Supplementary File of the paper "Objective Space-based Population Generation to Accelerate Evolutionary Algorithms for Large-scale Many-objective Optimization"

Qi Deng, Qi Kang, *Senior Member, IEEE*, Liang Zhang, MengChu Zhou, *Fellow, IEEE*, and Jing An, *Member, IEEE*

I. ADDITIONAL EXPERIMENTAL RESULTS

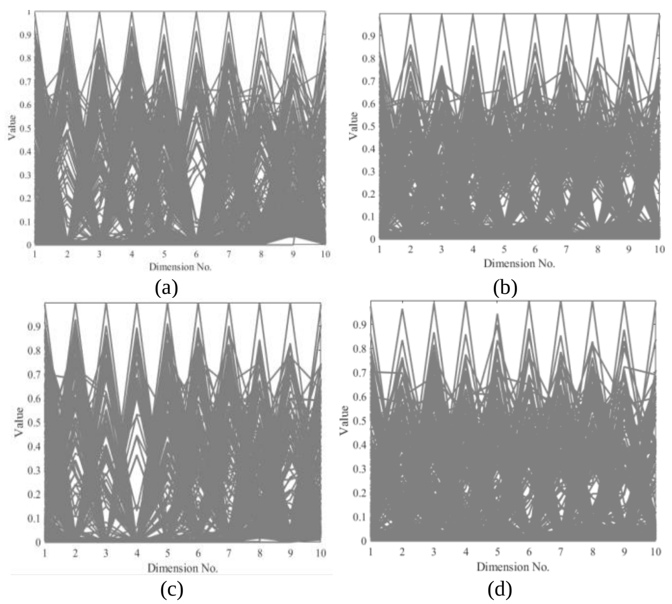


Fig. S1. The non-dominated solution obtained by EAGO and LMEA on ten-objective DTLZ2 problem. (a) EAGO on 500D. (b) LMEA on 500D. (c) EAGO on 1000D. (d) LMEA on 1000D.

Fig. S1 plots the non-dominated solutions. Figs. S1(a) and (b) represent non-dominated solutions of EAGO and LMEA under 10-objective and 500-dimensional DTLZ2 problem, and Figs. S1(c) and (d) represent those of 1,000-dimension problems. There is no obvious difference in their final non-dominated solution distribution, and the overall performance of our proposed EAGO algorithm can reach the level of the most advanced algorithm.

Fig. S2 shows the time iterative process of each algorithm on UF9 problem (the purplish red dotted line is the convergence line of IGD value). It can now be seen that under the same number of evaluations, the time for EAGO to converge and finally stop is the shortest.

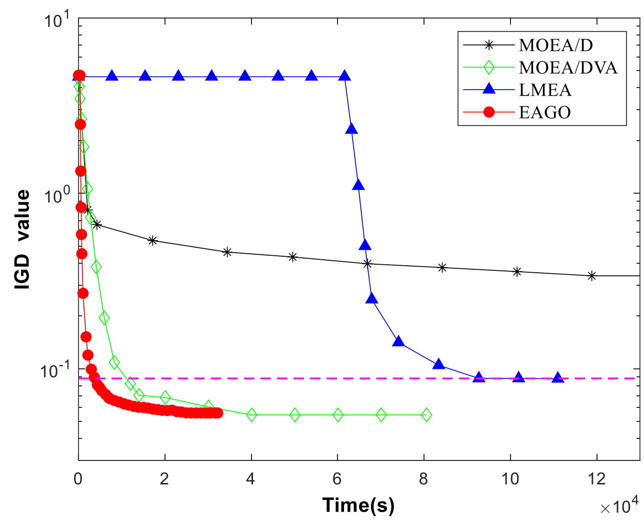


Fig. S2. The time iterative process obtained by four algorithms on three-objective 5000-dimensional UF9 problem with 230000000 evaluation times.

Table S-I shows the statistical results of the IGD values obtained by four algorithms on DTLZ1–DTLZ7 and UF9-10 problems. The Wilcoxon rank sum test is adopted at a significance level of 0.05. The symbols "+", "-" and "≈" respectively indicate that the results are significantly better, significantly worse and similar to the results obtained by EAGO. It can be seen that the proposed EAGO has exceeded or at least reached its peers' best level in overall performance.

Table S-II shows the statistical results of the IGD values obtained by four algorithms on LSMOP1–LSMOP9 problems, and Table S-III shows the running time of four algorithms. It can be seen that the proposed EAGO is still able to achieve a consistent performance on the LSMOP test problems with the DTLZ problems, which exceeds or reaches best level while achieving great saving in execution time.

Table S-IV lists the shortest time required for IGD to reach specified convergence value obtained by four algorithms on LSMOP1-9 problems. And Fig. S3 shows the time iterative process of the IGD values of four algorithms under 10-objective and 1000-dimensional LSMOP1-9 problems.

