Supplementary File of "Activation Function-assisted Objective Space Mapping to Enhance Evolutionary Algorithms for Large-scale Many-objective Optimization"

By Q. Deng, et. Al.

I. Algorithm Pseudocode

The procedure of I_{ϵ^+} indicator selection method in EAGOA is presented in Algorithm S1.

Algorithm S1. Q = **IndicatorSelect** (P, L_k , Q)

Input: P (current population), L_k (solutions set), Q (new population);
Output: Q (new population)

/* Utilize the I_{ϵ^+} indicator to select the last layer */
Calculate a Chebyshev distance matrix I between each two solutions in L_k;
$F(\boldsymbol{s}_1) = \sum_{\boldsymbol{s}_2 \in L_k \setminus \boldsymbol{s}_1} -e^{-\mathrm{I}(\boldsymbol{s}_2,\boldsymbol{s}_1)/\kappa}$ for all $\boldsymbol{s}_1 \in L_k$; /* Calculate fitness values of solutions in L_k, κ is a coefficient, usually 0.05 */
while |P| - |Q| < $|L_k|$ **do**
  Remove $\boldsymbol{s}_*$ with the smallest fitness value from L_k;
  $F(\boldsymbol{s}) = F(\boldsymbol{s}) + e^{-\mathrm{I}(\boldsymbol{s}_*,\boldsymbol{s})/\kappa}$ for all $\boldsymbol{s} \in L_k$; /* Update fitness values */
end while
Q ← Q ∪ L_k.

Specifically, for any $\boldsymbol{s}_1, \boldsymbol{s}_2 \in L_k$, a distance matrix is defined as:

$$\mathrm{I}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \max\left(f_j(\boldsymbol{s}_1) - f_j(\boldsymbol{s}_2), 1 \le j \le M\right) \quad (1)$$

where M is the dimensionality of objective space. Then, the fitness value of $\boldsymbol{s}_1$ is calculated as:

$$F(\boldsymbol{s}_1) = \sum_{\boldsymbol{s}_2 \in L_k \setminus \boldsymbol{s}_1} -e^{-\mathrm{I}(\boldsymbol{s}_2,\boldsymbol{s}_1)/\kappa} \quad (2)$$

where κ is a fitness scaling factor, usually is 0.05. The solutions from L_k are selected by I_{ϵ^+} indicator until the size of new population Q is equal to the size of population P.

II. Additional Experimental Results

A. Activation Functions and Inverse Mapping

Besides, we also add more comparison of different nonlinear characteristic problems and validate the effectiveness of the selection of f_L and f_G. The LSMOP7 instance with 10-objective 1000-dimension is chosen as a representative of LSMOPs problems. We generate five subpopulations from an initial population by using $\boldsymbol{x}' = f_\sigma((\boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{r}')\mathrm{T}) = f_\sigma((1-r)\boldsymbol{x})$ and different activation functions, $Tanh$, $Sigmoid$, $Relu$, f_L (3a) and f_G (3b).

$$f_L(x) = \begin{cases} \alpha x, & if\ x \ge 0 \\ x, & if\ x < 0 \end{cases} \quad (3a)$$

$$f_G(x) = \beta(e^x - 1) \quad (3b)$$

Our rank-selection method ("Value-sort" or "Dominated-indicator", more details can be seen in Algorithm 3 and 4) is executed once only to obtain five new populations, denoted by $\mathrm{P^T}$, $\mathrm{P^S}$, $\mathrm{P^R}$, $\mathrm{P^L}$ and $\mathrm{P^G}$ respectively. Fig. S1 shows the IGD values of five subpopulations in 10 independent runs. From Fig. S1, f_L and f_G that we define performs well, especially f_G. This answers our f_L and f_G are still effective even if the nonlinear characteristics are completely changed.

