Supplementary File of ‘Efficient Resource Allocation in Cooperative Co-evolution for Large-scale Global Optimization’

Ming Yang, Mohammad Nabi Omidvar, Changhe Li, *Member, IEEE*, Zhihua Cai, Xiaodong Li, *Senior Member, IEEE*, Borhan Kazimipour, and Xin Yao, *Fellow, IEEE*

CONTENTS

The work was supported in part by the National Natural Science Foundation of China (Grant Nos. 61305086, 61203306, 61329302 and 61305079) and EPSRC (Grant Nos. EP/I010297/1 and EP/K001523/1).

M. Yang, C. Li and Z. Cai are with the School of Computer Science, China University of Geosciences, Wuhan, 430074, China (e-mail: yangming0702@gmail.com, changhe.lw@gmail.com, zhcai@cug.edu.cn).

M. N. Omidvar and X. Yao are with the Centre of Excellence for Research in Computational Intelligence and Applications, School of Computer Science, University of Birmingham, Birmingham B15 2TT, U.K. (e-mail: m.omidvar@cs.bham.ac.uk, x.yao@cs.bham.ac.uk).

X. Li and B. Kazimipour are with the School of Computer Science and Information Technology, RMIT University, Melbourne, VIC 3001, Australia (e-mail: xiaodong.li@rmit.edu.au, borhan.kazimipour@rmit.edu.au).

TABLE I: The average fitness values ± standard deviations on the CEC'2010 and the CEC'2013 functions over 25 independent runs. The significant best results are in bold font (Wilcoxon rank sum test with Holm p-value correction, α=0.05). R^+, R^- and p-value are obtained through multiple-problem analysis by the Wilcoxon test between CCFR-I (U=D_i) and its competitors.

CEC'2010 Functions			
F	CCFR-I ($U = D_i$)	CCFR-I ($U = 2D_i$)	CCFR-I ($U = 10D_i$)
f_1	1.20e-05±4.89e-06	1.31e-05±5.19e-06	1.68e-05±6.54e-06↑
f_2	**2.75e+01±5.25e+00**	5.13e+01±5.04e+00↑	1.52e+02±7.22e+00↑
f_3	**4.56e+00±4.63e-01**	5.56e+00±4.63e-01↑	8.10e+00±4.65e-01↑
f_4	8.33e+10±6.16e+10	8.69e+10±4.68e+10	1.06e+11±4.31e+10↑
f_5	7.23e+07±1.32e+07	7.32e+07±1.22e+07	9.12e+07±1.74e+07↑
f_6	7.74e+05±7.15e+05	7.83e+05±8.28e+05	7.28e+05±8.51e+05
f_7	**1.49e-03±2.47e-04**	1.66e-03±2.78e-04↑	2.14e-03±3.90e-04↑
f_8	3.19e+05±1.08e+06	6.38e+05±1.46e+06	9.57e+05±1.70e+06↑
f_9	9.38e+06±1.18e+06	8.81e+06±1.05e+06	1.05e+07±1.44e+06↑
f_{10}	1.41e+03±1.01e+02	1.42e+03±7.83e+01	1.61e+03±1.10e+02↑
f_{11}	1.03e+01±2.71e+00	9.72e+00±2.11e+00	1.00e+01±2.59e+00
f_{12}	**1.17e+00±4.57e+00**	4.72e+00±1.75e+01↑	7.49e+00±2.30e+01↑
f_{13}	3.18e+02±9.91e+01	3.25e+02±1.01e+02	4.03e+02±9.45e+01↑
f_{14}	2.48e+07±2.85e+06	2.48e+07±2.85e+06	2.48e+07±2.85e+06
f_{15}	2.81e+03±1.31e+02	2.81e+03±1.31e+02	2.81e+03±1.31e+02
f_{16}	2.01e+01±2.62e+00	2.01e+01±2.62e+00	2.01e+01±2.62e+00
f_{17}	9.78e+00±1.09e+01	9.78e+00±1.09e+01	9.78e+00±1.09e+01
f_{18}	1.14e+03±1.82e+02	1.14e+03±1.82e+02	1.14e+03±1.82e+02
f_{19}	1.16e+06±9.47e+04	1.16e+06±9.47e+04	1.16e+06±9.47e+04
f_{20}	1.01e+09±8.96e+08	1.01e+09±8.96e+08	1.01e+09±8.96e+08
R^+	—	168.0	170.0
R^-	—	42.0	40.0
p-value	—	2.66e-02	1.71e-02
CEC'2013 Functions			
F	CCFR-I ($U = D_i$)	CCFR-I ($U = 2D_i$)	CCFR-I ($U = 10D_i$)
f_1	1.30e-05±3.18e-06	1.40e-05±3.49e-06	1.80e-05±4.65e-06↑
f_2	**5.51e-01±1.47e+00**	5.33e+01±1.70e+01↑	3.14e+02±2.05e+01↑
f_3	2.00e+01±3.06e-07	**2.00e+01±3.23e-07**↓	2.00e+01±3.89e-04↑
f_4	4.50e+07±1.66e+07	5.26e+07±2.22e+07	7.47e+07±2.31e+07↑
f_5	2.53e+06±2.67e+05	2.47e+06±3.75e+05	2.62e+06±3.88e+05
f_6	1.06e+06±1.19e+03	**1.06e+06±1.30e+03**↓	1.07e+06±1.64e+03↑
f_7	8.60e+06±1.90e+07	9.94e+06±2.64e+07	1.04e+07±1.85e+07
f_8	9.61e+09±1.59e+10	9.61e+09±1.59e+10	9.61e+09±1.59e+10
f_9	1.85e+08±2.79e+07	1.84e+08±2.70e+07	1.84e+08±2.73e+07
f_{10}	9.47e+07±1.86e+05	9.46e+07±3.84e+05	**9.43e+07±3.44e+05**↓
f_{11}	3.25e+08±3.24e+08	2.53e+08±3.33e+08	3.28e+08±3.38e+08
f_{12}	6.00e+08±7.09e+08	6.00e+08±7.09e+08	6.00e+08±7.09e+08
f_{13}	9.28e+08±5.33e+08	9.28e+08±5.33e+08	9.28e+08±5.33e+08
f_{14}	2.14e+09±2.11e+09	2.14e+09±2.11e+09	2.14e+09±2.11e+09
f_{15}	8.25e+06±3.28e+06	8.25e+06±3.28e+06	8.25e+06±3.28e+06
R^+	—	49.5	89.5
R^-	—	70.5	30.5
p-value	—	6.25e-01	1.60e-01

The symbols ↑ and ↓ denote that the CCFR-I ($U = D_i$) algorithm performs significantly better than and worse than this algorithm by the Wilcoxon rank sum test at the significance level of 0.05, respectively.