TABLE S-I
IGD METRIC VALUES OF THE FOUR ALGORITHMS ON DTLZ1–DTLZ7 AND UF9-10, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	MOEA/D	MOEA/DVA	LMEA	EAGO(**ours**)
DTLZ1	5	100	2.1909E+00(1.63E-03) −	1.0056E+01(5.72E-05) −	**6.2360E-02(4.22E-04)** +	2.8348E-01(4.04E-03)
		500	9.7995E+00(4.30E-02) −	4.1455E+01(1.63E-04) −	**6.4216E-02(5.16E-04)** +	8.1869E-01(6.16E-03)
		1000	4.9758E+00(8.84E-01) −	8.7046E+01(4.66E-01) −	**6.4562E-02(4.19E-04)** +	6.4086E-01(5.93E-03)
	10	100	3.1757E+01(4.90E+00) −	3.0555E+01(1.79E-02) −	**1.0524E-01(4.87E-03)** +	1.3804E-01(7.03E-03)
		500	3.2376E+02(1.01E+02) −	1.5349E+02(1.14E-02) −	**1.0813E-01(6.05E-04)** +	1.2807E-01(3.45E-04)
		1000	5.2867E+02(2.16E+00) −	2.3983E+02(1.20E+01) −	**1.1219E-01(8.75E-04)** +	1.2774E-01(2.20E-04)
DTLZ2	5	100	1.9490E-01(2.26E-08) −	2.9042E-01(9.26E-08) −	**1.9126E-01(2.14E-03)** ≈	1.9291E-01(7.50E-05)
		500	1.9490E-01(5.31E-08) −	2.9044E-01(3.17E-08) −	1.9674E-01(2.29E-03) −	**1.9322E-01(2.64E-03)**
		1000	1.9421E-01(6.90E-04) −	2.9047E-01(5.00E-06) −	1.9088E-01(1.10E-03) −	**1.8315E-01(5.00E-05)**
	10	100	4.2219E-01(1.81E-02) −	6.1598E-01(5.13E-02) −	3.9818E-01(1.48E-02) −	**3.9480E-01(1.30E-04)**
		500	4.2213E-01(8.63E-03) −	5.8377E-01(3.53E-04) −	**3.9339E-01(1.76E-02)** +	3.9883E-01(6.55E-04)
		1000	4.2314E-01(1.00E-03) −	6.0197E-01(1.60E-02) −	3.9338E-01(8.69E-04) −	**3.9100E-01(5.30E-04)**
DTLZ3	5	100	8.3591E+00(2.49E-03) −	2.9242E+01(5.92E-05) −	**1.9540E-01(2.14E-03)**+	3.6499E-01(6.54E-03)
		500	1.3009E+02(1.97E-01) −	1.5345E+02(1.01E-04) −	**2.0058E-01(4.44E-03)** +	4.1153E+00(2.47E-01)
		1000	1.6742E+02(2.93E+01) −	2.7859E+02(1.06E+01) −	**2.0518E-01(3.35E-04)** +	3.5745E+00(3.12E-01)
	10	100	1.5067E+02(4.87E-02) −	1.2187E+02(3.77E-02) −	**3.9735E-01(3.77E-02)** ≈	4.0103E-01(1.91E-03)
		500	1.3511E+03(2.78E+02) −	6.5384E+02(9.88E-02) −	4.0865E-01(9.88E-02) −	**4.0642E-01(2.80E-04)**
		1000	2.2226E+03(1.32E+01) −	1.0482E+03(7.50E+00) −	4.1485E-01(2.90E-04) −	**3.9108E-01(1.13E-03)**
DTLZ4	5	100	1.1081E+00(2.50E-01) −	6.3366E-01(1.29E-01) −	**2.5338E-01(1.55E-02)** +	3.3175E-01(2.42E-02)
		500	4.2067E-01(1.18E-01) −	6.3368E-01(1.29E-01) −	**1.9164E-01(2.46E-02)** +	2.6051E-01(6.20E-03)
		1000	3.0779E-01(1.13E-01) −	6.3474E-01(1.00E-03) −	5.7542E-01(6.99E-02) −	**2.6215E-01(1.41E-03)**
	10	100	6.2007E-01(3.18E-02) −	7.1535E-01(3.33E-02) −	**4.3188E-01(2.47e-02)** +	5.9721E-01(4.65E-03)
		500	6.1999E-01(2.93E-02) +	7.8950E-01(5.83E-05) ≈	**4.0043E-01(3.89e-02)** +	7.2801E-01(1.45E-03)
		1000	6.2287E-01(2.80E-03) −	7.0918E-01(2.27E-02) −	4.8719E-01(2.23E-02) −	**4.5746E-01(1.77E-03)**
DTLZ5	5	100	3.2795E-02(9.31E-07) −	4.8810E-01(5.06E-04) −	4.3544E-03(1.44E-04) −	**4.3206E-03(7.06E-05)**
		500	3.2793E-02(1.04E-06) −	1.8921E-01(5.20E-08) −	**4.0265E-03(1.48E-04)** +	4.7423E-03(1.14E-04)
		1000	3.2795E-02(5.00E-07) −	5.7320E-02(3.47E-02) −	4.2209E-03(1.73E-04) −	**4.1726E-03(1.02E-04)**
	10	100	1.9884E-02(2.41E-04) −	3.1085E+00(1.87E-04) −	2.5612E-03(6.95E-05) −	**2.4812E-03(8.00E-05)**
		500	1.9885E-02(4.61E-04) −	1.0974E+00(3.30E-04) −	**2.5242E-03(1.96E-05)** +	3.5642E-03(1.06E-04)
		1000	1.9878E-02(1.00E-06) −	1.1377E+00(1.53E-01) −	**2.4686E-03(7.35E-06)** +	2.8748E-03(4.20E-05)
DTLZ6	5	100	2.0741E-01(3.14E-02) −	5.9687E-01(2.43E-06) −	3.9526E-03(2.14E-04) ≈	**3.9479E-03(3.68E-04)**
		500	1.2960E+00(1.04E-01) −	2.9873E+00(4.25E-07) −	**4.0104E-03(1.22E-03)** +	6.2329E-03(1.28E-04)
		1000	5.2191E+00(1.62E-02) −	3.4444E+00(9.43E-06) −	4.4128E-03(3.16E-04) −	**3.9438E-03(1.34E-04)**
	10	100	1.1803E-01(9.81E-02) −	2.1714E+00(4.09E-02) −	2.2089E-03(5.11E-04) −	**1.7114E-03(1.40E-05)**
		500	1.1966E+00(2.52E-01) −	1.6027E+01(5.33E-02) −	1.5958E-03(3.55E-05) ≈	**1.5435E-03(5.05E-06)**
		1000	3.1999E+00(2.96E-01) −	4.4972E+01(2.72E+00) −	2.8410E-03(1.27E-03) −	**1.5603E-03(4.23E-05)**
DTLZ7	5	100	7.0560E-01(2.57E-02) −	3.8346E-01(2.51E-06) −	**3.2417E-01(1.10E-02)** ≈	3.2803E-01(2.45E-03)
		500	7.0574E-01(3.73E-07) −	3.8147E-01(4.57E-07) −	**3.3236E-01(8.53E-03)** +	3.3965E-01(1.18E-03)
		1000	7.0589E-01(1.00E-05) −	3.8148E-01(7.57E-07) −	3.2892E-01(6.80E-03) −	**3.1768E-01(8.55E-04)**
	10	100	2.8119E+00(1.37E+00) −	1.0258E+00(7.84E-02) −	8.9741E-01(6.40E-03) −	**8.3919E-01(3.02E-02)**
		500	2.5544E+00(6.94E-01) −	9.6536E-01(1.02E-01) +	8.8646E-01(1.27E-02) −	**8.6650E-01(7.25E-03)**
		1000	2.5550E+00(1.05E-03) −	9.6057E-01(3.70E-03) +	**8.8562E-01(1.96E-02)** ≈	8.8936E-01(1.24E-02)
UF9	3	5000	3.4451E-01(9.21E-02) −	**4.3421E-02(7.01E-06)** +	8.7621E-02(4.50E-03) −	5.5912E-02(3.26E-03)
UF10	3	5000	9.0791E-01(2.15E-01) −	**4.2932E-01(1.13E-03)** +	4.6523E-01(9.36E-03) −	4.4971E-01(5.24E-03)
+/−/≈			1 / 43 / 0	4 / 39 / 1	19 / 19 / 6	/
AvgRank			3.2955	3.5455	1.5682	1.5909
p-Value			0.0000	0.0000	1.0000	/

"+", "−" and "≈"respectively indicate that the result is better, worse and similar to that of EAGO.
"AvgRank" represents the average ranking results through Friedman test
" *p*-Value" indicates the significance difference after correction by the Bonferroni procedure

TABLE S-II
IGD METRIC VALUES OF THE FOUR ALGORITHMS ON LSMOP1–LSMOP9, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	MOEA/D	MOEA/DVA	LMEA	EAGO
LSMOP1	5	500	1.5198E-01(2.89E-03) −	1.9918E-01(7.47E-03) ≈	1.5635E-01(3.93E-03) ≈	**1.3851E-01(7.82E-03)**
		1000	1.5049E-01(4.16E-03) −	1.6052E-01(6.87E-03) −	1.7462E-01(2.22E-02) −	**1.2854E-01(1.86E-03)**
	10	500	**2.4726E-01(1.45E-03)** ≈	3.7106E-01(7.36E-02) ≈	3.8762E-01(6.81E-02) ≈	2.8823E-01(2.58E-01)
		1000	3.7361e-01(2.77E-02) −	3.4800E-01(5.78E-02) −	3.8146E-01(7.42E-02) −	**3.2069E-01(1.13E-02)**
LSMOP2	5	500	1.6312E-01(5.90E-04) −	2.1441E-01(2.16E-03) −	1.5032E-01(3.08E-03) −	**1.4415E-01(8.18E-03)**
		1000	1.5670E-01(1.07E-02) ≈	1.8957E-01(6.97E-03) −	**1.3856E-01(2.72E-03)** +	1.5799E-01(1.95E-04)
	10	500	**2.8025E-01(1.44E-03)**+	4.7619E-01(2.58E-02) ≈	4.6867E-01(1.63E-01) ≈	5.0923E-01(5.14E-02)
		1000	**2.5295e-01(8.98E-04)**+	4.3356E-01(2.35E-02) −	2.6684E-01(1.61E-01) +	2.9656E-01(1.37E-02)
LSMOP3	5	500	4.9613E-01(4.74E-02) +	5.4185E-01(4.55E-02) +	**3.6979E-01(9.90E-04)** ≈	6.8738E-01(2.65E-01)
		1000	9.5775e-01(4.00E-02) −	6.9438E-01(7.07E-02) −	**3.8518E-01(6.65E-02)** +	4.3269E-01(6.21E-02)
	10	500	**3.6680E-01(7.27E-03)** +	1.7428e+00(3.68E-02) −	1.4623E+00(7.08E-01) −	8.2385E-01(2.40E-01)
		1000	4.7779E-01(4.13E-02) −	1.6334e+00(8.34E-02) −	8.7072E-01(1.34E-01) −	**4.5733E-01(7.82E-03)**
LSMOP4	5	500	2.3304E-01(1.19E-03) −	1.9378E-01(4.40E-03) −	2.2776E-01(4.95E-02) −	**1.9008E-01(1.32E-02)**
		1000	1.8898E-01(4.55E-03) −	1.8523E-01(1.71E-03) −	1.7092E-01(8.44E-03) −	**1.6166E-01(1.34E-03)**
	10	500	**3.1226E-01(4.50E-03)** +	4.0636E-01(3.36E-03) +	5.5467E-01(1.48E-02) ≈	5.1321E-01(1.94E-02)
		1000	**2.6229E-01(1.38E-02)**≈	3.7235E-01(5.32E-02) −	4.0673E-01(1.86E-02) −	3.7010E-01(4.10E-02)
LSMOP5	5	500	8.2787E-01(8.31E-03) −	3.7166E-01(3.92E-02) −	7.4712E-01(9.87E-02) −	**3.2470E-01(7.01E-02)**
		1000	8.2497E-01(6.32E-02) −	2.8307E-01(4.75E-04) −	2.9276E-01(3.07E-02) −	**2.6556E-01(1.40E-02)**
	10	500	9.3518E-01(3.06E-02) −	6.2192E-01(2.34E-02) −	8.9686E-01(7.54E-01) −	**5.1374E-01(2.53E-02)**
		1000	9.7138E-01(2.47E-02) −	**5.8421E-01(3.31E-02)** +	6.6891E-01(1.10E-01) −	6.0031E-01(1.03E-02)
LSMOP6	5	500	**9.5750E-01(9.56E-02)** +	1.2273E+00(1.79E-01) ≈	1.3689E+00(3.39E-01) +	1.5263E+00(2.35E-01)
		1000	1.2629E+00(2.44E-01) −	1.6352E+00(4.09E-01) −	**9.3200E-01(1.51E-01)** +	1.1705E+00(6.60E-02)
	10	500	**1.0632E+00(2.84E-02)** +	1.7226E+00(1.52E-01) −	1.7249E+00(1.87E-01) −	1.2258E+00(2.95E-02)
		1000	**1.0807E+00(1.54E-02)** +	1.3902E+00(1.74E-01) −	1.3082E+00(2.29E-01) −	1.2049E+00(1.40E-02)
LSMOP7	5	500	8.9659E-01(4.38E-02) −	9.4612E-01(2.81E-02) −	5.9449E-01(8.13E-02) −	**4.6983E-01(4.57E-02)**
		1000	8.3295E-01(1.53E-01) −	9.0375E-01(5.27E-02) −	5.8573E-01(7.21E-02) −	**4.9597E-01(5.88E-02)**
	10	500	**1.1426E+00(8.50E-04)** +	8.7059E+00(3.24E-02) −	2.0985E+00(2.97E-01) −	1.8473E+00(4.03E-01)
		1000	1.7821E+00(2.01E-02)−	4.7198e+00(1.26E-02) −	**1.2239E+00(2.83E-01)** +	1.4792E+00(5.25E-01)
LSMOP8	5	500	8.7296E-01(5.60E-02) −	2.5087E-01(6.19E-02) ≈	2.4855E-01(2.81E-03) ≈	**2.3832E-01(1.06E-02)**
		1000	8.7677E-01(1.67E-02) −	2.7103E-01(5.51E-03) −	2.9536E-01(1.36E-02) −	**2.5170E-01(1.94E-02)**
	10	500	7.6708E-01(4.25E-04) −	6.4933E-01(2.69E-03) −	**4.5501E-01(6.21E-03)** ≈	4.9633E-01(4.76E-02)
		1000	7.7009E-01(1.07E-02) −	6.3258E-01(4.16E-02) −	7.1259E-01(8.09E-02) −	**5.6294E-01(9.93E-02)**
LSMOP9	5	500	7.9467E-01(5.10E-02) +	**4.6537E-01(1.13E-01)** +	7.6733E-01(1.96E-02) ≈	9.2894E-01(1.14E-01)
		1000	9.5151E-01(2.31E-01) +	**4.2847E-01(3.89E-02)** +	7.7786E-01(1.22E-02) +	1.5818E+00(2.51E-02)
	10	500	2.4748E+00(4.15E-03) −	2.9863E+00(2.64E-02) −	4.1357E+00(9.19E-01) −	**1.7948E+00(6.59E-02)**
		1000	2.6898E+00(6.65E-03) −	**1.7574e+00(3.14E-02)** +	3.0989E+00(2.47E-01) −	2.2605E+00(1.66E-01)
+/−/≈			11 / 22 / 3	6 / 25 / 5	7 / 21 / 8	/
AvgRank			2.5833	2.8611	2.6389	1.9167
p-Value			0.1708	0.0115	0.1057	/

