard deviation σ_i of each univariate marginal distribution $F_i(x_i)$

$$\sigma_i^{new} = \sqrt{(1-\beta)\left(\left(\mu_i^{new}-\mu_i^R\right)^2 + \left(\sigma_i^R\right)^2\right) + \beta\left(\left(\mu_i^{new}-\mu_i^I\right)^2 + \left(\sigma_i^I\right)^2\right)} \tag{4}$$

- Learning the correlation matrix value R_{ij}

$$R_{ij}^{new} = (1-\beta)R_{ij}^R + \beta R_{ij}^I \tag{5}$$

We have chosen the coefficient β as

$$\beta = \begin{cases} \frac{fit^R}{fit^R+fit^I} & fit^I \le fit^R \\ 0.1 & \text{otherwise} \end{cases} \tag{6}$$

where fit^R or fit^I represents the fitness value of the resident or the immigrant model.

V. Improved Correlation-Triggered Adaptive Variance Scaling

The Correlation-Triggered Adaptive Variance Scaling (CT-AVS) [22] is a technique which is supposed to have the capability of preventing premature convergence. It allows to distinguish between *peak* and *slopes* regions of fitness landscape. At the peak regions, the normal marginal pdf can match contour-lines of the fitness landscape appropriately as it always concentrates the search around its mean and therefore can contract around a peak.

At the slope-like regions, contour-lines of slopes cannot be matched with the normal pdf. Thus, the probabilistic representation of the structure based on mean and variance is different from the true structure, so estimates from the normal pdf are a much less reliable source of information for guiding the search compared to exploring a single peak.

This problem can be partially solved using *variance scaling* with parameter c^{AVS}. When the fitness landscape in actually searched region is identified to be a slope-like region, the variance is increased beyond its maximum-likelihood estimate, to move out of this region. To identify the type of region,

i-CT-AVS initialization:

1) Assign $c^{\mathrm{AVS-MIN}} = 1/c^{\mathrm{AVS-MAX}}$.
2) Assign $\eta^{\mathrm{INC}} = 1/\eta^{\mathrm{DEC}}$.
3) Assign $c^{\mathrm{AVS}}[d] = 1 \quad \forall d = 1, \ldots, D$.

i-CT-AVS in generation t for each dimension d:

1) Store best fitness found in b^t.
2) If $b^t = b^{t-1}$ then
   (a) Assign $c^{\mathrm{AVS}}[d] = c^{\mathrm{AVS}}[d] \cdot \eta^{\mathrm{DEC}}$.
   else
   (b) Assign $c^{\mathrm{AVS}}[d] = c^{\mathrm{AVS}}[d] \cdot \eta^{\mathrm{INC}}$.
3) If $c^{\mathrm{AVS}}[d] < c^{\mathrm{AVS-MIN}}$ or $c^{\mathrm{AVS}}[d] > c^{\mathrm{AVS-MAX}}$ then
   (a) assign $c^{\mathrm{AVS}}[d] = c^{\mathrm{AVS-MAX}}$.
4) Compute ranked correlation coefficient $r[d]$.
5) If $r[d] > \theta^{corr}$ then
   (a) Assign $\sigma[d] = c^{\mathrm{AVS}}[d] \cdot \sigma[d]$.

Fig. 3. Principle of iCTAVS.

correlation between fitness and the density of modelled normal distribution is used. There is a strong correlation in the peak regions (positive correlation for maximization tasks, negative for minimization ones). Rank correlation Spearman's ρ_S is used. Variance scaling factor c^{AVS} is adaptive, it is increased if the best fitness value b^t improves in one generation, and decreased in the opposite case via coefficients η^{INC}, η^{DEC}.

In [22], the variance scaling is executed in all dimensions by the same factor c^{AVS}. We designed an improved version of this technique (iCTAVS): we use factors $c^{\mathrm{AVS}}[d]$ calculated independently in each dimension d, see Fig. 3.

VI. Results of Experiments

For the experimental testing we used two sets of benchmarks, all these functions have been modified to get the minimization task with optimal fitness value 0:

1) the popular classical set of benchmarks (Sphere, Schwefel's, Ackley's, Rastrigin's, Rosenbrock's, Griewank's).
2) the complex set CEC 2013 [30] with rotated and composition function, including 28 test functions.

We have proposed eight variants of copula based algorithms, see Fig. 4, labeled by the abbreviations as a composition of the following variant of components:

- two different copulas: Gaussian copula (GC-) or Clayton copula (CC-)
- two types of migration: migration of individuals (-mi-) or migration of probability model (-mm-)
- two types of marginal distributions: normal pdf (-) and normal pdf modified by iCTAVS (-avs).

All proposed algorithms used the following parameters setting:

- Number of islands: 4.
- Topology: bi-directional ring.
- Population size of each island: 250.
- Selection: truncation selection, with a selection proportion of 0.2, i.e. 50 individuals.

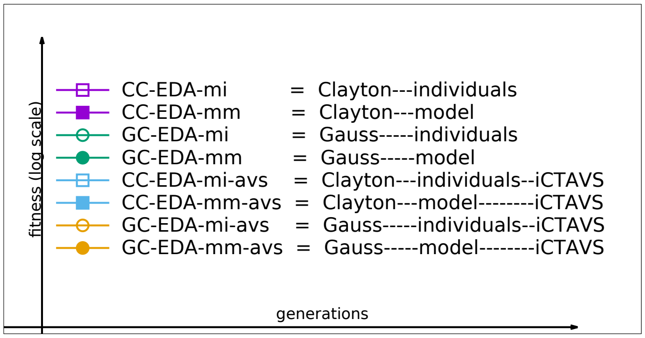

Fig. 4. The list of the proposed variants of our algorithm and their shortcuts and graphical key valid for Fig. 5 and Fig. 6.

- Replace count (for the migration of individuals): 40.
- Migration rate: after every 10 generations.
- Number of independent runs: 51.

A. *Mutual comparison of the proposed algorithms*

In the case of classical benchmarks optimization curves are drawn in Fig. 5.