I. The Sensitivity Study of the Parameter U of CCFR

Table I summarizes the results of CCFR-I with different values of the parameter U (see Eq. (6a) in the paper) on the CEC'2010 and the CEC'2013 large-scale functions [1], [2]. D_i is the dimensionality of a subcomponent.

For the functions with separable variables (i.e., the CEC'2010 functions f_1–f_{13} and the CEC'2013 functions f_1–f_7), the smaller value of U, the better performance of CCFR in general. This is because CCFR with a small value of U can stop early evolution for the stagnant subpopulations. It can save more computational resources on the separable and stagnant variables than CCFR with a larger value of U. Therefore, we use $U = D_i$ as the default setting of U. For the functions without separable variables, the subpopulations hardly enter the stagnant state, so there is no difference between CCFR-I with different values of U. Overall, CCFR-I with different values of U had similar performances on most of the CEC'2010 and the CEC'2013 functions.

II. The Scale-up Study of CCFR

We use the block-rotated ellipsoid function [3] to study the performances of CCFR-I, CBCC1-I, CBCC2-I and CC-I with the scale-up dimensionality of the function and the number of the subcomponents. The dimensionality of the function ranges

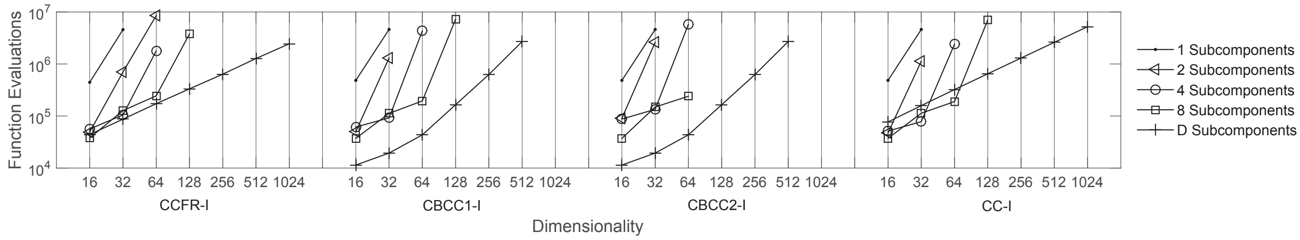


Fig. 1: For the block-rotated ellipsoid function with variant numbers of subcomponents, the average function evaluations used by CCFR-I, CBCC1-I, CBCC2-I and CC-I over successful runs out of 10 runs.

from 2^4 to 2^{10}. The numbers of the subcomponents are $\{1,2,4,8,D\}$, where D is the dimensionality. Within 10^7 function evaluations, if the best overall objective value is smaller than a target value (i.e., 0.1) in a run, CCFR-I stops running and this run is considered successful. Fig. 1 shows the average number of function evaluations over successful runs out of 10 runs. CCFR-I can reach the target value within 10^7 function evaluations when there are less than 64 variables in a subcomponent. When the number of the variables in a subcomponent is equal to or smaller than 8, the number of function evaluations linearly increases as the dimensionality of the function and the number of the subcomponents increase. When there are more than 8 variables in a subcomponent, the number of function evaluations increases rapidly and linearly as the dimensionality of the function and the number of the subcomponents increase. It can be seen in Fig. 1 that CBCC1-I, CBCC2-I and CC-I have similar performances to CCFR-I, but for CCFR-I, as the dimensionality of the function and the number of the subcomponents increase, the number of function evaluations increases less rapidly than the other three CC algorithms.

III. The Performance of CCFR with Groupings

In order to study the effect of decomposition on the performance of CCFR, we test CCFR with two grouping methods (DG [4] and IDG2 [5]). DG is a differential grouping method with a theoretical foundation, which is able to group the interacting variables with a high accuracy. In DG, the parameter ϵ was set to 10^{-3}, which was recommended in [4]. IDG2 is an improved variant of DG, which is able to group the interacting variables better than DG. Table II summarizes the grouping results of IDG2 and DG.

Table III summarizes the optimization results of CCFR, CBCC1 [6], CBCC2 [6] and DECC [4] with IDG2 and DG. Note that, for the algorithms with IDG2 and DG, the function evaluations spent by groupings (see the 'FEs' column in Table II) are counted into the entire function evaluations. The multiple-problem analysis results show that CCFR-IDG2 and CCFR-DG performed better than the other peer algorithms on the CEC'2010 and the CEC'2013 functions.

CCFR-DG performed significantly better than the other peer algorithms with DG on most of the separable functions f_1–f_3. For almost all the partially separable functions (the CEC'2010 functions f_4–f_{18}; the CEC'2013 functions f_4–f_{11}), the difference between the results of the algorithms with DG is not significant. For the CEC'2010 functions f_7, f_8 and f_{13}, DG is not able to identify the interdependence between variables. There is interdependence between the subcomponents formed by DG. CCFR-DG performed worse than CBCC1-DG and DECC-DG by several orders of magnitude. This indicates that if there is interdependence between subcomponents, it may be a good way to optimize each subcomponent one by one.

CCFR-IDG2 significantly outperformed the other peer algorithms on most of the separable functions f_1–f_3 by several orders of magnitude. CCFR-IDG2 outperformed the other peer algorithms on most of the partially separable functions (the CEC'2010 functions f_4–f_{18}; the CEC'2013 functions f_4–f_{11}). For the partially separable functions on which CCFR-IDG2 performed worse, the difference between the results of CCFR and the other peer algorithms is not significant. For the functions on which CCFR-IDG2 performed better, the difference is significant. For the nonseparable functions (the CEC'2010 functions f_{19}–f_{20}; the CEC'2013 functions f_{12}–f_{15}), all the variables are grouped into one subcomponent. Therefore, there is no significant difference between the algorithms with IDG2 on these nonseparable functions.