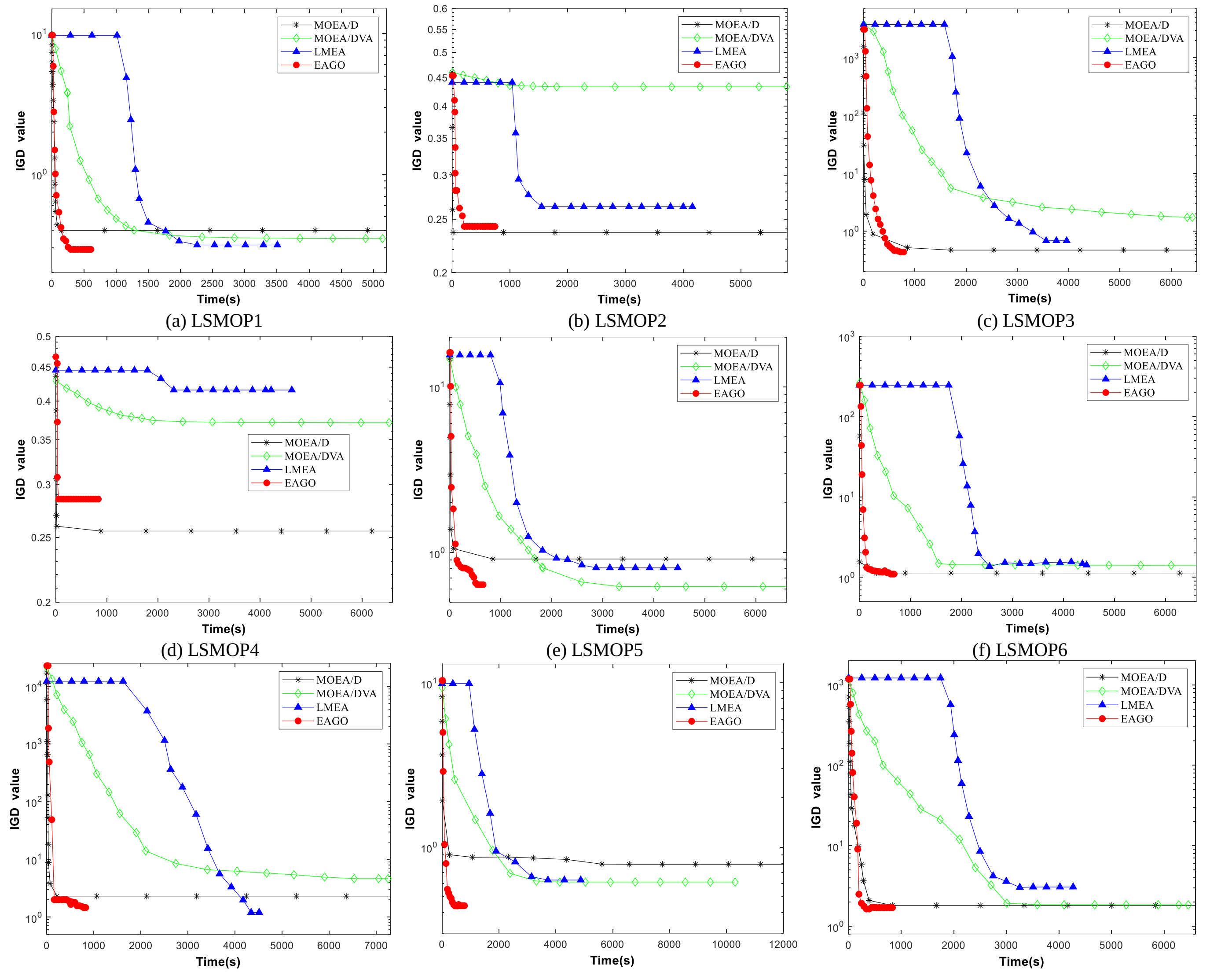
TABLE S-III
THE RUNNING TIME OF THE FOUR ALGORITHMS, TAKING THE BEST VALUES OF EACH INSTANCE

Problem	M	D	MOEA/D	MOEA/DVA	LMEA	EAGO
LSMOP1	5	500	3075.01s−	1693.30s−	969.31s−	**164.14s**
		1000	12844.82s−	4422.39s−	3526.38s−	**542.47s**
	10	500	3711.86s−	1947.65s−	946.08s−	**151.14s**
		1000	13292.18s−	5041.47s−	3667.26s−	**594.45s**
LSMOP2	5	500	3548.23s−	1770.90s−	884.52s−	**188.23s**
		1000	14414.91s−	4444.21s−	4048.30s−	**628.59s**
	10	500	4384.31s−	2210.21s−	921.89s−	**213.61s**
		1000	15120.83s−	5253.37s−	4158.52s−	**744.02s**
LSMOP3	5	500	3313.21s−	1842.78s−	923.38s−	**195.97s**
		1000	13169.61s−	4865.18s−	3531.05s−	**691.96s**
	10	500	4206.56s−	2339.23s−	931.86s−	**240.85s**
		1000	15578.27s−	6409.04s−	3896.86s−	**747.17s**
LSMOP4	5	500	3513.72s−	1949.28s−	1002.97s−	**230.13s**
		1000	13106.44s−	5140.03s−	3860.89s−	**758.71s**
	10	500	4517.56s−	2487.19s−	915.92s−	**267.09s**
		1000	15694.65s−	6247.92s−	4249.03s−	**834.18s**
LSMOP5	5	500	3749.84s−	2006.13s−	978.81s−	**161.61s**
		1000	11753.56s−	5479.15s−	3615.69s−	**541.35s**
	10	500	4498.57s−	2329.05s−	953.72s−	**154.21s**
		1000	14020.38s−	6066.91s−	4094.54s−	**595.58s**
LSMOP6	5	500	4110.94s−	2192.66s−	936.09s−	**166.16s**
		1000	12563.27s−	5629.22s−	4171.18s−	**536.47s**
	10	500	4937.14s−	2690.89s−	960.19s−	**218.46s**
		1000	15240.84s−	6681.43s−	3984.84s−	**661.01s**
LSMOP7	5	500	4131.33s−	2204.07s−	1018.22s−	**204.21s**
		1000	12891.07s−	5920.00s−	3889.83s−	**689.69s**
	10	500	5078.94s−	2836.40s−	1007.34s−	**243.05s**
		1000	15480.03s−	7250.41s−	4156.37s−	**833.75s**
LSMOP8	5	500	4318.58s−	2324.44s−	1001.99s−	**192.14s**
		1000	13131.56s−	6180.25s−	4244.74s−	**648.07s**
	10	500	5050.57s−	2826.02s−	936.27s−	**230.81s**
		1000	15377.88s−	10259.71s−	4249.61s−	**752.01s**
LSMOP9	5	500	3787.84s−	2044.83s−	987.42s−	**196.99s**
		1000	11930.12s−	5743.29s−	3819.01s−	**655.22s**
	10	500	4704.12s−	2417.59s−	984.38s−	**209.09s**
		1000	14182.91s−	6464.36s−	4185.90s−	**789.63s**
+/−/≈			0 / 36 / 0	0 / 36 / 0	0 / 36 / 0	/
AvgRank			4	3	2	1
p-Value			0.0000	0.0000	0.0061	/