For the Spheric problem, all eight variants are able to reach optima with satisfactory precision (10^{-8}), the best is CC-EDA-mi, the variants using iCTAVS adaptation are the worst.

In the case of Schwefel's function, the performance of all variants is unsatisfactory. The only GC-EDA-mm and GC-EDA-mi are relatively good.

In the case of Ackley's, Griewank's and Rastrigin's functions: CC-EDA-mm-avs and GC-EDA-mm-avs get stuck in local optima, while other algorithm variants (GC-EDA-mi, CC-EDA-mi, GC-EDA-mm, CC-EDA-mm, GC-EDA-mi-avs, CC-EDA-mi-avs) are able to reach global optima. The variants without iCTAVS component converge faster.

The most interesting results were obtained for Rosenbrock's problem. On this well-known benchmark, many algorithms get stuck in local optima near to 10^0. Only GC-EDA-mi-avs and GC-EDA-mm-avs are capable to reach global optima.

In the case of CEC 2013 benchmarks the optimization curves are drawn in Fig. 6. It can be recognized that the variants with Gaussian copula have the best performance on the most of the test functions.

Generally, GC-EDA-mm can be recognized as the best variant. To confirm this conclusion we executed statistical tests comparing GC-EDA-mm with other seven variants. Results of Mann–Whitney U test (Wilcoxon rank-sum test) are listed in Tab. I.

Results for GC-EDA-mm on benchmark set CEC 2013 for 10, 30, 50 dimensions are arranged in Tab. III, IV and V according to rules provided in [30] (values smaller than 10^{-8} are taken as zero).

B. *Comparison with recently published algorithms*

Two variants of our proposed algorithm are selected for an extra comparison with algorithms of other researchers: primarily the GC-EDA-mm is the our best variant; the GC-EDA-mi-avs shows remarkably good performance but on a few functions only.

TABLE I

U-VALUE OF MANN–WHITNEY U TEST (WILCOXON RANK-SUM TEST) BETWEEN **GC-EDA-MM** AND THE OTHER VARIANTS ON 10 DIMENSIONAL BENCHMARKS. $U^{Crit}_{n_1=n_2=51,p=0.05} = 1054$, THE MARK "▼" DENOTES THE CASES WHERE GC-EDA-MM PERFORMS SIGNIFICANTLY BETTER ($U \leq U^{Crit}$), THE MARK "△" DENOTES THE CASES WHERE OTHER VARIANT IS SIGNIFICANTLY BETTER AND NO SIGNIFICANT DIFFERENCE IS DENOTED BY THE MARK "○".

Fun.	CC-mi	CC-mm	GC-mi	CC-mi -avs	CC-mm -avs	GC-mi -avs	GC-mm -avs
Sphere	2601 △	757 ▼	1997 △	0 ▼	0 ▼	0 ▼	0 ▼
Ackley	2601 △	751 ▼	2187 △	0 ▼	0 ▼	1 ▼	0 ▼
Schwefel	551 ▼	854 ▼	928 ▼	111 ▼	35 ▼	13 ▼	3 ▼
Rastrigin	1080 ○	509 ▼	979 ▼	220 ▼	169 ▼	207 ▼	165 ▼
Rosenbrock	110 ▼	525 ▼	922 ▼	1020 ▼	636 ▼	2601 △	2276 △
Griewank	1300 ○	1300 ○	1045 ▼	25 ▼	0 ▼	0 ▼	0 ▼
No. 1	2375 △	2046 △	907 ▼	384 ▼	377 ▼	385 ▼	369 ▼
No. 2	0 ▼	0 ▼	1019 ▼	0 ▼	0 ▼	258 ▼	394 ▼
No. 3	53 ▼	121 ▼	523 ▼	2017 △	1793 △	2502 △	1284 ○
No. 4	0 ▼	0 ▼	2058 △	3 ▼	4 ▼	52 ▼	30 ▼
No. 5	2544 △	2529 △	638 ▼	1112 ○	963 ▼	1146 ○	944 ▼
No. 6	0 ▼	0 ▼	862 ▼	0 ▼	0 ▼	2225 △	1884 △
No. 7	0 ▼	0 ▼	540 ▼	4 ▼	0 ▼	1043 ▼	270 ▼
No. 8	922 ▼	712 ▼	1385 ○	1152 ○	945 ▼	1367 ○	1283 ○
No. 9	14 ▼	54 ▼	1072 ○	1 ▼	0 ▼	507 ▼	186 ▼
No. 10	2 ▼	3 ▼	1080 ○	2 ▼	4 ▼	441 ▼	385 ▼
No. 11	1116 ○	1292 ○	935 ▼	1165 ○	1074 ○	713 ▼	985 ▼
No. 12	107 ▼	111 ▼	1294 ○	359 ▼	138 ▼	1113 ○	1050 ▼
No. 13	174 ▼	346 ▼	1256 ○	670 ▼	696 ▼	1588 △	1668 △
No. 14	176 ▼	188 ▼	1455 ○	39 ▼	38 ▼	1 ▼	0 ▼
No. 15	157 ▼	239 ▼	1468 ○	32 ▼	0 ▼	172 ▼	39 ▼
No. 16	1054 ▼	854 ▼	1267 ○	1162 ○	984 ▼	1614 △	1304 ○
No. 17	2282 △	2218 △	1759 △	2093 △	1032 ▼	1587 △	1204 ○
No. 18	2459 △	2516 △	1269 ○	1857 △	1064 ○	1975 △	1246 ○
No. 19	2137 △	2584 △	1221 ○	1837 △	1676 △	825 ▼	723 ▼
No. 20	0 ▼	0 ▼	1804 △	1 ▼	0 ▼	0 ▼	0 ▼
No. 21	1326 ○	1326 ○	1248 ○	1326 ○	1326 ○	1376 ○	1351 ○
No. 22	1150 ○	1019 ▼	1522 ○	275 ▼	110 ▼	181 ▼	86 ▼
No. 23	167 ▼	169 ▼	1748 △	0 ▼	0 ▼	10 ▼	0 ▼
No. 24	1 ▼	128 ▼	145 ▼	0 ▼	0 ▼	0 ▼	51 ▼
No. 25	16 ▼	129 ▼	202 ▼	51 ▼	153 ▼	0 ▼	0 ▼
No. 26	99 ▼	148 ▼	1282 ○	6 ▼	15 ▼	22 ▼	6 ▼
No. 27	1010 ▼	2016 △	240 ▼	0 ▼	0 ▼	0 ▼	0 ▼
No. 28	2575 △	2575 △	1096 ○	2563 △	2550 △	2568 △	2550 △