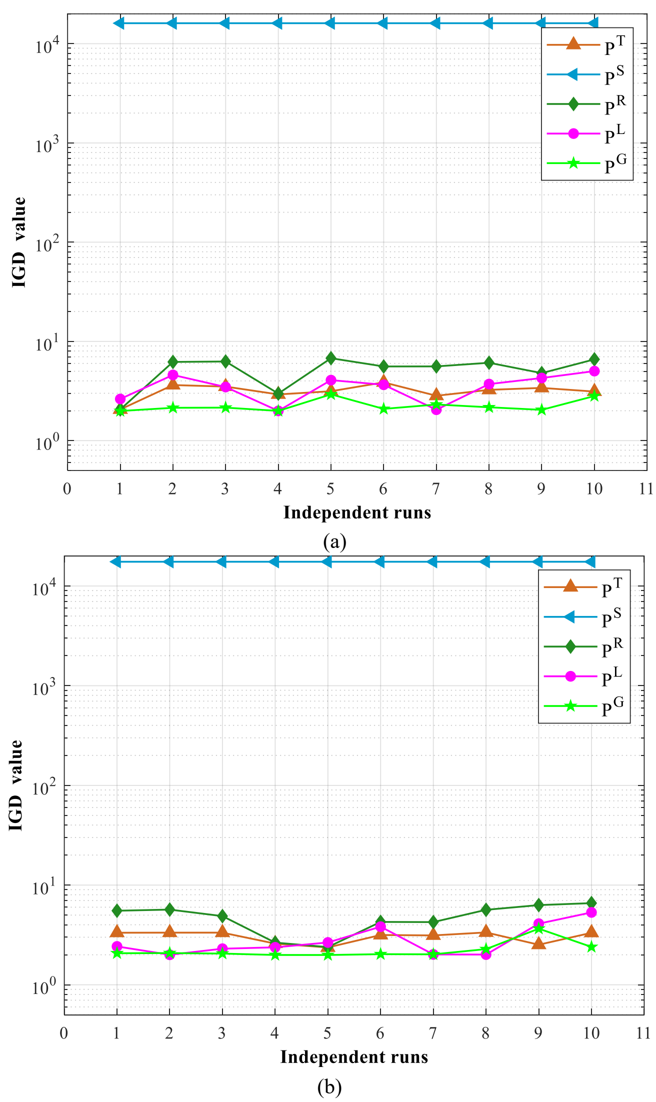

Fig. S1. The IGD value of five new populations after surviving our rank-selection method once, denoted by $\mathrm{P^T}$, $\mathrm{P^S}$, $\mathrm{P^R}$, $\mathrm{P^L}$ and $\mathrm{P^G}$ respectively. (a) rank-selection method is "Value-sort"; and (b) rank-selection method is "Dominated-indicator".

Furthermore, to make a comparison of different nonlinear characteristic problems and validate the effectiveness of a nonlinear activation function, the 3-objective 100-dimensional DTLZ1-3, LSMOP2, SDTLZ2 and MaF2 instances are also used in the experiment. After a population is initialized, first-generation subpopulations are generated by $(\boldsymbol{y}-\boldsymbol{y}\otimes\boldsymbol{r}')\times\mathrm{T}$ and $f_\sigma((\boldsymbol{y}-\boldsymbol{y}\otimes\boldsymbol{r}')\mathrm{T})$, where f_σ can be sigmoid. The resulting subpopulations are denoted by Linear and Sigmoid respectively. Fig. S2 shows the solution distribution of subpopulations mentioned above. Obviously, the solution distribution generated using a nonlinear activation function is better in most cases.

To make a further discussion of nonlinear mapping model, we perform more experiments on 3-objective 10-dimensional DTLZ1-7 problems. In the experiments, population is initialized randomly and the population size is set to 105. We choose two representative nonlinear mapping methods, i.e., Radial Basis Function Neural Network and Gaussian Process. These two nonlinear methods together with Activation Function-assisted Inverse Mapping $f_\sigma((\boldsymbol{y}-\boldsymbol{y}\otimes\boldsymbol{r}')\mathrm{T})$ (f_σ can be sigmoid) are used as the mapping methods to get the first-generation new population. The IGD of these population are shown in Table S-I, denoted by GP, RBFNN and Activation, respectively.