TABLE S-IV
THE SHORTEST TIME REQUIRED FOR IGD TO REACH CONVERGENCE VALUE

Problem	convergent IGD	EAGO	LMEA	MOEA/DVA	MOEA/D
LSMOP1	4.0E-01	155.99s	1774.28s	1419.62s	**151.52s**
LSMOP2	3.0E-01	54.97s	1152.73s	/	**4.38s**
LSMOP3	7.0E-01	448.49s	3626.99s	/	**324.45s**
LSMOP4	3.8E-01	31.49s	/	1601.29s	**3.01s**
LSMOP5	7.0E-01	**524.42s**	2810.44s	2390.19s	/
LSMOP6	1.5E+00	132.69s	4372.53s	1867.73s	**25.17s**
LSMOP7	2.3E+00	**105.53s**	2752.21s	/	223.10s
LSMOP8	7.0E-01	**142.78s**	2984.81s	2367.47s	/
LSMOP9	3.1E+00	**183.06s**	3533.65s	3191.14s	307.30s

"/" indicates that the algorithm cannot reach specified convergence value.



(a) LSMOP1 (b) LSMOP2 (c) LSMOP3
(d) LSMOP4 (e) LSMOP5 (f) LSMOP6
(g) LSMOP7 (h) LSMOP8 (i) LSMOP9

Fig. S3. The time iterative process obtained by four algorithms on ten-objective 1000-dimensional LSMOP1-9 problems with 17000000 evaluation times. (a) IGD time iterative process on LSMOP1. (b) IGD time iterative process on LSMOP2. (c) IGD time iterative process on LSMOP3. (d) IGD time iterative process on LSMOP4. (e) IGD time iterative process on LSMOP5. (f) IGD time iterative process on LSMOP6. (g) IGD time iterative process on LSMOP7. (h) IGD time iterative process on LSMOP8. (g) IGD time iterative process on LSMOP9.

II. Parameter Analysis

To apply the Taguchi's parameter design method to our work, we replace four factors (n_s, n_d, $\check{c}_f$, $\hat{c}_f$) with A, B, C, D as shown in Table S-V, each of which has three levels. n_s and n_d are used for decision variable analysis. $\check{c}_f$ and $\hat{c}_f$ are the lower and upper bound coefficients of the mutation space.

TABLE S-V
LEVEL OF DESIGN VARIABLE

Parameter	Level 1	Level 2	Level 3
A: n_s	3	5	10
B: n_d	10	30	50
C: $\check{c}_f$	1	0.1	0.01
D: $\hat{c}_f$	1	10	50

TABLE S-VI
ORTHOGONAL EXPERIMENTAL DESIGN

Experiment	A	B	C	D	IGD	HV
1	3	10	1	1	1.5151E-01	9.1410E-01
2	3	30	0.1	10	1.4176E-01	9.3251E-01
3	3	50	0.01	50	1.3602E-01	9.4890E-01
4	5	10	0.1	10	1.5415E-01	9.3668E-01
5	5	30	0.01	1	1.2729E-01	9.2996E-01
6	5	50	1	50	1.4152E-01	9.4533E-01
7	10	10	0.01	10	1.5092E-01	9.4798E-01
8	10	30	1	50	1.2896E-01	9.6346E-01
9	10	50	0.1	1	1.4314E-01	9.2505E-01
Average					1.4170E-01	9.3822E-01

TABLE S-VII
AVERAGE IGD FOR ALL LEVELS OF ALL FACTORS

	A	B	C	D
Level 1	1.4310E-01	1.5219E-01	1.4066E-01	1.4065E-01
Level 2	1.4099E-01	1.3267E-01	1.4635E-01	1.4894E-01
Level 3	1.4101E-01	1.4023E-01	1.3808E-01	1.3550E-01

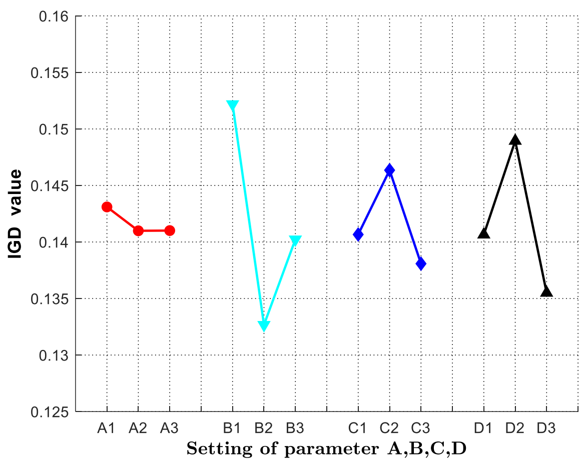


Fig. S4. Main factor effects on IGD

TABLE S-VIII
AVERAGE HV FOR ALL LEVELS OF ALL FACTORS

	A	B	C	D
Level 1	9.3184E-01	9.3292E-01	9.4096E-01	9.2304E-01
Level 2	9.3732E-01	9.4198E-01	9.3141E-01	9.3906E-01
Level 3	9.4550E-01	9.3976E-01	9.4228E-01	9.5256E-01

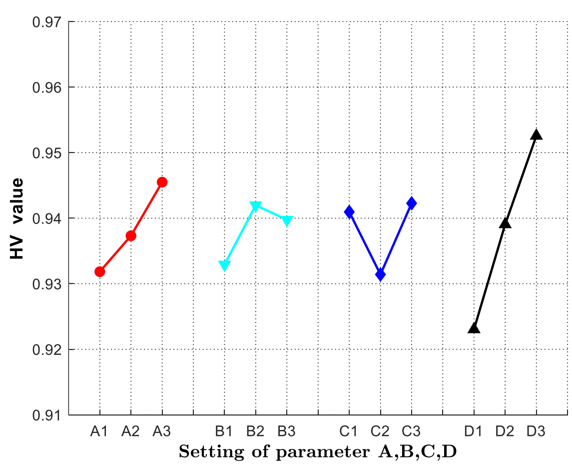


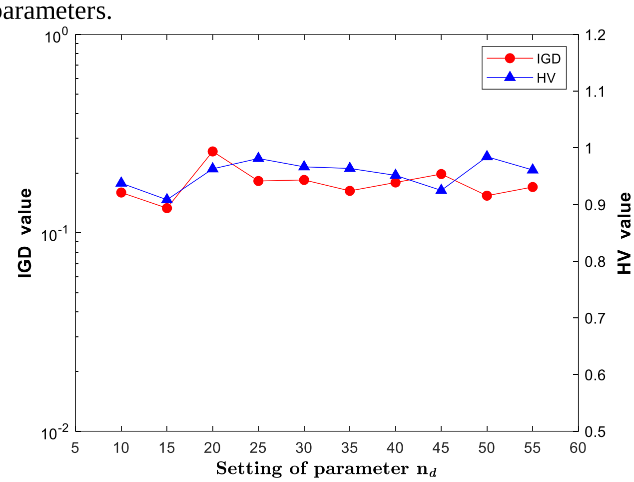
Fig. S5. Main factor effects on HV

TABLE S-IX
EFFECTS OF VARIOUS FACTORS

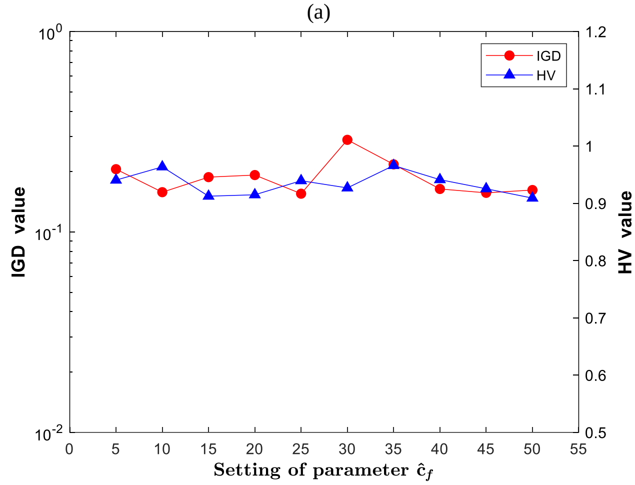
	IGD		HV	
	SSF	Factor effects (%)	SSF	Factor effects (%)
A	9.80E-07	0.91	3.15E-05	14.62
B	6.46E-05	59.71	1.49E-05	6.90
C	1.19E-05	11.04	2.35E-05	10.89
D	3.06E-05	28.34	1.46E-04	67.59
Sum	1.08E-04	100	2.15E-04	100

We carry out experiments using different parameter settings on the 5-objective 500 dimensional LSMOP1 problem to get the value of IGD and HV. Standard experiments with an orthogonal array are listed in Table S-VI. Average IGD and HV for all level of all factors are shown in Tables S-VII and S-VIII, respectively. Figs. S4 and S5 illustrate the main factor effects on IGD and HV. It can be seen that parameter combination (A2, B2, C3, D3) minimizes IGD value and (A3, B2, C3, D3) maximizes HV value.