For most of the functions, the algorithms with IDG2 performed better than the ones with DG. This is because IDG2 can identify the interdependence between variables with higher accuracies than DG. The multiple-problem analysis results show that compared with DG, IDG2 made CCFR performed much better than the other peer algorithms. The performances of CCFR-IDG2 and CCFR-DG do not differ greatly on most of the functions where CCFR-IDG2 performed worse than CCFR-DG. For most of the functions where CCFR-IDG2 performed better than CCFR-DG, CCFR-IDG2 outperformed greatly CCFR-DG by several orders of magnitude due to its higher grouping accuracy for the nonseparable variables on these functions (i.e., the CEC'2010 functions f_7, f_8, f_{13} and f_{18}; the CEC'2013 functions f_4, f_7, f_8 and f_{11}). The experimental results show that the performance of CCFR is dependent on the decomposition method. A high grouping accuracy, especially for the nonseparable variables, can improve the performance of CCFR.

TABLE II: The grouping results on the CEC'2010 and the CEC'2013 functions. The values of IDG2 and DG are separated by "/". The bold font denotes IDG2 performed better than DG; the gray background denotes IDG2 performed worse than DG.

CEC'2010 Functions

F	Sep Vars	Non-Sep		IDG2 / DG ($\epsilon = 10^{-3}$)						
				FEs	Sep			Non-sep		
		Vars	Groups		Formed Vars	Captured Vars	Accuracy	Formed Subcomponents	Captured Subcomponents	Accuracy
f_1	1000	0	0	500501 / 1001000	1000 / 1000	1000 / 1000	100.0% / 100.0%	0 / 0	0 / 0	100.0% / 100.0%
f_2	1000	0	0	500501 / 1001000	1000 / 1000	1000 / 1000	100.0% / 100.0%	0 / 0	0 / 0	100.0% / 100.0%
f_3	1000	0	0	500501 / 1001000	0 / 1000	0 / 1000	0.0% / 100.0%	1 / 0	0 / 0	100.0% / 100.0%
f_4	950	50	1	500501 / 14554	950 / 33	950 / 33	**100.0% / 3.5%**	1 / 10	1 / 1	100.0% / 100.0%
f_5	950	50	1	500501 / 905450	950 / 950	950 / 950	100.0% / 100.0%	1 / 1	1 / 1	100.0% / 100.0%
f_6	950	50	1	500501 / 906332	854 / 950	854 / 950	89.9% / 100.0%	2 / 1	1 / 1	100.0% / 100.0%
f_7	950	50	1	500501 / 67742	950 / 248	950 / 248	**100.0% / 26.1%**	1 / 4	1 / 0	**100.0% / 0.0%**
f_8	950	50	1	500501 / 23286	950 / 134	950 / 133	**100.0% / 14.0%**	1 / 5	1 / 0	**100.0% / 0.0%**
f_9	500	500	10	500501 / 270802	500 / 500	500 / 500	100.0% / 100.0%	10 / 10	10 / 10	100.0% / 100.0%
f_{10}	500	500	10	500501 / 272958	500 / 500	500 / 500	100.0% / 100.0%	10 / 10	10 / 10	100.0% / 100.0%
f_{11}	500	500	10	500501 / 270640	0 / 501	0 / 500	0.0% / 100.0%	11 / 10	10 / 9	**100.0% / 90.0%**
f_{12}	500	500	10	500501 / 271390	500 / 500	500 / 500	100.0% / 100.0%	10 / 10	10 / 10	100.0% / 100.0%
f_{13}	500	500	10	500501 / 50328	500 / 131	500 / 107	**100.0% / 21.4%**	10 / 34	10 / 0	**100.0% / 0.0%**
f_{14}	0	1000	20	500501 / 21000	0 / 0	0 / 0	100.0% / 100.0%	20 / 20	20 / 20	100.0% / 100.0%
f_{15}	0	1000	20	500501 / 21000	0 / 0	0 / 0	100.0% / 100.0%	20 / 20	20 / 20	100.0% / 100.0%
f_{16}	0	1000	20	500501 / 21128	0 / 4	0 / 0	100.0% / 100.0%	20 / 20	20 / 16	**100.0% / 80.0%**
f_{17}	0	1000	20	500501 / 21000	0 / 0	0 / 0	100.0% / 100.0%	20 / 20	20 / 20	100.0% / 100.0%
f_{18}	0	1000	20	500501 / 39624	0 / 78	0 / 0	100.0% / 100.0%	20 / 50	20 / 0	**100.0% / 0.0%**
f_{19}	0	1000	1	500501 / 2000	0 / 0	0 / 0	100.0% / 100.0%	1 / 1	1 / 1	100.0% / 100.0%
f_{20}	0	1000	1	500501 / 155430	0 / 33	0 / 0	100.0% / 100.0%	1 / 241	1 / 0	**100.0% / 0.0%**