From Table S-I, we can see that when D (decision variable dimension) is relatively small, the solutions generated using our nonlinear inverse model is worse than those with Gaussian Process but is better than Radial Basis Function Neural Network. In other words, Gaussian Process is indeed more excellent than our nonlinear inverse model when D (decision variable dimension) is small.

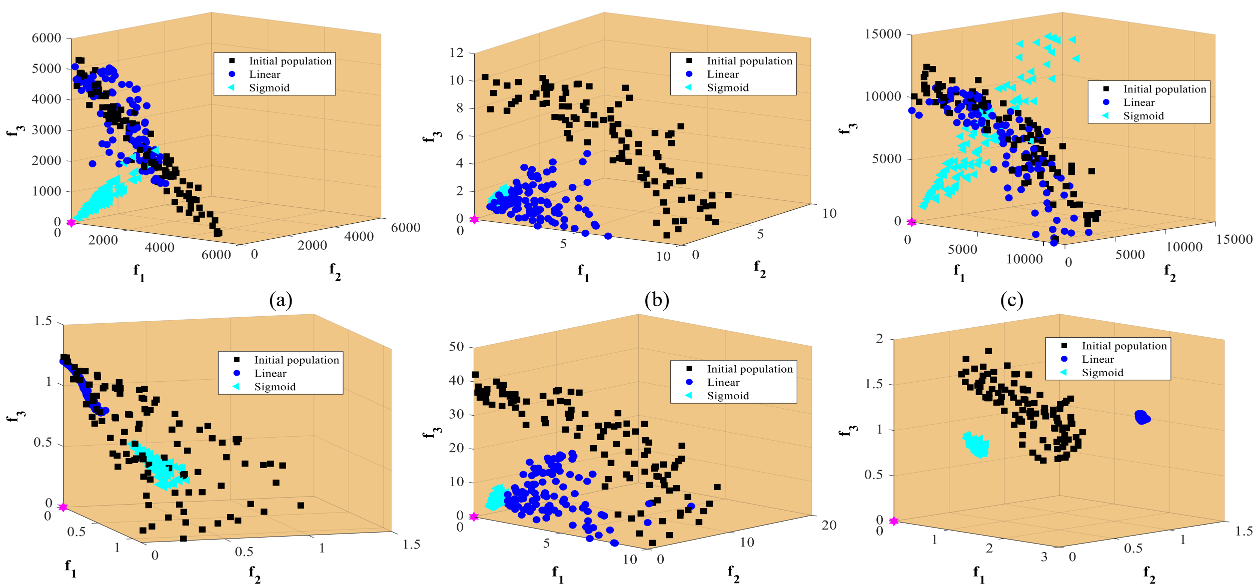

Fig. S2. The solution distribution generated using linear mapping $(\boldsymbol{y}-\boldsymbol{y}\otimes\boldsymbol{r}')\mathrm{T}$ and nonlinear activation function $f_\sigma((\boldsymbol{y}-\boldsymbol{y}\otimes\boldsymbol{r}')\mathrm{T})$ on a 3-objective 100-dimensional DTLZ1-3, LSMOP2, SDTLZ2 and MaF2 problems, denoted by Linear and Sigmoid respectively. Magenta points are the ideal ones. (a) DTLZ1, (b) DTLZ2, (c) DTLZ3, (d) LSMOP2, (e) SDTLZ2, (f) MaF2.