Analysis of variance is introduced to show the influence of parameters on IGD and HV. The sum of squares of various factors (SSF) can be shown in Table S-IX, which shows that factor B, the number of solutions generated by disturbing a single decision variable ("n_d"), has larger effect on IGD value and factor D, the upper limit coefficient ("$\hat{c}_f$"), has larger effect on HV value. In other words, the results are more sensitive to n_d and $\hat{c}_f$. Fig. S6 shows the specific values of IGD and HV when n_d and $\hat{c}_f$ are changed respectively while fixing other parameters.



(a)



(b)

Fig. S6. The effect of a single parameter changes on the results. (a) Effect of parameter n_d on IGD and HV value. (b) Effect of parameter $\hat{c}_f$ on IGD and HV value.

III. Comparison with Another Inverse Model and Additional Experimental on SDTLZ Test Suites

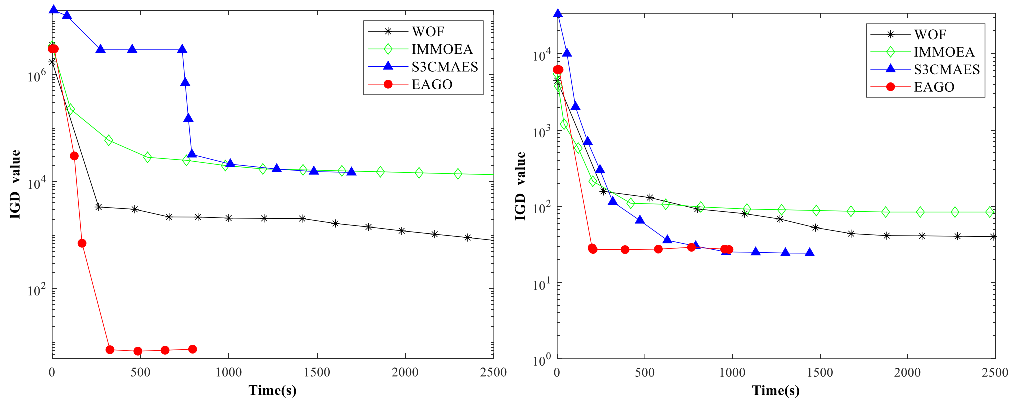
RBFNNs is a representative method to build the inverse model proposed by Giagkiozis and Fleming [59], known as the Pareto estimation method. According to Fleming [59], there are two algorithms using RBFNNs, which can be called MOEA/D-NN and NSGAII-NN. Table S-X shows the experimental results of EAGO and two algorithms on the 5-objective 500 dimensional LSMOP1-9 problems and Table S-XI shows the running time. It can be clearly seen that the performance of EAGO is better than MOEA/D-NN and NSGAII-NN, and EAGO EAGO is faster than them.

TABLE S-X
IGD METRIC VALUES OF THE THREE ALGORITHMS, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	M	D	MOEA/D-NN	NSGAII-NN	EAGO
DTLZ1	5	500	1.9341E+01(3.26E+00) −	1.3955E+03(5.24E+02) −	**8.1869E-01(6.16E-03)**
	10	500	7.6287E+02(4.29E+01) −	3.1432E+03(4.28E+02) −	**1.2807E-01(3.45E-04)**
DTLZ2	5	500	**9.2071E-02(4.56E-03)** +	2.2290E-01(7.24E-02) −	1.9322E-01(2.64E-03)
	10	500	**2.6709E-01(5.24E-02)** +	8.6984E+01(7.24E+00) −	3.9883E-01(6.55E-04)
DTLZ3	5	500	1.7541E+02(6.12E+01) −	4.3753E+03(3.21E+02) −	**4.1153E+00(2.47E-01)**
	10	500	1.5247E+03(2.65E+02) −	4.0331E+04(2.87E+03) −	**4.0642E-01(2.80E-04)**
DTLZ4	5	500	3.9426E-01(4.31E-02) −	3.2578E-01(7.21E-02) −	**2.6051E-01(6.20E-03)**
	10	500	**5.2002E-01(1.68E-02)** ≈	4.3525E+01(4.12E+00) −	7.2801E-01(1.45E-03)
DTLZ5	5	500	1.3031E-02(3.56E-03) −	2.7565E+01(1.58E+00) −	**4.7423E-03(1.14E-04)**
	10	500	2.3641E-02(6.14E-03) −	5.4669E+01(3.24E+00) −	**3.5642E-03(1.06E-04)**
DTLZ6	5	500	1.3724E+00(2.64E-01) −	3.3324E+02(3.14E+01) −	**6.2329E-03(1.28E-04)**
	10	500	2.3768E+00(3.16E-01) −	1.4338E+02(2.57E+01) −	**1.5435E-03(5.05E-06)**
DTLZ7	5	500	3.5685E-01(2.46E-02) ≈	3.6916E-01(3.14E-02) ≈	**3.3965E-01(1.18E-03)**
	10	500	4.3272E+00(2.14E-01) −	2.1024E+01(3.47E+00) −	**8.6650E-01(7.25E-03)**
+/−/≈			2 / 10 / 2	0 / 13 / 1	/
AvgRank			1.8571	2.9286	1.2143
p-Value			0.2669	0.0000	/

TABLE S-XI
THE RUNNING TIME OF THE THREE ALGORITHMS, TAKING THE BEST VALUES OF EACH INSTANCE

Problem	M	D	MOEA/D-NN	NSGAII-NN	EAGO
DTLZ1	5	500	2420.38s−	447.06s−	**195.70s**
	10	500	3210.22s−	523.29s−	**203.91s**
DTLZ2	5	500	1740.57s−	432.01s−	**174.83s**
	10	500	2539.59s−	494.38s−	**183.34s**
DTLZ3	5	500	2406.22s−	451.08s−	**194.03s**
	10	500	3220.62s−	518.30s−	**203.20s**
DTLZ4	5	500	1765.74s−	439.22s−	**177.63s**
	10	500	2586.31s−	500.75s−	**197.36s**
DTLZ5	5	500	1760.21s−	433.45s−	**190.34s**
	10	500	2586.78s−	504.95s−	**239.94s**
DTLZ6	5	500	2209.26s−	477.47s−	**155.24s**
	10	500	2764.42s−	590.08s−	**219.42s**
DTLZ7	5	500	2021.72s−	435.53s−	**133.10s**
	10	500	2205.38s−	617.63s−	**180.84s**
+/−/≈			0 / 14 / 0	0 / 14 / 0	/
AvgRank			3	2	1
p-Value			0.0000	0.0245	/



(a) (b)

Fig. S7. The IGD convergence process obtained by four methods on ten-objective 1000D SDTLZ1-2 problems. (a) IGD on SDTLZ1. (b) IGD on SDTLZ2.

We perform some experiments on SDTLZ1-2 where objectives have entirely different scales. The IGD values and convergence process are illustrated in Table S-XII and Fig. S7 respectively. It can be clearly seen that EAGO can converge faster than other algorithms under the premise of ensuring the same or better overall performance.