In Tab. II we arranged a comparison of our prime variant with other algorithms on CEC 2013 benchmarks (unfortunately we did not found any Copula EDA tested on these benchmarks). We performed unpaired two-sample t-test at the significance level 0.05 to show statistical significance of our comparison, see Tab. II. It can be seen that GC-EDA-mm is significantly better especially for the complex composition functions (No. 21–28).

In Tables VI, VII and VIII we arranged a comparison (mean fitness values) of both variants, GC-EDA-mm and GC-EDA-mi-avs, with the algorithms that used different versions of copulas. The comparison is done with the same subset of classical benchmarks with 10 dimensions for the same number of fitness evaluations. GC-EDA-mm is better in most cases. In the case of Rosenbrock's function GC-EDA-mi-avs is the best.

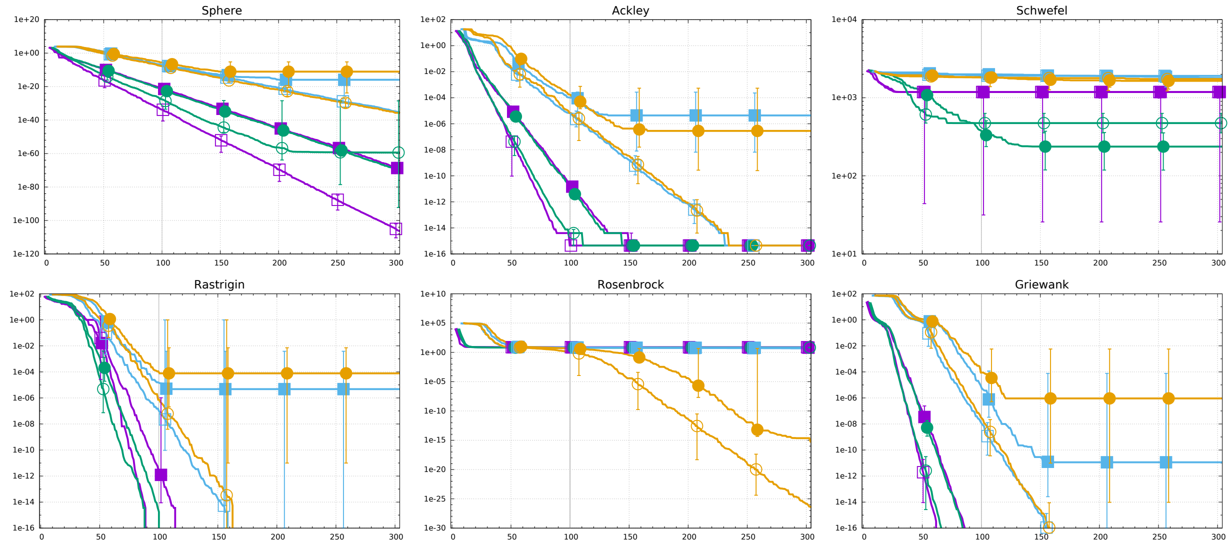

Fig. 5. Result for classical benchmarks. For key, please see Fig. 4

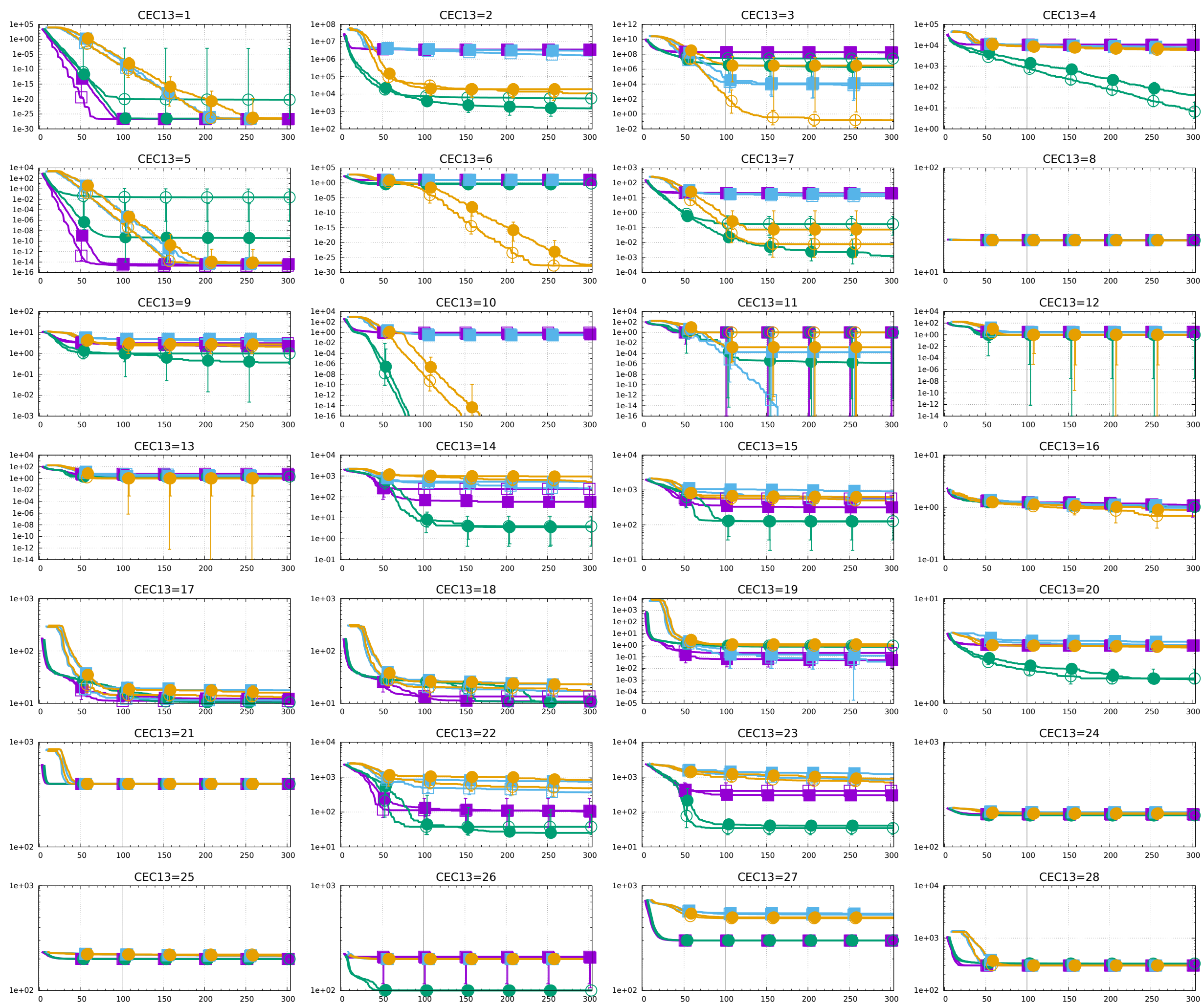

Fig. 6. Result for benchmark set CEC 2013 in 10 dimensions. For key, please see Fig. 4

TABLE II

COMPARISON (MEDIANS) OF **GC-EDA-MM** WITH OTHER PUBLISHED ALGORITHMS ON CEC 2013 BENCHMARK SET IN 10 DIMENSIONS AFTER 100000 FITNESS EVALUATION, BEST PERFORMING ALGORITHM IS INDICATED BY BOLDFACE FONT. RESULTS OF T-TEST, THE MARK "▼" DENOTES THE CASES WHERE GC-EDA-MM PERFORMS SIGNIFICANTLY BETTER, THE MARK "△" DENOTES THE CASES WHERE OTHER ALGORITHM IS SIGNIFICANTLY BETTER AND NO SIGNIFICANT DIFFERENCE IS DENOTED BY THE MARK "○".

No.	GC-mm	IPOP-CMA-ES [31]	iCMAES-ILS [32]	CMA-ES-RIS [33]	SPSO2011 [34]	PSO [35]	GA-TPC [36]	LaF [37]	SPAM-AOS [38]
1	**0.00e+00**	**0.00e+00** ○	1.00e-08 ○	**0.00e+00** ○	**0.00e+00** ○	**0.00e+00** ○	**0.00e+00** ○	**0.00e+00** ○	**0.00e+00** ○
2	3.84e+03	**0.00e+00** △	1.00e-08 △	**0.00e+00** △	1.59e+04 ▼	3.63e+04 ▼	**0.00e+00** △	9.65e+04 ▼	**0.00e+00** △
3	3.42e+06	**0.00e+00** △	1.00e-08 △	**0.00e+00** △	2.31e+03 △	2.68e+05 △	**0.00e+00** △	2.60e+05 ○	1.31e+00 △
4	1.41e+03	**0.00e+00** △	1.00e-08 △	**0.00e+00** △	1.29e+03 △	8.87e+03 ▼	**0.00e+00** △	5.57e+02 △	**0.00e+00** △
5	**0.00e+00**	**0.00e+00** △	1.00e-08 △	**0.00e+00** △	**0.00e+00** △	**0.00e+00** △	**0.00e+00** △	**0.00e+00** △	1.14e-13 △
6	3.18e-01	**0.00e+00** △	6.57e-05 ▼	2.42e-03 ○	9.80e+00 ▼	9.80e+00 ▼	**0.00e+00** △	9.81e+00 ▼	4.11e+00 ▼