TABLE S-I
IGD METRIC VALUES OF THE THREE POPULATIONS ON LSMOP1–LSMOP9 PROBLEMS, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

	Training data size is 105(N)		
Problem	GP	RBFNN	Activation
DTLZ1	1.3083E+02(2.68E+01) −	7.2865E+01(5.27E+00) −	**8.5308E+00(2.17E-01)**
DTLZ2	**3.5952E-01(8.39E-02)** +	2.2749E+00(2.58E-01) −	4.5438E-01(3.20E-02)
DTLZ3	3.9519E+02(6.04E+01) −	2.0046E+02(3.76E+01) −	**9.7655E+01(2.67E+00)**
DTLZ4	**7.0851E-01(3.73E-02)** +	2.3747E+00(4.37E-01) −	9.4600E-01(1.86E-02)
DTLZ5	**2.8603E-01(9.03E-02)** +	1.8806E+00(3.07E-01) −	3.0591E-01(5.37E-02)
DTLZ6	6.7590E+00(7.09E-01) +	**3.4452E+00(3.12E-01)** +	7.5992E+00(1.75E-01)
DTLZ7	**7.5403E+00(5.37E-01)** +	1.2925E+01(3.48E+00) +	1.4151E+01(2.78E+00)
+/−/≈	5 / 2 / 0	2 / 5 / 0	/