CEC'2013 Functions

F	Sep Vars	Non-Sep		IDG2 / DG ($\epsilon = 10^{-3}$)						
				FEs	Sep			Non-sep		
		Vars	Groups		Formed Vars	Captured Vars	Accuracy	Formed Subcomponents	Captured Subcomponents	Accuracy
f_1	1000	0	0	500501 / 1001000	1000 / 1000	1000 / 1000	100.0% / 100.0%	0 / 0	0 / 0	100.0% / 100.0%
f_2	1000	0	0	500501 / 1001000	1000 / 1000	1000 / 1000	100.0% / 100.0%	0 / 0	0 / 0	100.0% / 100.0%
f_3	1000	0	0	500501 / 1001000	0 / 1000	0 / 1000	0.0% / 100.0%	1 / 0	0 / 0	100.0% / 100.0%
f_4	700	300	7	500501 / 15792	700 / 40	700 / 40	**100.0% / 5.7%**	7 / 13	7 / 3	**100.0% / 58.3%**
f_5	700	300	7	500501 / 527026	700 / 707	700 / 700	100.0% / 100.0%	7 / 10	7 / 6	**100.0% / 66.7%**
f_6	700	300	7	500501 / 579848	0 / 752	0 / 700	0.0% / 100.0%	8 / 5	7 / 3	**100.0% / 50.0%**
f_7	700	300	7	500501 / 11452	700 / 64	700 / 64	**100.0% / 9.1%**	7 / 10	7 / 0	**100.0% / 0.0%**
f_8	0	1000	20	500501 / 22682	200 / 4	0 / 0	100.0% / 100.0%	18 / 25	18 / 14	**80.0% / 65.0%**
f_9	0	1000	20	500501 / 17650	0 / 0	0 / 0	100.0% / 100.0%	20 / 20	20 / 20	100.0% / 100.0%
f_{10}	0	1000	20	500501 / 48650	0 / 152	0 / 0	100.0% / 100.0%	20 / 18	20 / 14	**100.0% / 65.0%**
f_{11}	0	1000	20	500501 / 9102	0 / 1	0 / 0	100.0% / 100.0%	20 / 18	20 / 0	**100.0% / 0.0%**
f_{12}	0	1000	1	500501 / 149894	0 / 50	0 / 0	100.0% / 100.0%	1 / 222	1 / 0	**100.0% / 0.0%**
f_{13}	0	905	1	409966 / 18786	0 / 0	0 / 0	100.0% / 100.0%	1 / 20	1 / 0	**100.0% / 0.0%**
f_{14}	0	905	1	409966 / 26698	0 / 0	0 / 0	100.0% / 100.0%	1 / 19	1 / 0	**100.0% / 0.0%**
f_{15}	0	1000	1	500501 / 2000	0 / 0	0 / 0	100.0% / 100.0%	1 / 1	1 / 1	100.0% / 100.0%

IV. Comparison Between CCFR-IDG2 and Non-CC Algorithms

Table IV summarizes the results of CCFR-IDG2, MA-SW-Chains [7] and MOS-CEC2013 [8]. MA-SW-Chains and MOS-CEC2013 were ranked the first in the IEEE CEC'2010 and the IEEE CEC'2013 competitions on large-scale global optimization, respectively. For the partially separable functions (the CEC'2010 functions f_4–f_{18}; the CEC'2013 functions f_4–f_{11}) on which CCFR-IDG2 performed better than MA-SW-Chains, the difference between the results of CCFR-IDG2 and MA-SW-Chains is very significant. For the partially separable functions on which CCFR-IDG2 performed worse than MA-SW-Chains, the difference is not significant except for the CEC'2010 function f_{12}. CCFR-IDG2 performed worse than MOS-CEC2013 on most of the CEC'2010 and the CEC'2013 functions. For the nonseparable functions (the CEC'2010 functions f_{19}–f_{20}; the CEC'2013 functions f_{12}–f_{15}), CCFR-IDG2 optimized all the decision variables together and performed significantly worse than MA-SW-Chains and MOS-CEC2013. This indicates that the optimizer used by CCFR-IDG2 (i.e., SaNSDE) is inferior to MA-SW-Chains and MOS-CEC2013. The multiple-problem analysis shows that CCFR-IDG2 performed worse than MA-SW-Chains and MOS-CEC2013 on the CEC'2013 functions. This may be because that the optimizer used by CCFR-IDG2 is worse than MA-SW-Chains and MOS-CEC2013. The previous experimental results have shown that for a given optimizer (i.e., SaNSDE), CCFR is superior to the other peer algorithms.

Fig. 2 shows the convergence behavior of CCFR-IDG2, MA-SW-Chains and MOS-CEC2013. Because CCFR-IDG2 spends 500501 function evaluations grouping the decision variables, in Fig. 2 the convergence lines of CCFR-IDG2 start from 500502 function evaluations. For the separable function f_1, CCFR-IDG2 optimized each separable variable one by one and converged

TABLE III: The average fitness values ± standard deviations on the CEC'2010 and the CEC'2013 functions over 25 independent runs. The significant best results are in bold font (Wilcoxon rank sum test with Holm p-value correction, α=0.05). R^+, R^- and p-value have similar meanings as in Table I.