7	2.27e-02	**0.00e+00** △	3.56e-07 △	4.01e+01 ▼	4.00e-01 ○	2.11e+01 ▼	1.42e-03 ○	1.11e+01 ▼	6.04e+01 ▼
8	2.04e+01	2.04e+01 △	2.04e+01 ○	2.03e+01 △	**2.03e+01** △	2.03e+01 △	2.04e+01 ○	2.04e+01 ○	2.04e+01 ○
9	9.73e-01	1.99e-01 ○	**4.18e-05** △	3.63e+00 ▼	2.10e+00 ▼	4.80e+00 ▼	2.60e+00 ▼	3.79e+00 ▼	6.73e+00 ▼
10	**0.00e+00**	**0.00e+00** ○	1.00e-08 ○	9.86e-03 ▼	1.00e-01 ▼	3.00e-01 ▼	3.69e-02 ▼	3.25e-01 ▼	1.48e-02 ▼
11	1.99e-04	**0.00e+00** ○	1.30e-07 ○	2.98e+00 ▼	4.80e+00 ▼	1.09e+01 ▼	**0.00e+00** △	9.95e-01 ▼	6.30e+00 ▼
12	9.95e-01	**0.00e+00** △	1.00e-08 △	1.29e+01 ▼	4.40e+00 ▼	1.39e+01 ▼	5.97e+00 ▼	1.19e+01 ▼	1.80e+01 ▼
13	1.99e+00	**0.00e+00** △	1.00e-08 △	2.64e+01 ▼	5.50e+00 ▼	2.08e+01 ▼	8.52e+00 ▼	2.31e+01 ▼	3.55e+01 ▼
14	**8.30e+00**	1.85e+01 ○	1.19e+01 ▼	1.13e+02 ▼	6.38e+02 ▼	8.34e+02 ▼	1.86e+01 ○	2.06e+02 ▼	2.18e+02 ▼
15	1.32e+02	1.85e+01 △	**1.20e+01** △	6.33e+02 ▼	5.20e+02 ▼	7.74e+02 ▼	8.51e+02 ▼	5.48e+02 ▼	1.00e+03 ▼
16	1.19e+00	1.12e+00 ○	2.36e-01 △	**1.76e-01** △	7.00e-01 △	5.00e-01 △	1.34e+00 ▼	1.13e+00 ○	2.92e-01 △
17	1.78e+01	**1.10e+01** △	1.12e+01 △	1.18e+01 △	1.73e+01 ○	1.89e+01 ○	1.11e+01 △	1.13e+01 △	1.58e+01 △
18	2.56e+01	**1.10e+01** △	1.12e+01 △	2.95e+01 ▼	1.83e+01 △	1.78e+01 △	1.74e+01 △	1.95e+01 △	4.17e+01 ▼
19	8.26e-01	6.46e-01 △	7.09e-01 △	7.61e-01 ○	8.00e-01 ○	9.00e-01 ○	**5.11e-01** △	5.81e-01 △	7.33e-01 △
20	**2.30e+00**	3.02e+00 ▼	2.84e+00 ▼	4.25e+00 ▼	2.50e+00 ○	3.40e+00 ▼	3.21e+00 ▼	2.96e+00 ▼	4.11e+00 ▼
21	4.00e+02	4.00e+02 △	2.01e+02 △	**2.00e+02** △	4.00e+02 ○	4.00e+02 ○	3.00e+02 △	3.00e+02 △	2.86e+02 △
22	**4.45e+01**	5.65e+01 △	1.46e+02 ○	2.68e+02 ▼	4.68e+02 ▼	9.06e+02 ▼	7.51e+01 △	2.44e+02 ▼	3.34e+02 ▼
23	4.43e+01	5.92e+01 ▼	**3.66e+01** △	8.34e+02 ▼	3.45e+02 ▼	9.10e+02 ▼	9.09e+02 ▼	6.67e+02 ▼	1.52e+03 ▼
24	2.00e+02	2.09e+02 ▼	1.19e+02 △	**1.18e+02** △	2.01e+02 ○	2.14e+02 ▼	2.14e+02 ▼	2.11e+02 ▼	1.93e+02 ○
25	**2.00e+02**	2.04e+02 ▼	**2.00e+02** △	2.07e+02 ○	2.01e+02 ▼	2.09e+02 ▼	2.19e+02 ▼	2.04e+02 ▼	2.15e+02 ▼
26	**1.00e+02**	2.06e+02 ▼	1.16e+02 ▼	2.00e+02 ▼	1.08e+02 ○	2.00e+02 ▼	2.00e+02 ▼	1.15e+02 ▼	1.67e+02 ▼
27	**3.00e+02**	4.47e+02 ▼	**3.00e+02** ▼	3.08e+02 ▼	3.01e+02 ○	3.36e+02 ▼	4.15e+02 ▼	3.35e+02 ▼	3.81e+02 ▼
28	3.24e+02	3.00e+02 △	3.00e+02 △	3.00e+02 △	3.00e+02 △	3.00e+02 △	3.00e+02 △	3.00e+02 ○	**2.92e+02** △