TABLE S-XII

IGD VALUES OF THE FOUR ALGORITHMS, TAKING THE BEST VALUES OF EACH INSTANCE

Problem	M	D	S3-CMA-ES	WOF-SMPSO	IM-MOEA	EAGO
SDTLZ1	5	500	7.3113E+03(5.23E+02) −	4.0672E+01(3.14E+00) −	2.8927E+03(3.24E+02) −	**2.6644E+00(2.35E-01)**
	5	1000	1.5277E+04(3.24E+03) −	9.7179E+01(5.24E+00) −	5.1749E+03(2.34E+02) −	**2.2978E+00(6.41E-01)**
	10	500	7.3861E+03(7.15E+02) −	7.5413E+02(4.21E+01) −	4.3298E+03(1.68E+02) −	**6.7515E+00(2.15E-01)**
	10	1000	1.4873E+04(2.54E+03) −	5.7119E+02(3.15E+01) −	9.0610E+03(3.24E+02) −	**7.3077E+00(4.25E-01)**
SDTLZ2	5	500	**9.5078E-01(2.47E-02)** +	1.1995E+00(4.58E-01) −	4.0657E+00(3.24E-01) −	9.9528E-01(5.62E-02)
	5	1000	9.7582E-01(6.24E-02) −	1.1984E+00(3.61E-01) −	4.2245E+00(8.29E-01) −	**9.5117E-01(4.32E-02)**
	10	500	**2.4119E+01(3.65E+00)** +	4.0198E+01(5.63E+00) −	7.5898E+01(6.24E+00) −	2.5899E+01(4.67E+00)
	10	1000	**2.4257E+01(4.35E+00)** +	4.0039E+01(6.24E+00) −	8.1365E+01(5.27E+00) −	2.7573E+01(5.27E+00)
+/−/≈			3 / 5 / 0	0 / 8 / 0	0 / 8 / 0	/
AvgRank			2.6250	2.5000	3.5000	1.3750
p-Value			0.3168	0.4882	0.0060	/

IV. THEORETICAL ANALYSIS AND EFFECTIVENESS RESEARCH OF THE GENERATION PROCESS

We examine the generation process for a single individual. For solution s_s, it has a decision vector $\boldsymbol{x} = (x_1, \cdots, x_i, \cdots, x_D)$ and an objective vector $\boldsymbol{y} = (y_1, \cdots, y_j, \cdots, y_M)$. The current transition matrix is T and $\boldsymbol{y} \times \mathrm{T} = \boldsymbol{x}$. We only consider "objective-decrease" and "objective-increase", and use a simple version $\boldsymbol{r} = (r, \cdots, r, \cdots, r)$, where r is a random number vector between 0 and 1. If a generation operator is "objective-decrease", then

$$\begin{aligned}
\boldsymbol{x}' &= (\boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{r}) \times \mathrm{T} \\
&= (1-r, \cdots, 1-r, \cdots, 1-r) \\
&\quad \otimes (y_1, \cdots, y_j, \cdots, y_M) \times \begin{pmatrix} t_{11} & \cdots & t_{1D} \\ \vdots & & \\ t_{j1} & \cdots & t_{jD} \\ \vdots & & \\ t_{M1} & \cdots & t_{MD} \end{pmatrix} \\
&= (((1-r)y_1 t_{11} + \cdots + (1 \\
&\quad - r)y_M t_{M1}), \cdots, ((1-r)y_1 t_{1D} + \cdots + (1 \\
&\quad - r)y_M t_{MD})) \\
&= (1-r)((y_1 t_{11} + \cdots + y_M t_{M1}), \cdots, (y_1 t_{1D} + \cdots \\
&\quad + y_M t_{MD})) \\
&= (1-r)\boldsymbol{x} \qquad \text{(S1)}
\end{aligned}$$

where $\boldsymbol{x}'$ is a new decision vector and $\otimes$ is element-wise product. Similarly, if it is "objective-increase", $\boldsymbol{x}' = (\boldsymbol{y} + \boldsymbol{y} \otimes \boldsymbol{r}) \times \mathrm{T} = (\boldsymbol{1} + \boldsymbol{r}) \otimes \boldsymbol{y} \times \mathrm{T} = (1+r)\boldsymbol{x}$. So the operation of objective space can lead to the change of the absolute value of a decision vector. If we define a wrong decision, the newly generated decision vector may not reduce any objective function value. For objective function f_1, its multivariate Taylor expansion is $f_1(\boldsymbol{x}') = f_1(\boldsymbol{x}) + \sum_{i=1}^{D}(x_i' - x_i)\frac{\partial f_1(x)}{\partial x_i} + \frac{1}{2!}\sum_{i,k=1}^{D}(x_i' - x_i)(x_k' - x_k)\frac{\partial^2 f_1}{\partial x_i \partial x_k} + O^D$. Taking "objective-decrease" as an example, if it is a wrong decision for solution s_s, $f_j(\boldsymbol{x}') \geq f_j(\boldsymbol{x})$ for all $j = 1 \cdots M$. Or All elements in $H = ((-r)\sum_{i=1}^{D}\frac{\partial f_1(x)}{\partial x_i} + \frac{(-r)^2}{2!}\sum_{i,k=1}^{D}\frac{\partial^2 f_1}{\partial x_i \partial x_k} + O^D, \cdots, (-r)\sum_{i=1}^{D}\frac{\partial f_M(x)}{\partial x_i} + \frac{(-r)^2}{2!}\sum_{i,k=1}^{D}\frac{\partial^2 f_M}{\partial x_i \partial x_k} + O^D \quad)$ are non-negative numbers, to be determined by the partial derivatives of all objectives in all decision dimensions based on the current function point. H can be further simplified to

$\left((-\mathrm{r})\sum_{\mathrm{i}=1}^{\mathrm{D}} \frac{\partial \mathrm{f}_1(\mathrm{x})}{\partial \mathrm{x}_\mathrm{i}} + 0^\mathrm{D}, \cdots, (-\mathrm{r})\sum_{\mathrm{i}=1}^{\mathrm{D}} \frac{\partial \mathrm{f}_\mathrm{M}(\mathrm{x})}{\partial \mathrm{x}_\mathrm{i}} + 0^\mathrm{D} \right)$. When the partial derivative and the current position are uncertain, it is equally possible that each term of H is positive or negative. The case in which H is completely nonnegative can only occur in extreme cases. The probability of such occurrence is relatively small and can be found as $(\frac{1}{2})^M$. So for a single solution, the probability of making a completely wrong decision is relatively small.

Next, we make an illustration of the effectiveness of our model and a more in-depth discussion of nonlinear and linear mappings. In fact, our proposed model can generate new population (solutions) more efficiently, thus ensuring that our model works well. To show this point clearly, we carry out some further experiments to compare our objective space mapping model with a traditional population generating model in decision space. In the experiments, 3-objective 100-dimensional LSMOP1-9 instances are used as an example, where the population size is set to 105. Population is initialized randomly. We use traditional crossover-mutation to generate a new population in the decision space, denoted by traditional model. Besides, we construct our proposed linear mapping matrix T and "objective-decrease" strategy to generate a new population in the objective space, denoted by linear mapping. The first-generation new population's distribution in the objective space of two models are shown in Fig. S8 and their solutions are not sorted or selected (The first-generation means its parent population is the initial population). We also compare the IGD values of two models, which are shown in Table XIII.