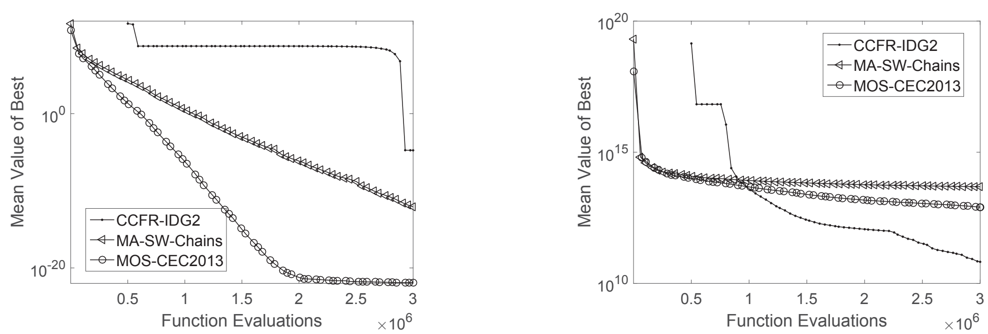
CEC'2010 Functions

F	CCFR-IDG2	CBCC1-IDG2	CBCC2-IDG2	DECC-IDG2	CCFR-DG	CBCC1-DG	CBCC2-DG	DECC-DG
f_1	**1.6e-05±6.5e-06**	1.7e+07±2.1e+07↑	1.7e+07±2.1e+07↑	1.7e+07±2.1e+07↑	4.8e+08±9.8e+07	2.9e+07±3.1e+07↓	2.9e+07±3.1e+07↓	2.9e+07±3.1e+07↓
f_2	**1.7e+02±8.6e+00**	4.7e+03±4.8e+02↑	4.7e+03±4.8e+02↑	4.7e+03±4.8e+02↑	**3.2e+02±1.7e+01**	4.7e+03±4.8e+02↑	4.7e+03±4.8e+02↑	4.7e+03±4.8e+02↑
f_3	1.2e+01±3.7e-01	1.2e+01±3.7e-01	1.2e+01±3.7e-01	1.2e+01±3.7e-01	**1.1e+01±3.8e-01**	1.2e+01±3.7e-01↑	1.2e+01±3.7e-01↑	1.2e+01±3.7e-01↑
f_4	1.3e+11±7.5e+10	7.4e+10±4.8e+10↓	1.1e+11±2.9e+10	8.9e+10±4.6e+10↓	4.3e+10±1.6e+10	3.5e+11±2.0e+11↑	5.1e+10±3.1e+10	7.8e+11±5.5e+11↑
f_5	9.2e+07±1.6e+07	6.8e+07±1.1e+07↓	6.8e+07±9.4e+06↓	6.7e+07±1.0e+07↓	4.9e+08±2.4e+07	6.9e+07±1.0e+07↓	6.9e+07±1.0e+07↓	6.9e+07±1.1e+07↓
f_6	6.8e+05±7.1e+05	1.1e+06±7.9e+05↑	1.1e+06±6.9e+05↑	6.4e+05±6.8e+05	1.1e+07±7.5e+05	1.3e+06±6.4e+05↓	1.3e+06±6.4e+05↓	**8.1e+05±7.2e+05**↓
f_7	**2.0e-03±3.5e-04**	7.9e+04±1.0e+04↑	1.1e+05±1.8e+04↑	4.2e+04±1.2e+04↑	2.7e+07±7.0e+07	1.1e+05±8.5e+04↓	7.6e+09±6.6e+09↑	**6.0e+04±3.3e+04**↓
f_8	**3.2e+05±1.1e+06**	8.8e+05±1.6e+06↑	1.1e+06±1.7e+06↑	5.2e+05±1.3e+06↑	2.6e+08±1.9e+08	**4.6e+06±8.8e+06**↓	6.3e+07±6.0e+07↓	1.5e+07±2.3e+07↓
f_9	1.3e+07±1.7e+06	2.1e+07±2.2e+07	4.4e+09±7.0e+08↑	5.4e+07±6.4e+07↑	1.1e+07±1.4e+06	1.8e+07±2.1e+07	1.8e+07±2.1e+07	3.3e+07±2.0e+07↑
f_{10}	**1.8e+03±1.4e+02**	3.4e+03±1.7e+02↑	4.6e+03±7.7e+02↑	4.3e+03±1.8e+02↑	**1.6e+03±1.2e+02**	3.2e+03±1.7e+02↑	3.2e+03±1.7e+02↑	4.1e+03±1.7e+02↑
f_{11}	**2.0e+01±3.3e+00**	2.4e+01±2.4e+00↑	2.5e+01±2.3e+00↑	2.3e+01±2.1e+00↑	**1.1e+01±2.5e+00**	2.3e+01±2.2e+00↑	2.3e+01±2.1e+00↑	2.3e+01±2.7e+00↑
f_{12}	**2.0e+01±2.2e+01**	2.6e+04±7.4e+03↑	3.7e+04±9.7e+03↑	2.3e+04±8.8e+03↑	**4.6e+00±6.9e+00**	2.2e+04±6.3e+03↑	2.2e+04±6.3e+03↑	1.9e+04±7.3e+03↑
f_{13}	**5.3e+02±1.0e+02**	2.6e+04±7.8e+03↑	3.9e+04±6.2e+03↑	2.5e+04±7.8e+03↑	2.8e+06±9.2e+05	**5.8e+03±4.4e+03**↓	1.6e+04±7.8e+03↓	8.7e+03±3.9e+03↓
f_{14}	**3.1e+07±3.3e+06**	3.5e+07±2.6e+06↑	9.5e+09±5.2e+08↑	3.3e+07±2.7e+06↑	**2.5e+07±2.9e+06**	2.8e+07±2.1e+06↑	2.8e+07±2.1e+06↑	2.7e+07±2.2e+06↑
f_{15}	**3.2e+03±1.5e+02**	4.4e+03±1.5e+02↑	4.6e+03±1.7e+02↑	4.4e+03±1.9e+02↑	**2.8e+03±1.3e+02**	4.0e+03±1.5e+02↑	4.0e+03±1.5e+02↑	4.0e+03±1.6e+02↑
f_{16}	2.0e+01±2.6e+00	1.9e+01±3.2e+00	2.0e+01±3.4e+00	2.0e+01±4.0e+00	2.4e+01±4.3e+00	2.0e+01±3.4e+00↓	2.1e+01±3.1e+00	2.1e+01±3.4e+00
f_{17}	**6.7e+01±8.7e+01**	1.3e+02±8.9e+01↑	7.2e+02±3.4e+02↑	8.0e+01±5.2e+01↑	1.1e+01±1.1e+01	3.6e+01±4.9e+01↑	3.6e+01±4.9e+01↑	2.4e+01±3.7e+01
f_{18}	1.4e+03±1.9e+02	1.3e+03±1.9e+02	1.7e+03±2.4e+02↑	1.2e+03±1.5e+02↓	**1.3e+08±9.9e+07**	6.9e+09±2.3e+09↑	1.4e+10±2.0e+09↑	2.1e+10±3.9e+09↑
f_{19}	1.3e+06±1.0e+05	1.3e+06±1.0e+05	1.3e+06±1.0e+05	1.3e+06±1.0e+05	1.2e+06±9.5e+04	1.2e+06±9.5e+04	1.2e+06±9.5e+04	1.2e+06±9.5e+04
f_{20}	2.0e+09±1.8e+09	2.0e+09±1.8e+09	2.0e+09±1.8e+09	2.0e+09±1.8e+09	**3.1e+07±6.6e+06**	1.4e+10±2.7e+09↑	1.6e+08±1.5e+08↑	3.3e+10±5.9e+09↑
R^+	—	165.0	174.0	153.0	—	123.0	137.0	123.0
R^-	—	45.0	36.0	57.0	—	87.0	73.0	87.0
p-value	—	2.51e-02	1.00e-02	7.31e-02	—	5.02e-01	2.32e-01	5.02e-01