VII. CONCLUSIONS

In this paper we deal with parallel copula EDA (Estimation of Distribution Algorithm) for optimization in continuous domain. We proposed eight variants of copula EDA using the concept of island-based algorithm. We examined the influence of two variants of migration (migration of model or migration of individuals) and two variant of copulas (Gaussian or Clayton) and two marginal distributions (normal pdf or normal pdf modified by iCTAVS component). To discover the strong points and weaknesses of them we used the traditional benchmarks (Ackley's, Rastrigin's, Sphere, Schwefel's, Rosenbrock's, Griewank's) and the new hard set CEC 2013 benchmarks.

The most important parameter is the selection of proper copula. From the mutual comparison of the proposed variants of our algorithm it follows that the utilization of Gaussian copula provides significantly better results compared to Clayton copula.

The effect of iCTAVS variance adaptation is ambiguous. Especially in combination with Gaussian copula and migration of individuals this component provides good results for some benchmarks (in the case of Rosenbrock's problem GC-EDA-mi-avs is significantly better than the other algorithms). But in many cases iCTAVS is distinctly less satisfactory.

In generally, the best performance is evident in the case of *Gaussian copula EDA with the migration of model* (GC-EDA-mm). This algorithm was compared with other algorithms on the complex benchmarks CEC 2013 and the statistical significance of the comparison was confirmed utilizing two-sample t-test.

The future research will be focused on increasing the efficiency of the model migration and incorporating more efficient techniques preventing the premature convergence.

ACKNOWLEDGMENT

This work was supported by Brno University of Technology under number FIT-S-14-2297.

REFERENCES

[1] M. Pelikan and H. Mühlenbein, "Marginal distributions in evolutionary algorithms," in *Proceedings of the International Conference on Genetic Algorithms Mendel 98*, 1999, pp. 90–95.

[2] ——, "The bivariate marginal distribution algorithm," in *Advances in Soft Computing*. Springer London, 1999, pp. 521–535.

[3] J. S. De Bonet, C. L. Isbell, and P. A. Viola, "MIMIC: Finding optima by estimating probability densities," in *Advances in Neural Information Processing Systems*, vol. 9. The MIT Press, Cambridge, 1997, pp. 424–430.

[4] M. Pelikan, D. Goldberg, and E. Cantú-Paz, "BOA: The bayesian optimization algorithm," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-99)*, vol. I, 1999, pp. 525–532 also IlliGAL Report no. 99003.

[5] P. Larrañaga and J. A. Lozano, *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Norwell, MA, USA: Kluwer Academic Publishers, 2001.

[6] M. Hauschild and M. Pelikan, "An introduction and survey of estimation of distribution algorithms," *Swarm and Evolutionary Computation*, vol. 1, no. 3, pp. 111 – 128, 2011.

TABLE III

RESULTS OF **GC-EDA-MM** FOR CEC 2013 BENCHMARK PROBLEMS IN **10 DIMENSIONS** AFTER 100000 FITNESS EVALUATION.