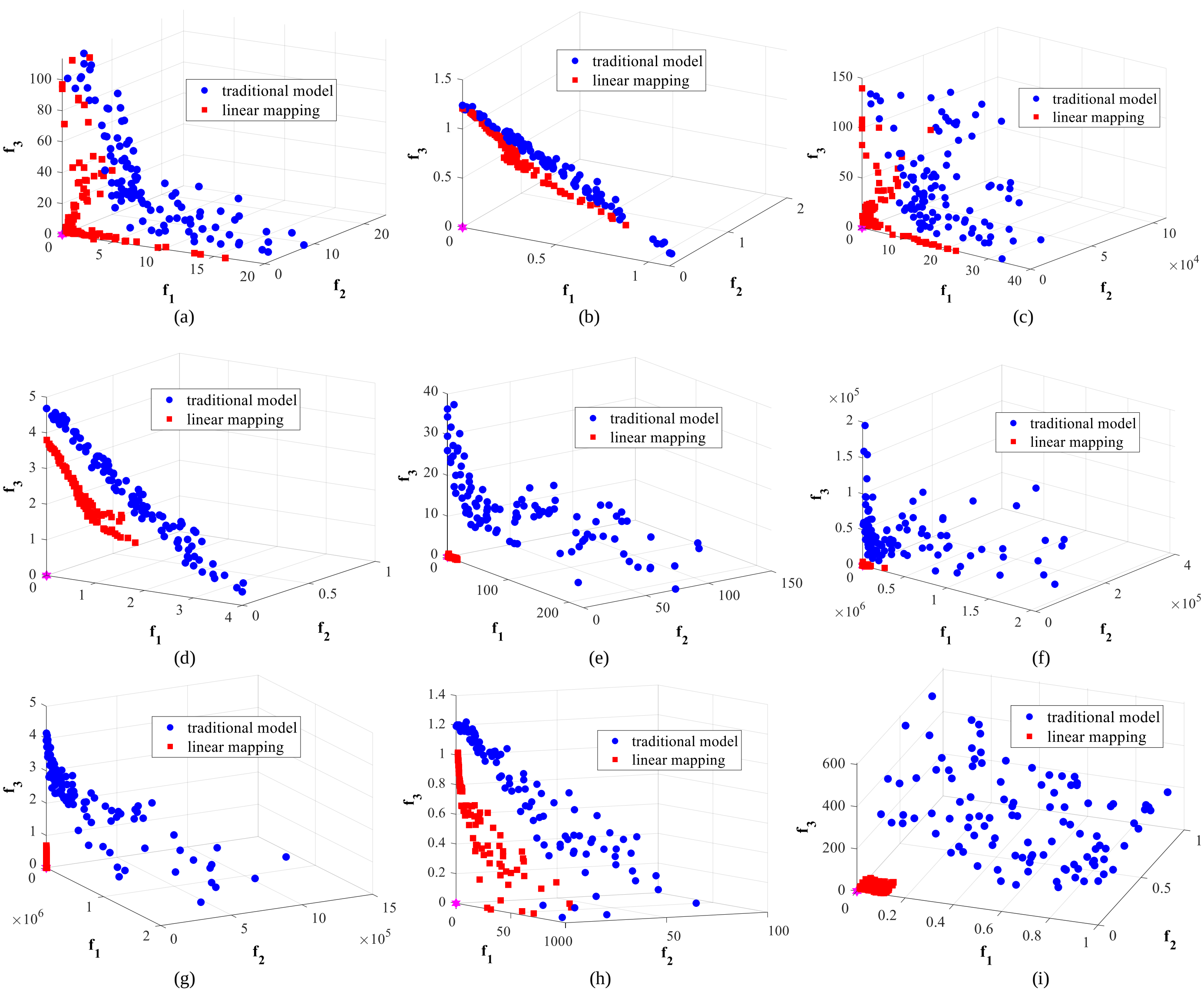


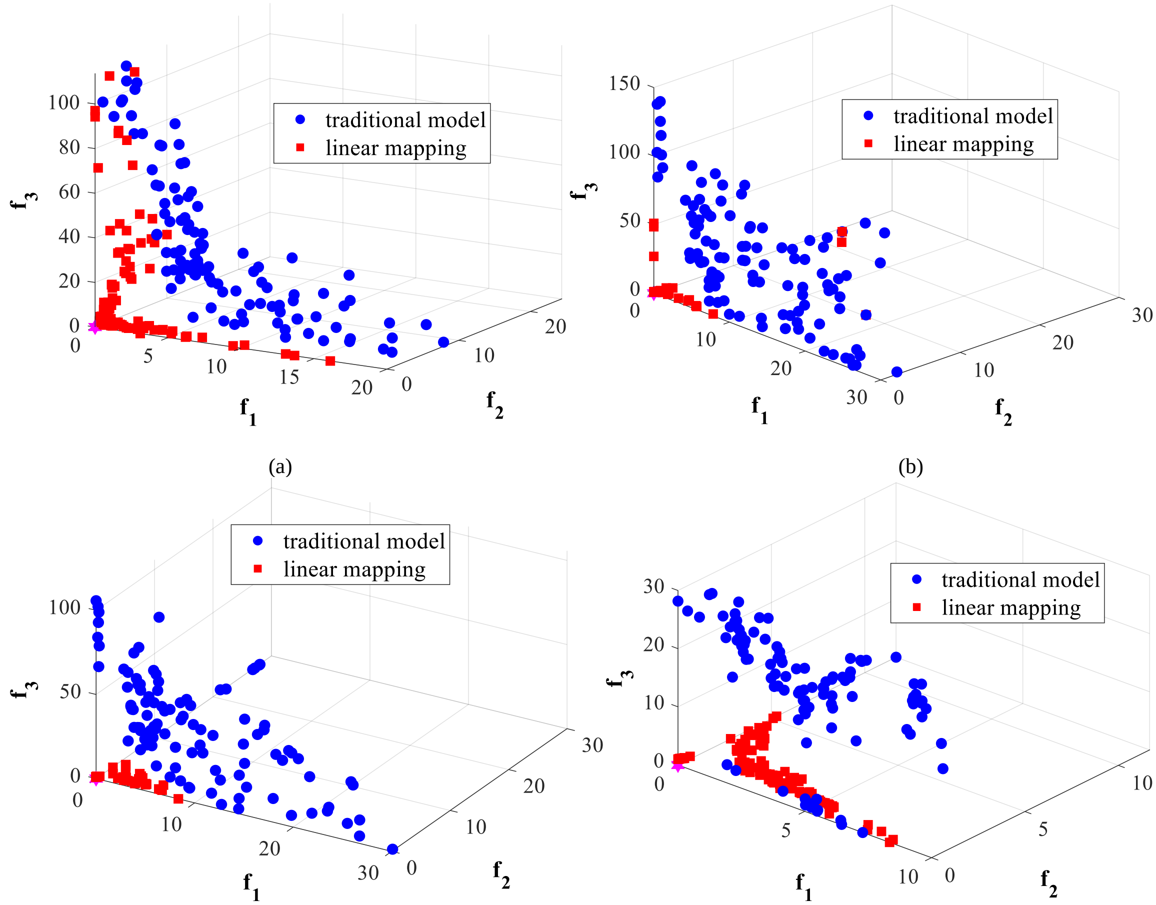
Fig. S8. The first-generation populations' distribution of traditional model and our model (linear mapping) on 3-objective 100-dimensional LSMOP1-9 problems, Magenta points are the ideal point. (a) LSMOP1. (b) LSMOP2. (c) LSMOP3. (d) LSMOP4. (e) LSMOP5. (f) LSMOP6. (g) LSMOP7. (h) LSMOP8. (i) LSMOP9.

TABLE XIII

IGD METRIC VALUES OF TWO POPULATIONS ON LSMOP1–LSMOP9, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	traditional model	linear mapping
LSMOP1	9.7482E+00(6.79E-01) −	**7.1703E-01(2.12E-01)**
LSMOP2	**2.2280E-01(3.41E-03)** +	2.7947E-01(2.70E-01)
LSMOP3	2.0684E+02(1.87E+02) −	**5.0486E+00(3.65E+00)**
LSMOP4	**5.8946E-01(1.51E-02)** +	6.1335E-01(1.76E-01)
LSMOP5	1.5502E+01(1.84E+00) −	**1.4086E+00(5.25E-01)**
LSMOP6	2.3827E+04(5.46E+03) −	**3.8013E+00(1.67E+00)**
LSMOP7	1.2238E+03(1.60E+03) −	**2.6041E+01(3.14E+01)**
LSMOP8	8.6773E-01(8.82E-02) −	**7.7767E-01(3.84E-02)**
LSMOP9	1.1286E+02(7.80E+00) −	**7.5286E+00(5.44E+00)**
+/−/≈	2 / 7 / 0	/

From Fig. S8, we can see that the distribution of solutions generated by using our objective space mapping model is closer to the ideal point. From Table XIII, our model has a clear advantage, with an order of magnitude gap in IGD values in most instances. All of the above illustrate the improvement brought by our objective space mapping model over traditional methods. In addition, we observe the population distribution generated by traditional model and our model (linear mapping) in different generations in the process of evolution, which is shown in Fig. S9 (Population is selected by non-dominated sorting and crowding distance to make it evolve). The figure illustrates that our model is more advanced than traditional population generation method based on decision space. In other words, our objective space mapping model is effective in the generation of new populations. Hence, it ensures the effectiveness of our EAGO.



(c) (d)

Fig. S9. The newly generated solution distribution of traditional model and our model (linear mapping) on 3-objective 100-dimensional LSMOP1 problem in different generations, Magenta points are the ideal point. (a) first-generation. (b) second-generation. (c) 5th-generation. (d) 50th-generation.