CEC'2013 Functions

F	CCFR-IDG2	CBCC1-IDG2	CBCC2-IDG2	DECC-IDG2	CCFR-DG	CBCC1-DG	CBCC2-DG	DECC-DG
f_1	**1.8e-05±4.5e-06**	4.6e+07±1.3e+08↑	4.6e+07±1.3e+08↑	4.6e+07±1.3e+08↑	4.8e+08±6.9e+07	6.2e+07±1.3e+08↓	6.2e+07±1.3e+08↓	6.2e+07±1.3e+08↓
f_2	**3.6e+02±2.1e+01**	2.1e+04±1.0e+03↑	2.1e+04±1.0e+03↑	2.1e+04±1.0e+03↑	**7.4e+02±4.0e+01**	2.1e+04±1.0e+03↑	2.1e+04±1.0e+03↑	2.1e+04±1.0e+03↑
f_3	2.1e+01±1.2e-02	2.1e+01±1.2e-02	2.1e+01±1.2e-02	2.1e+01±1.2e-02	**2.0e+01±6.0e-07**	2.1e+01±1.1e-02↑	2.1e+01±1.1e-02↑	2.1e+01±1.1e-02↑
f_4	**9.6e+07±4.0e+07**	2.2e+08±6.0e+07↑	6.6e+10±5.6e+09↑	2.9e+08±9.7e+07↑	9.1e+10±5.6e+10	8.7e+10±5.1e+10	4.6e+11±2.8e+11↑	8.3e+10±4.7e+10
f_5	2.8e+06±3.2e+05	2.6e+06±4.3e+05	2.5e+06±4.7e+05↓	3.0e+06±4.7e+05	3.0e+06±5.2e+05	2.8e+06±3.6e+05	2.6e+06±4.4e+05↓	3.3e+06±4.0e+05↑
f_6	1.1e+06±1.0e+03	1.1e+06±1.7e+03↓	1.1e+06±1.8e+03↓	1.1e+06±1.6e+03↓	1.1e+06±1.6e+03	1.1e+06±2.1e+03↓	1.1e+06±1.5e+03↓	1.1e+06±2.3e+03↓
f_7	2.0e+07±2.9e+07	2.2e+07±2.6e+07	9.9e+07±3.7e+08	2.4e+07±3.8e+07	1.4e+08±9.7e+07	1.2e+08±3.9e+07	1.6e+10±1.4e+10↑	1.4e+08±7.1e+07
f_8	**6.6e+10±9.5e+10**	2.3e+13±1.6e+13↑	1.1e+12±1.7e+11↑	7.4e+13±5.8e+13↑	1.6e+15±1.0e+15	2.0e+15±1.5e+15	5.9e+15±4.3e+15↑	2.0e+15±1.4e+15
f_9	**1.9e+08±2.8e+07**	2.6e+08±4.0e+07↑	2.3e+08±3.0e+07↑	3.0e+08±5.7e+07↑	**1.9e+08±2.8e+07**	2.5e+08±3.8e+07↑	2.2e+08±2.9e+07↑	2.9e+08±5.2e+07↑
f_{10}	9.5e+07±1.8e+05	9.4e+07±2.8e+05↓	9.4e+07±2.5e+05↓	9.5e+07±3.0e+05↓	9.5e+07±3.1e+05	9.4e+07±6.1e+05↓	9.4e+07±6.6e+05↓	9.4e+07±2.4e+05↓
f_{11}	4.2e+08±3.4e+08	5.0e+09±1.5e+10	7.3e+10±1.2e+11↑	2.8e+09±1.1e+10	**2.8e+10±6.0e+10**	4.5e+10±6.1e+10↑	5.2e+12±3.7e+12↑	4.7e+10±5.7e+10↑
f_{12}	1.6e+09±1.6e+09	1.6e+09±1.6e+09	1.6e+09±1.6e+09	1.6e+09±1.6e+09	**8.0e+07±8.3e+06**	6.0e+10±8.3e+09↑	6.6e+08±1.3e+08↑	1.2e+11±1.4e+10↑
f_{13}	1.2e+09±6.0e+08	1.2e+09±6.0e+08	1.2e+09±6.0e+08	1.2e+09±6.0e+08	**2.0e+09±1.0e+09**	4.0e+09±1.5e+09↑	4.1e+10±2.7e+10↑	6.3e+09±1.9e+09↑
f_{14}	3.4e+09±3.1e+09	3.5e+09±3.2e+09	3.5e+09±3.2e+09	3.5e+09±3.2e+09	7.4e+09±8.5e+09	1.3e+10±1.2e+10↑	5.0e+11±1.2e+12↑	8.9e+09±6.8e+09
f_{15}	9.8e+06±3.7e+06	9.9e+06±3.7e+06	9.9e+06±3.7e+06	9.9e+06±3.7e+06	8.3e+06±3.3e+06	8.3e+06±3.3e+06	8.3e+06±3.3e+06	8.3e+06±3.3e+06
R^+	—	107.0	107.0	112.0	—	80.0	99.0	91.0
R^-	—	13.0	13.0	8.0	—	40.0	21.0	29.0
p-value	—	5.37e-03	5.37e-03	1.53e-03	—	2.77e-01	2.56e-02	8.33e-02

The symbols ↑ and ↓ have similar meanings as in Table I.



(a) f_1 (b) f_8

Fig. 2: The average convergence graph over 25 independent runs on the selected CEC'2013 functions.

slowly, but when CCFR-IDG2 finished evolving the last variable with the largest weight value, the best overall objective value dropped sharply. f_8 is a partially separable function with imbalance between subcomponents. For f_8, compared with MA-SW-Chains and MOS-CEC2013, in the beginning of the evolutionary process, CCFR-IDG2 converged very slowly. When the first evolutionary cycle ended (about 0.8×10^6 function evaluations), CCFR-IDG2 started to allocate most computational resources to the subpopulation which made the greatest improvement of the best overall objective value. CCFR-IDG2 converged much faster than MA-SW-Chains and MOS-CEC2013. This indicates that if the optimizer used by CCFR-IDG2 performs well on a function, CCFR might outperform MA-SW-Chains and MOS-CEC2013 on this function.

To show a better performance of CCFR-IDG2, we replaced SaNSDE with a better optimizer (i.e., CMAES [9]). Table V summarizes the results of CCFR-IDG2 with CMAES. CCFR-IDG2 with CMAES significantly outperformed MA-SW-Chains

TABLE IV: The average errors $\pm$ standard deviations on the CEC'2010 and the CEC'2013 functions over 25 independent runs. The significant best results are in bold font (Wilcoxon rank sum test with Holm p-value correction, α=0.05). R^+, R^- and p-value have similar meanings as in Table I.