No.	Best	Median	Worst	Average	Std.dev
1	0.00e+00	0.00e+00	3.03e-01	9.58e-03	4.43e-02
2	5.29e+02	3.84e+03	3.10e+04	5.70e+03	5.28e+03
3	9.73e+01	3.42e+06	1.12e+08	1.04e+07	1.99e+07
4	3.33e+02	1.41e+03	4.50e+03	1.78e+03	9.97e+02
5	0.00e+00	0.00e+00	2.11e+00	1.14e-01	3.74e-01
6	0.00e+00	3.18e-01	7.94e+00	5.41e-01	1.11e+00
7	7.65e-04	2.27e-02	2.60e-01	4.67e-02	5.26e-02
8	2.02e+01	2.04e+01	2.05e+01	2.04e+01	6.72e-02
9	1.19e-03	9.73e-01	2.44e+00	7.48e-01	6.05e-01
10	0.00e+00	0.00e+00	5.67e-02	2.49e-03	9.30e-03
11	0.00e+00	1.99e-04	2.98e+00	5.49e-01	7.41e-01
12	0.00e+00	9.95e-01	2.98e+00	8.80e-01	7.80e-01
13	0.00e+00	1.99e+00	5.32e+00	1.62e+00	1.04e+00
14	3.12e-01	8.30e+00	1.18e+02	1.80e+01	2.52e+01
15	6.00e-01	1.32e+02	3.90e+02	1.30e+02	8.95e+01
16	5.69e-01	1.19e+00	1.41e+00	1.14e+00	1.79e-01
17	1.28e+01	1.78e+01	2.38e+01	1.79e+01	2.69e+00
18	2.06e+01	2.56e+01	3.27e+01	2.60e+01	2.82e+00
19	6.51e-01	8.26e-01	1.25e+00	8.51e-01	1.33e-01
20	1.84e+00	2.30e+00	3.11e+00	2.35e+00	2.62e-01
21	4.00e+02	4.00e+02	4.00e+02	4.00e+02	1.39e-04
22	1.41e+01	4.45e+01	7.80e+02	1.59e+02	1.90e+02
23	1.69e+01	4.43e+01	1.70e+02	5.17e+01	2.86e+01
24	2.00e+02	2.00e+02	2.00e+02	2.00e+02	4.30e-02
25	2.00e+02	2.00e+02	2.00e+02	2.00e+02	2.10e-02
26	1.00e+02	1.00e+02	2.00e+02	1.05e+02	1.52e+01
27	3.00e+02	3.00e+02	3.01e+02	3.00e+02	1.71e-01
28	3.00e+02	3.24e+02	3.76e+02	3.28e+02	1.96e+01

TABLE IV

RESULTS OF **GC-EDA-MM** FOR CEC 2013 BENCHMARK PROBLEMS IN **30 DIMENSIONS** AFTER 300000 FITNESS EVALUATION.

No.	Best	Median	Worst	Average	Std.dev
1	1.13e+03	1.98e+03	0.00e+00	1.98e+03	4.96e+02
2	1.62e+06	4.68e+06	1.01e+07	4.97e+06	2.02e+06
3	5.67e+08	3.13e+09	1.00e+10	3.65e+09	2.23e+09
4	3.28e+04	5.58e+04	7.29e+04	5.56e+04	7.22e+03
5	1.51e+02	4.66e+02	9.13e+02	4.64e+02	1.62e+02
6	8.11e+01	1.34e+02	2.40e+02	1.36e+02	3.64e+01
7	8.43e+00	2.56e+01	5.19e+01	2.71e+01	8.92e+00
8	2.10e+01	2.12e+01	2.13e+01	2.12e+01	6.34e-02
9	7.81e+00	1.15e+01	1.45e+01	1.16e+01	1.57e+00
10	9.55e+01	1.63e+02	3.21e+02	1.75e+02	5.06e+01
11	3.31e+01	5.03e+01	7.70e+01	5.12e+01	9.42e+00
12	3.55e+01	6.58e+01	9.33e+01	6.75e+01	1.27e+01
13	5.87e+01	1.15e+02	1.63e+02	1.13e+02	1.94e+01
14	5.75e+01	4.56e+02	1.11e+03	4.56e+02	2.32e+02
15	5.65e+02	1.15e+03	1.76e+03	1.17e+03	3.13e+02
16	2.50e+00	4.59e+00	5.91e+00	4.56e+00	6.97e-01
17	3.68e+01	4.53e+01	5.46e+01	4.52e+01	4.11e+00
18	4.05e+01	4.68e+01	6.36e+01	4.79e+01	5.89e+00
19	1.98e+01	7.23e+01	3.43e+02	1.01e+02	6.99e+01
20	7.89e+00	1.13e+01	1.24e+01	1.08e+01	1.20e+00
21	4.19e+02	6.30e+02	1.03e+03	6.75e+02	1.59e+02
22	1.81e+02	3.97e+02	1.09e+03	4.21e+02	1.96e+02
23	6.39e+02	1.27e+03	2.35e+03	1.36e+03	3.67e+02
24	2.12e+02	2.28e+02	2.40e+02	2.28e+02	6.01e+00
25	2.17e+02	2.40e+02	2.76e+02	2.45e+02	1.83e+01
26	2.00e+02	3.19e+02	3.25e+02	3.05e+02	3.84e+01
27	4.32e+02	5.04e+02	6.19e+02	5.11e+02	4.39e+01
28	4.15e+02	1.04e+03	1.54e+03	9.92e+02	2.78e+02

TABLE V

RESULTS OF **GC-EDA-MM** FOR CEC 2013 BENCHMARK PROBLEMS IN **50 DIMENSIONS** AFTER 500000 FITNESS EVALUATION.