TABLE XIV
IGD METRIC VALUES OF THE THREE POPULATIONS ON LSMOP1–LSMOP9 WITH DIFFERENT TRAINING DATA, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Training data size is 105(N)			
Problem	RBFNN	GP	Linear
LSMOP1	2.4862E+00(2.87E-01) −	1.0017E+01(5.62E-01) −	**7.1703E-01(2.12E-01)**
LSMOP2	**1.6918E-01(2.38E-03)** +	2.2558E-01(3.46E-03) +	2.7947E-01(2.70E-01)
LSMOP3	2.6012E+01(6.73E+00) −	4.0067E+02(3.81E+02) −	**5.0486E+00(3.65E+00)**
LSMOP4	**5.3212E-01(3.65E-02)** +	6.1282E-01(3.38E-02) +	6.1335E-01(1.76E-01)
LSMOP5	5.1564E+00(9.73E-01) −	1.4775E+01(2.07E+00) −	**1.4086E+00(5.25E-01)**
LSMOP6	4.1506E+04(3.64E+04) −	2.9343E+04(9.09E+03) −	**3.8013E+00(1.67E+00)**
LSMOP7	3.5376E+02(5.30E+02) −	3.5061E+03(3.13E+03) −	**2.6041E+01(3.14E+01)**
LSMOP8	9.8371E-01(6.63E-02) −	9.4472E-01(1.84E-01) −	**7.7767E-01(3.84E-02)**
LSMOP9	9.8017E+01(4.40E+01) −	1.0779E+02(1.19E+01) −	**7.5286E+00(5.44E+00)**
+/−/≈	**2 / 7 / 0**	**2 / 7 / 0**	**/**
Training data size is 1050(10N)			
Problem	RBFNN	GP	Linear
LSMOP1	1.1173E+00(3.27E-01) −	9.6702E+00(2.65E-01) −	**7.1703E-01(2.12E-01)**
LSMOP2	**1.3271E-01(1.38E-03)** +	2.1801E-01(1.86E-03) +	2.7947E-01(2.70E-01)
LSMOP3	3.3168E+01(8.70E+00) −	3.1084E+02(4.18E+02) −	**5.0486E+00(3.65E+00)**
LSMOP4	**4.4354E-01(5.36E-02)** +	5.8252E-01(5.34E-02) +	6.1335E-01(1.76E-01)
LSMOP5	3.0038E+00(8.39E-01) −	1.0602E+01(3.15E+00) −	**1.4086E+00(5.25E-01)**
LSMOP6	2.1149E+03(2.46E+03) −	2.6374E+04(4.25E+03) −	**3.8013E+00(1.67E+00)**
LSMOP7	**3.3003E+00(3.25E+00)** +	6.7936E+02(1.25E+02) −	2.6041E+01(3.14E+01)
LSMOP8	7.8370E-01(4.25E-02) −	8.1333E-01(3.27E-01) −	**7.7767E-01(3.84E-02)**
LSMOP9	3.0177E+01(2.45E+01) −	1.1071E+02(2.58E+01) −	**7.5286E+00(5.44E+00)**
+/−/≈	**3 / 6 / 0**	**2 / 7 / 0**	**/**
Training data size is 10500(100N)			
Problem	RBFNN	GP	Linear
LSMOP1	**6.9654E-01(3.16E-01)** +	9.1527E+00(7.25E-01) −	7.1703E-01(2.12E-01)
LSMOP2	**1.1774E-01(1.22E-03)** +	2.1934E-01(1.43E-03) +	2.7947E-01(2.70E-01)
LSMOP3	9.5669E+00(2.67E+00) −	3.9497E+02(9.38E+01) −	**5.0486E+00(3.65E+00)**
LSMOP4	**4.3971E-01(4.68E-02)** +	5.7898E-01(8.54E-02) +	6.1335E-01(1.76E-01)
LSMOP5	2.0196E+00(6.37E-01) −	1.6032E+01(2.35E+00) −	**1.4086E+00(5.25E-01)**
LSMOP6	5.8024E+02(1.23E+02) −	1.8107E+04(5.70E+03) −	**3.8013E+00(1.67E+00)**
LSMOP7	**3.3099E+00(2.35E+00)** +	1.5684E+02(1.14E+02) −	2.6041E+01(3.14E+01)
LSMOP8	**7.1463E-01(8.52E-02)** +	7.8059E-01(7.21E-01) −	7.7767E-01(3.84E-02)
LSMOP9	2.4779E+01(1.27E+01) −	1.0104E+02(6.87E+01) −	**7.5286E+00(5.44E+00)**
+/−/≈	**5 / 4 / 0**	**2 / 7 / 0**	**/**

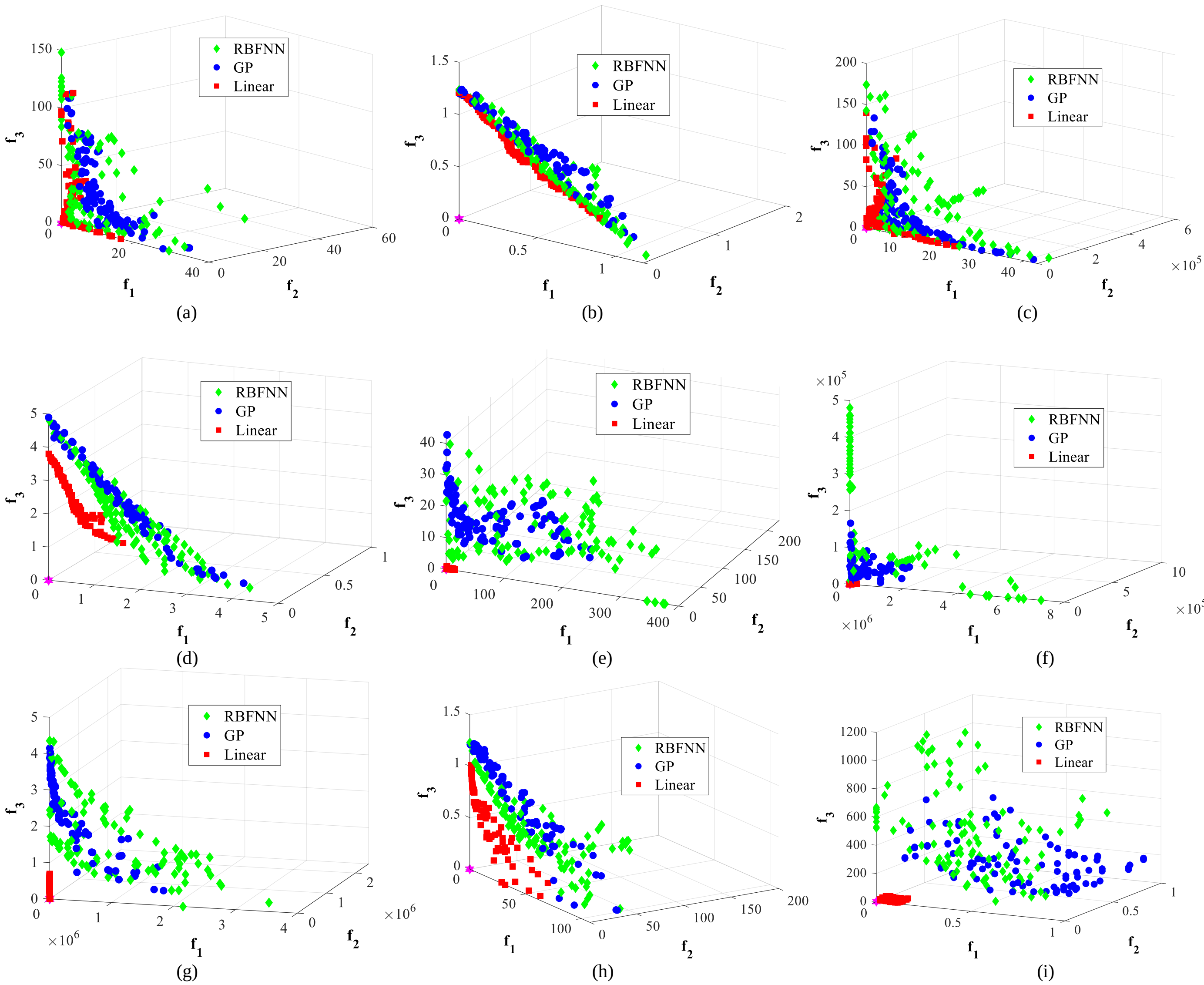


Fig. S10. The first-generation populations' distribution in the objective space of GP, RBFNN and Linear on 3-objective 100-dimensional LSMOP1-9 problems when training data is N, Magenta points are the ideal point. (a) LSMOP1. (b) LSMOP2. (c) LSMOP3. (d) LSMOP4. (e) LSMOP5. (f) LSMOP6. (g) LSMOP7. (h) LSMOP8. (i) LSMOP9.

To make a further discussion of the nonlinear mapping and linear mapping, we continue the above experiments on 3-objective 100-dimensional LSMOP1-9 problems. We choose two representative nonlinear mapping methods, i.e., Radial Basis Function Neural Network (RBFNN) and Gaussian Process (GP). These two nonlinear methods together with the Linear Transition Matrix (Linear) are used as the mapping methods with different size training data are shown in Table XIV (*N* is the population size, and a certain amount of data is needed to train the nonlinear models). When training data is relatively small, the solutions generated using our linear mapping is much better than those with the nonlinear mappings. The distributions of the new population generated by three mapping models in the objective space are shown in Fig. S10, here training data size equals to N=105. We can see that the distribution of the solutions generated by using the linear model is closer to the ideal point (objective) than that with two nonlinear models.

In terms of the inverse function in the absolute numerical sense, the accuracy of our linear mapping does not change over time and the ability does not grow. As the amount of training data increases, a nonlinear mapping model may be able to express the complex function relationship better, and thus the population quality generated by the nonlinear mapping can be improved. For example, when training data size is 100*N*= 10500, the performance of the first-generation population generated by RBFNN can reach or even exceed that of our linear mapping model.

If we enhance the expressive capacity of the mapping model, and further obtain a more general function approximator, the evolutionary search based on objective space mapping will be more efficient. But existing nonlinear models, i.e., deep neural networks, require a large amount of training data and training time for large-scale and many-objective problems, which is very costly. On the premise of ensuring the rapid construction of stable and usable mapping bridge, the inverse model should also have more accurate mapping, especially nonlinear expression ability. This may be a good research topic that is worth studying further.