CEC'2010 Functions			
F	CCFR-IDG2	MA-SW-Chains	MOS-CEC2013
f_1	1.62e-05±6.55e-06	3.88e-14±3.59e-14↓	**0.00e+00±0.00e+00**↓
f_2	1.73e+02±8.62e+00	8.63e+02±5.84e+01↑	**0.00e+00±0.00e+00**↓
f_3	1.22e+01±3.66e-01	**5.41e-13±2.13e-13**↓	1.65e-12±6.73e-14↓
f_4	1.26e+11±7.50e+10	2.94e+11±9.32e+10↑	**1.56e+10±6.02e+09**↓
f_5	**9.15e+07±1.61e+07**	1.75e+08±1.03e+08↑	1.11e+08±2.25e+07↑
f_6	6.85e+05±7.05e+05	3.52e+04±1.72e+05	**1.22e-07±6.43e-08**↓
f_7	2.04e-03±3.45e-04	3.30e+02±1.40e+03	**0.00e+00±0.00e+00**↓
f_8	3.19e+05±1.08e+06	9.28e+06±2.36e+07↑	**1.95e+00±8.03e+00**↓
f_9	1.34e+07±1.68e+06	1.45e+07±1.59e+06	**3.46e+06±4.49e+05**↓
f_{10}	**1.81e+03±1.43e+02**	2.06e+03±1.19e+02↑	3.78e+03±1.47e+02↑
f_{11}	**1.99e+01±3.26e+00**	3.69e+01±8.24e+00↑	1.91e+02±4.07e-01↑
f_{12}	2.03e+01±2.23e+01	3.19e-06±5.32e-07↓	**0.00e+00±0.00e+00**↓
f_{13}	5.26e+02±1.04e+02	1.09e+03±6.29e+02↑	7.14e+02±5.68e+02
f_{14}	3.08e+07±3.35e+06	3.34e+07±3.37e+06↑	**9.80e+06±6.03e+05**↓
f_{15}	3.18e+03±1.51e+02	**2.69e+03±9.75e+01**↓	7.44e+03±1.84e+02↑
f_{16}	**2.01e+01±2.62e+00**	1.08e+02±1.51e+01↑	3.82e+02±1.55e+01↑
f_{17}	6.72e+01±8.68e+01	1.26e+00±9.45e-02↓	**2.83e-07±7.97e-08**↓
f_{18}	1.37e+03±1.93e+02	1.87e+03±5.79e+02↑	1.54e+03±7.46e+02
f_{19}	1.28e+06±1.01e+05	2.85e+05±1.74e+04↓	**2.91e+04±2.14e+03**↓
f_{20}	1.97e+09±1.83e+09	1.05e+03±7.59e+01↓	**3.52e+02±4.43e+02**↓
R^+	—	143.0	73.0
R^-	—	67.0	137.0
p-value	—	1.56e-01	2.32e-01
CEC'2013 Functions			
F	CCFR-IDG2	MA-SW-Chains	MOS-CEC2013
f_1	1.77e-05±4.52e-06	8.49e-13±1.09e-12↓	**1.27e-22±7.41e-23**↓
f_2	**3.64e+02±2.06e+01**	1.22e+03±1.14e+02↑	8.32e+02±4.48e+01↑
f_3	2.07e+01±1.21e-02	2.14e+01±5.62e-02↑	**9.18e-13±5.12e-14**↓
f_4	**9.56e+07±4.03e+07**	4.58e+09±2.46e+09↑	1.74e+08±7.87e+07↑
f_5	2.80e+06±3.18e+05	**1.87e+06±3.06e+05**↓	6.94e+06±8.85e+05↑
f_6	1.06e+06±1.05e+03	1.01e+06±1.53e+04↓	**1.48e+05±6.43e+04**↓
f_7	2.03e+07±2.94e+07	3.45e+06±1.27e+06	**1.62e+04±9.10e+03**↓
f_8	**6.63e+10±9.52e+10**	4.85e+13±1.02e+13↑	8.00e+12±3.07e+12↑
f_9	1.89e+08±2.83e+07	**1.07e+08±1.68e+07**↓	3.83e+08±6.29e+07↑
f_{10}	9.48e+07±1.82e+05	9.18e+07±1.06e+06↓	**9.02e+05±5.07e+05**↓
f_{11}	4.17e+08±3.43e+08	2.19e+08±2.98e+07	**5.22e+07±2.05e+07**↓
f_{12}	1.56e+09±1.58e+09	1.25e+03±1.05e+02↓	**2.47e+02±2.54e+02**↓
f_{13}	1.21e+09±6.00e+08	1.98e+07±1.82e+06↓	**3.40e+06±1.06e+06**↓
f_{14}	3.39e+09±3.06e+09	1.36e+08±2.11e+07↓	**2.56e+07±7.94e+06**↓
f_{15}	9.82e+06±3.69e+06	5.71e+06±7.57e+05↓	**2.35e+06±1.94e+05**↓
R^+	—	34.0	41.0
R^-	—	86.0	79.0
p-value	—	1.51e-01	3.03e-01

The symbols ↑ and ↓ have similar meanings as in Table I.

on almost all the CEC'2010 and the CEC'2013 functions. CCFR-IDG2 with CMAES performed significantly better than MOS-CEC2013 by several orders of magnitude on most of the partially separable functions (the CEC'2010 functions f_4–f_{18}; the CEC'2013 functions f_4–f_{11}).

References

[1] K. Tang, X. Li, P. N. Suganthan, Z. Yang, and T. Weise, "Benchmark functions for the CEC'2010 special session and competition on large-scale global optimization," Nature Inspired Computation and Applications Laboratory, Tech. Rep., 2010.

[2] X. Li, K. Tang, M. N. Omidvar, Z. Yang, and K. Qin, "Benchmark functions for the CEC'2013 special session and competition on large scale global optimization," Evolutionary Computation and Machine Learning Group, RMIT University, Australia, Tech. Rep., 2013.

[3] R. Ros and N. Hansen, "A simple modification in cma-es achieving linear time and space complexity," in *Parallel Problem Solving from Nature–PPSN X*. Springer, 2008, pp. 296–305.

[4] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 18, no. 3, pp. 378–393, 2014.

[5] M. N. Omidvar, M. Yang, Y. Mei, X. Li, and X. Yao, "IDG: A faster and more accurate differential grouping algorithm," University of Birmingham, School of Computer Science, Tech. Rep. CSR-15-04, September 2015. [Online]. Available: ftp://ftp.cs.bham.ac.uk/pub/tech-reports/2015/CSR-15-04.pdf

[6] M. N. Omidvar, X. Li, and X. Yao, "Smart use of computational resources based on contribution for cooperative co-evolutionary algorithms," in *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. ACM, 2011, pp. 1115–1122.