No.	Best	Median	Worst	Average	Std.dev
1	4.06e+03	6.48e+03	0.00e+00	6.39e+03	1.22e+03
2	6.58e+06	1.78e+07	3.29e+07	1.80e+07	5.29e+06
3	4.27e+09	1.23e+10	2.11e+10	1.22e+10	3.79e+09
4	5.85e+04	7.36e+04	8.99e+04	7.42e+04	7.37e+03
5	4.76e+02	9.51e+02	1.55e+03	9.52e+02	2.17e+02
6	2.17e+02	3.07e+02	3.84e+02	3.03e+02	3.01e+01
7	2.73e+01	4.30e+01	5.80e+01	4.37e+01	6.72e+00
8	2.12e+01	2.13e+01	2.14e+01	2.13e+01	5.55e-02
9	2.56e+01	3.04e+01	3.59e+01	3.06e+01	2.30e+00
10	4.38e+02	6.06e+02	9.03e+02	6.05e+02	9.68e+01
11	8.99e+01	1.38e+02	1.74e+02	1.37e+02	1.98e+01
12	1.08e+02	1.81e+02	2.23e+02	1.81e+02	2.09e+01
13	1.88e+02	2.74e+02	3.66e+02	2.76e+02	3.75e+01
14	4.84e+02	1.30e+03	2.72e+03	1.43e+03	5.21e+02
15	2.50e+03	4.08e+03	5.42e+03	4.03e+03	6.28e+02
16	3.94e+00	5.49e+00	6.88e+00	5.46e+00	5.62e-01
17	1.03e+02	1.25e+02	1.60e+02	1.25e+02	1.36e+01
18	1.09e+02	1.39e+02	1.71e+02	1.41e+02	1.36e+01
19	4.18e+02	1.47e+03	5.00e+03	1.62e+03	9.37e+02
20	1.50e+01	1.76e+01	2.10e+01	1.74e+01	9.65e-01
21	1.40e+03	2.02e+03	3.13e+03	2.16e+03	4.55e+02
22	4.86e+02	1.51e+03	3.00e+03	1.55e+03	5.71e+02
23	2.71e+03	4.47e+03	7.15e+03	4.71e+03	1.07e+03
24	2.71e+02	2.93e+02	3.12e+02	2.93e+02	8.17e+00
25	3.19e+02	3.39e+02	3.75e+02	3.41e+02	9.84e+00
26	3.53e+02	3.67e+02	3.88e+02	3.66e+02	6.83e+00
27	9.24e+02	1.04e+03	1.36e+03	1.07e+03	9.41e+01
28	9.50e+02	1.73e+03	2.79e+03	1.74e+03	3.26e+02

TABLE VI

COMPARISON (AVERAGE) WITH COPULA BAYESIAN NETWORK (CBN) [17], 100000 FITNESS EVALUATIONS.

	Rastrigin's	Ackley's	Rosenbrock's
CBN	2.39e+00	3.71e-02	1.05e+01
GC-EDA-mm	1.37e-01	8.17e-12	6.22e+00
GC-EDA-mi-avs	5.21e-01	1.06e-05	7.63e-01

TABLE VII

COMPARISON (AVERAGE) WITH: COPULA EDA (CEDA), COPULA EDA OF DYNAMIC K-S TEST (CEDA-KS) [39]; CLAYTON, GUMBEL, SN-EDA [40], 300000 FITNESS EVALUATIONS.

	Sphere	Rastrigin's	Rosenbrock's
cEDA	4.62e-08	6.45e-08	6.52e+00
cEDA-KS	1.16e-08	2.60e-08	7.05e+00
Clayton	1.45e-07	7.00e-08	8.36e+00
Gumbel	3.59e-09	5.49e-09	6.62e+00
Sn-EDA	1.22e-09	9.52e-09	6.54e+00
GC-EDA-mm	7.59e-65	1.37e-01	6.07e+00
GC-EDA-mi-avs	2.16e-33	5.10e-01	3.84e-24

TABLE VIII

COMPARISON (AVERAGE) WITH [41], 300000 FITNESS EVALUATIONS.

	Sphere	Rosenbr.	Schwefel's	Griewank's
UMDAc	2.54e-23	7.56e+00	1.81e-02	4.26e-03
MIMICc	6.88e-24	6.54e+00	1.88e-03	1.90e-02
cEDA(empirical)	2.12e-10	6.78e+00	1.14e-04	3.41e-02
cEDA(normal)	1.17e-14	6.09e+00	1.14e-07	1.69e-10
cEDA(adaptive)	8.01e-13	4.24e+00	4.40e-07	4.25e-13
GC-EDA-mm	7.59e-65	6.07e+00	2.40e+02	0.00e+00
GC-EDA-mi-avs	2.16e-33	3.84e-24	1.53e+03	0.00e+00

[7] J. Mai and M. Scherer, *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, ser. Series in quantitative finance. Imperial College Press, 2012, vol. 4.

[8] R. B. Nelsen, *An Introduction to Copulas*, ser. Springer Series in Statistics. Springer New York, 2006.

[9] U. Cherubini, E. Luciano, and W. Vecchiato, *Copula Methods in Finance*. Hoboken, NJ: John Wiley & Sons, 2004.

[10] K. Aas, C. Czado, A. Frigessi, and H. Bakken, "Pair-copula constructions of multiple dependence," *Insurance: Mathematics and Economics*, vol. 44, no. 2, pp. 182 – 198, 2009.

[11] M. Rey and V. Roth, "Copula mixture model for dependency-seeking clustering," in *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, J. Langford and J. Pineau, Eds. New York, NY, USA: ACM, 2012, pp. 927–934. [Online]. Available: http://icml.cc/2012/papers/486.pdf

[12] B. Póczos, Z. Ghahramani, and J. Schneider, "Copula-based kernel dependency measures," in *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, J. Langford and J. Pineau, Eds. New York, NY, USA: ACM, 2012, pp. 775–782. [Online]. Available: http://icml.cc/2012/papers/417.pdf

[13] L.-F. Wang, J.-C. Zeng, and Y. Hong, "Estimation of distribution algorithm based on copula theory," in *Evolutionary Computation, 2009. CEC '09. IEEE Congress on*, May 2009, pp. 1057–1063.