TABLE V: The average errors ± standard deviations on the CEC'2010 and the CEC'2013 functions over 25 independent runs. The significant best results are in bold font (Wilcoxon rank sum test with Holm p-value correction, α=0.05). R^+, R^- and p-value have similar meanings as in Table I.

CEC'2010 Functions			
F	CCFR-IDG2 (CMAES)	MA-SW-Chains	MOS-CEC2013
f_1	5.50e-17±2.15e-17	3.88e-14±3.59e-14↑	**0.00e+00±0.00e+00**↓
f_2	5.41e+02±4.80e+01	8.63e+02±5.84e+01↑	**0.00e+00±0.00e+00**↓
f_3	1.02e+00±3.98e-01	**5.41e-13±2.13e-13**↓	1.65e-12±6.73e-14↓
f_4	**4.28e-03±4.98e-03**	2.94e+11±9.32e+10↑	1.56e+10±6.02e+09↑
f_5	1.10e+08±1.60e+07	1.75e+08±1.03e+08↑	1.11e+08±2.25e+07
f_6	9.58e+00±8.51e-01	3.52e+04±1.72e+05↑	**1.22e-07±6.43e-08**↓
f_7	4.47e-07±1.73e-06	3.30e+02±1.40e+03↑	**0.00e+00±0.00e+00**↓
f_8	1.25e+06±1.85e+06	9.28e+06±2.36e+07↑	1.95e+00±8.03e+00
f_9	**9.28e-06±5.47e-06**	1.45e+07±1.59e+06↑	3.46e+06±4.49e+05↑
f_{10}	**1.29e+03±6.14e+01**	2.06e+03±1.19e+02↑	3.78e+03±1.47e+02↑
f_{11}	**2.35e-01±4.08e-01**	3.69e+01±8.24e+00↑	1.91e+02±4.07e-01↑
f_{12}	1.28e-10±9.64e-11	3.19e-06±5.32e-07↑	**0.00e+00±0.00e+00**↓
f_{13}	**4.73e+00±3.79e+00**	1.09e+03±6.29e+02↑	7.14e+02±5.68e+02↑
f_{14}	**2.61e-19±3.26e-20**	3.34e+07±3.37e+06↑	9.80e+06±6.03e+05↑
f_{15}	**2.04e+03±8.22e+01**	2.69e+03±9.75e+01↑	7.44e+03±1.84e+02↑
f_{16}	**8.07e-13±2.60e-14**	1.08e+02±1.51e+01↑	3.82e+02±1.55e+01↑
f_{17}	**7.42e-24±1.63e-25**	1.26e+00±9.45e-02↑	2.83e-07±7.97e-08↑
f_{18}	**1.09e+01±6.87e+00**	1.87e+03±5.79e+02↑	1.54e+03±7.46e+02↑
f_{19}	**2.12e+04±2.21e+03**	2.85e+05±1.74e+04↑	2.91e+04±2.14e+03↑
f_{20}	8.50e+02±2.50e+01	1.05e+03±7.59e+01↑	**3.52e+02±4.43e+02**↓
R^+	—	207.0	157.0
R^-	—	3.0	53.0
p-value	—	1.40e-04	5.22e-02

CEC'2013 Functions			
F	CCFR-IDG2 (CMAES)	MA-SW-Chains	MOS-CEC2013
f_1	5.52e-17±5.70e-18	8.49e-13±1.09e-12↑	**1.27e-22±7.41e-23**↓
f_2	**4.35e+02±3.55e+01**	1.22e+03±1.14e+02↑	8.32e+02±4.48e+01↑
f_3	2.04e+01±5.30e-02	2.14e+01±5.62e-02↑	**9.18e-13±5.12e-14**↓
f_4	**5.58e+03±2.73e+04**	4.58e+09±2.46e+09↑	1.74e+08±7.87e+07↑
f_5	2.19e+06±3.11e+05	**1.87e+06±3.06e+05**↓	6.94e+06±8.85e+05↑
f_6	9.99e+05±1.26e+04	1.01e+06±1.53e+04↑	**1.48e+05±6.43e+04**↓
f_7	**2.22e-08±4.21e-08**	3.45e+06±1.27e+06↑	1.62e+04±9.10e+03↑
f_8	**4.89e+03±1.23e+03**	4.85e+13±1.02e+13↑	8.00e+12±3.07e+12↑
f_9	1.59e+08±3.33e+07	**1.07e+08±1.68e+07**↓	3.83e+08±6.29e+07↑
f_{10}	9.11e+07±1.35e+06	9.18e+07±1.06e+06↑	**9.02e+05±5.07e+05**↓
f_{11}	**4.64e-05±7.47e-05**	2.19e+08±2.98e+07↑	5.22e+07±2.05e+07↑
f_{12}	1.01e+03±5.20e+01	1.25e+03±1.05e+02↑	**2.47e+02±2.54e+02**↓
f_{13}	**2.58e+06±3.00e+05**	1.98e+07±1.82e+06↑	3.40e+06±1.06e+06↑
f_{14}	3.63e+07±3.21e+06	1.36e+08±2.11e+07↑	**2.56e+07±7.94e+06**↓
f_{15}	2.80e+06±2.77e+05	5.71e+06±7.57e+05↑	**2.35e+06±1.94e+05**↓
R^+	—	103.0	77.0
R^-	—	17.0	43.0
p-value	—	1.25e-02	3.59e-01

The symbols ↑ and ↓ have similar meanings as in Table I.

[7] D. Molina, M. Lozano, and F. Herrera, "MA-SW-Chains: Memetic algorithm based on local search chains for large scale continuous global optimization," in *Evolutionary Computation, IEEE Congress on*, July 2010, pp. 1–8.

[8] A. LaTorre, S. Muelas, and J.-M. Pena, "Large scale global optimization: Experimental results with mos-based hybrid algorithms," in *IEEE Congress on Evolutionary Computation*, 2013, pp. 2742–2749.

[9] N. Hansen, S. D. Müller, and P. Koumoutsakos, "Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es)," *Evol. Comput.*, vol. 11, no. 1, pp. 1–18, Mar. 2003.