[14] L.-F. Wang, J.-C. Zeng, Y. Hong, and X. Guo, "Copula estimation of distribution algorithm sampling from clayton copula," *Journal of Computational Information Systems*, vol. 6, no. 7, pp. 2431–2440, 2010.

[15] L.-F. Wang, X. Guo, J.-C. Zeng, and Y. Hong, "Using gumbel copula and empirical marginal distribution in estimation of distribution algorithm," in *Advanced Computational Intelligence (IWACI), 2010 Third International Workshop on*. IEEE, 2010, pp. 583–587.

[16] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. R. Villa-Diharce, "Using copulas in estimation of distribution algorithms," in *MICAI 2009: Advances in Artificial Intelligence*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2009, vol. 5845, pp. 658–668.

[17] M. Méndez and R. Landa, "An EDA based on bayesian networks constructed with archimedean copulas," in *2012 Fourth World Congress on Nature and Biologically Inspired Computing (NaBIC)*, 2012, pp. 188–193.

[18] R. Salinas-Gutiérrez, A. Hernández-Aguirre, and E. R. Villa-Diharce, "Estimation of distribution algorithms based on copula functions," in *Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation*, ser. GECCO '11. New York, NY, USA: ACM, 2011, pp. 795–798.

[19] M. Soto, Y. González-Fernández, and A. Ochoa, "Modeling with copulas and vines in estimation of distribution algorithms," *CoRR*, vol. abs/1210.5500, 2012.

[20] Y. Gonzalez-Fernandez and M. Soto, "A survey of estimation of distribution algorithms based on copulas," Tech. Rep., 2012, technical Report ICIMAF 2012-679.

[21] M. Hyrš and J. Schwarz, "Elliptical and archimedean copulas in estimation of distribution algorithm with model migration," in *Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015)*. SciTePress - Science and Technology Publications, 2015, pp. 212–219.

[22] J. Grahl, P. A. Bosman, and F. Rothlauf, "The correlation-triggered adaptive variance scaling idea," in *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '06. New York, NY, USA: ACM, 2006, pp. 397–404. [Online]. Available: http://doi.acm.org/10.1145/1143997.1144071

[23] A. Sklar, "Fonctions de répartition à n dimensions et leurs marges," *Publications de l'Institut de Statistique de l'Université de Paris*, vol. 8, pp. 229–231, 1959.

[24] K. Michalak, *Sim-EDA: A Multipopulation Estimation of Distribution Algorithm Based on Problem Similarity*. Cham: Springer International Publishing, 2016, pp. 235–250.

[25] J. Madera, E. Alba, and A. Ochoa, *A Parallel Island Model for Estimation of Distribution Algorithms*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 159–186.

[26] L. delaOssa, J. A. Gámez, and J. M. Puerta, "Migration of probability models instead of individuals: An alternative when applying the island model to edas," in *Parallel Problem Solving from Nature - PPSN VIII*, ser. Lecture Notes in Computer Science. Springer, 2004, vol. 3242 of LNCS, pp. 242–252.

[27] J. Schwarz and J. Jaroš, "Parallel bivariate marginal distribution algorithm with probability model migration," in *Linkage in Evolutionary Computation*, ser. Studies in Computational Intelligence. Springer Berlin Heidelberg, 2008, vol. 157, pp. 3–23.

[28] M. El-Abd and M. S. Kamel, "A cooperative particle swarm optimizer with migration of heterogeneous probabilistic models," *Swarm Intelligence*, vol. 4, no. 1, pp. 57–89, 2010.

[29] S. Frühwirth-Schnatter, *Finite Mixture and Markov Switching Models*. New York: Springer, 2006.

[30] J. Liang, B. Qu, P. Suganthan, and A. G. Hernández-Díaz, "Problem definitions and evaluation criteria for the cec 2013 special session on real-parameter optimization," *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore, Technical Report*, vol. 201212, 2013.

[31] I. Loshchilov, "Cma-es with restarts for solving cec 2013 benchmark problems," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 369–376.

[32] T. Liao and T. Sttzle, "Benchmark results for a simple hybrid algorithm on the cec 2013 benchmark set for real-parameter optimization," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 1938–1944.

[33] F. Caraffini, G. Iacca, F. Neri, L. Picinali, and E. Mininno, "A cma-es super-fit scheme for the re-sampled inheritance search," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 1123–1130.

[34] R. Ugolotti and S. Cagnoni, *A Fair Comparison Between Standard PSO Versions*. Cham: Springer International Publishing, 2016, pp. 3–14.

[35] M. Zambrano-Bigiarini, M. Clerc, and R. Rojas, "Standard particle swarm optimisation 2011 at cec-2013: A baseline for future pso improvements," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 2337–2344.

[36] S. M. Elsayed, R. A. Sarker, and D. L. Essam, "A genetic algorithm for solving the cec'2013 competition problems on real-parameter optimization," in *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 356–360.

[37] S. Chen and Y. Gonzalez-Fernandez, "Leaders and followers on the cec2013 real-parameter optimization benchmark functionss," Tech. Rep., 2015, technical Report School of Information Technology York University.

[38] M. G. Epitropakis, F. Caraffini, F. Neri, and E. K. Burke, "A separability prototype for automatic memes with adaptive operator selection," in *Foundations of Computational Intelligence (FOCI), 2014 IEEE Symposium on*, Dec 2014, pp. 70–77.

[39] H. Zhao and L. Wang, "Marginal distribution in copula estimation of distribution algorithm based dynamic K-S test," in *IJCSI International Journal of Computer Science Issues*, no. 3, 2012, pp. 507–514.

[40] B. Jia, L. Wang, and Z. Cui, "Copula for estimation of distribution algorithm based on goodness-of-fit test," in *Journal of Theoretical and Applied Information Technology*, no. 3, 2013, pp. 1128–1132.

[41] R. Yang, L.-F. Wang, and J.-C. Zeng, *Adaptive Variance Scaling in Clayton Copula EDA*. Dordrecht: Springer Netherlands, 2012, pp. 1295–1302.