TABLE S-II

IGD RESULTS OBTAINED BY DIFFERENT ALGORITHMS ON LSMOP1–LSMOP9, AND BOLD INDICATES THE BEST RESULTS

Problem	M	D	MOEA/DVA	LMEA	S^3-CMA-ES	WOF	EAGO	**EAGOA(ours)**
LSMOP1	5	500	1.9918E-01(7.47E-03) −	1.5635E-01(3.93E-03) +	1.9823E-01(8.54E-02) −	2.7661E-01(3.54E-02) −	**1.5497E-01(6.37E-02)**+	1.7780E-01(5.17E-02)
		1000	1.6052E-01(6.87E-03)+	1.7462E-01(2.22E-02)+	3.0621E-01(3.14E-02)+	2.9522E-01(2.36E-02)+	**1.4863E-01(5.26E-02)**+	7.7859E-01(3.46E-02)
	10	500	3.7106E-01(7.36E-02) −	3.8762E-01(6.81E-02) −	**2.8183E-01(6.21E-02)** +	4.9067E-01(5.24E-02) −	2.8823E-01(2.58E-01) +	3.2904E-01(3.70E-02)
		1000	3.4800E-01(5.78E-02) ≈	3.8146E-01(7.42E-02) −	4.7312E-01(4.32E-02) −	5.9192E-01(3.21E-02) −	**3.2069E-01(1.13E-02)** ≈	3.4070E-01(2.02E-02)
LSMOP2	5	500	2.1441E-01(2.16E-03)−	1.5032E-01(3.08E-03) −	1.4113E-01(2.54E-02)−	1.4017E-01(2.64E-03)−	1.5066E-01(4.22E-02)−	**1.3677E-01(5.82E-02)**
		1000	1.8957E-01(6.97E-03) −	1.3856E-01(2.72E-03) −	**1.2012E-01(3.21E-02)** +	1.3521E-01(3.26E-03) ≈	1.6370E-01(1.53E-02)−	1.3518E-01(1.09E-02)
	10	500	4.7619E-01(2.58E-02) −	4.6867E-01(1.63E-01) −	**2.4062E-01(2.87E-02)** +	2.6823E-01(8.31E-03) +	5.0923E-01(5.14E-02) −	2.9087E-01(1.57E-02)
		1000	4.3356E-01(2.35E-02) −	2.6684E-01(1.61E-01) −	**2.0784E-01(2.35E-03)** +	2.4533E-01(2.65E-03) −	2.9656E-01(1.37E-02) −	2.4391E-01(1.36E-02)
LSMOP3	5	500	5.4185E-01(4.55E-02) −	**3.6979E-01(9.90E-04)**+	2.7220E+00(9.31E-01) −	5.5574E-01(5.24E-01) −	4.2223E-01(2.68E-02)−	3.7793E-01(4.25E-02)
		1000	6.9438E-01(7.07E-02)−	3.8518E-01(6.65E-02)−	3.0925E+00(3.24E-01) −	5.2347E-01(6.31E-01)−	1.3970E+00(6.20E-01)−	**3.3695E-01(2.85E-02)**
	10	500	1.7428e+00(3.68E-02) −	1.4623E+00(7.08E-01) −	1.9257E+00(6.21E-01) −	1.0062E+00(9.36E-01) +	**9.3375E-01(2.40E-01)** +	1.3403E+00(2.14E-01)
		1000	1.6334e+00(8.34E-02) +	8.7072E-01(1.34E-01) +	2.7134E+00(2.35E-02) −	8.6159E-01(4.23E-02) +	**4.8732E-01(7.82E-03)** +	2.0341E+00(3.73E-01)
LSMOP4	5	500	1.9378E-01(4.40E-03)+	2.2776E-01(4.95E-02)+	2.1662E-01(3.21E-02)+	2.0142E-01(4.25E-02)+	**1.4827E-01(5.11E-02)**+	2.5843E-01(3.59E-02)
		1000	1.8523E-01(1.71E-03) −	1.7092E-01(8.44E-03) −	2.2106E-01(2.24E-02) −	1.7273E-01(8.36E-02) −	**1.3049E-01(7.80E-02)**≈	1.5844E-01(5.27E-02)
	10	500	4.0636E-01(3.36E-03) −	5.5467E-01(1.48E-02) −	3.2902E-01(4.52E-02) −	3.1795E-01(7.26E-02) −	5.1321E-01(1.94E-02) −	**3.1102E-01(3.06E-02)**
		1000	3.7235E-01(5.32E-02) −	4.0673E-01(1.86E-02) −	**2.5178E-01(1.24E-02)** +	2.7182E-01(4.28E-02) −	3.7010E-01(4.10E-02) −	2.6861E-01(4.97E-02)
LSMOP5	5	500	3.7166E-01(3.92E-02) +	7.4712E-01(9.87E-02) −	6.4394E-01(6.24E-02) −	**3.1467E-01**(3.45E-02) +	3.8812E-01(3.54E-02)+	4.1303E-01(3.56E-02)
		1000	2.8307E-01(4.75E-04)−	2.9276E-01(3.07E-02)−	5.7478E-01(3.67E-02) −	3.7241E-01(7.24E-02)−	4.2730E-01(4.26E-02)−	**2.1135E-01(3.22E-02)**
	10	500	6.2192E-01(2.34E-02) +	8.9686E-01(7.54E-01) −	8.4524E-01(6.28E-02) −	5.9956E-01(4.28E-02) +	**5.1374E-01(2.53E-02)** +	7.0694E-01(2.16E-01)
		1000	5.8421E-01(3.31E-02) −	6.6891E-01(1.10E-01) −	7.3717E-01(4.52E-02) −	8.2863E-01(1.02E-02) −	6.0031E-01(1.03E-02) −	**4.9705E-01(1.88E-01)**
LSMOP6	5	500	1.2273E+00(1.79E-01) +	1.3689E+00(3.39E-01)+	1.3287E+00(4.68E-01)+	**1.0274E+00(4.25E-01)** +	1.9816E+00(1.22E-01)−	1.8188E+00(8.23E-01)
		1000	1.6352E+00(4.09E-01) +	**9.3200E-01(1.51E-01)** +	1.4782E+00(5.21E-01) +	1.1179E+00(1.32E-01) +	4.0794E+00(9.01E-01)−	3.2367E+00(2.15E-01)
	10	500	1.7226E+00(1.52E-01) −	1.7249E+00(1.87E-01) −	1.2475E+00(5.78E-01) −	1.2827E+00(6.57E-01) −	1.2258E+00(2.95E-02) −	**1.1804E+00(3.19E-01)**
		1000	1.3902E+00(1.74E-01) −	1.3082E+00(2.29E-01) −	1.2417E+00(3.14E-01) −	1.0735E+00(2.36E-01) −	1.2049E+00(1.40E-02) −	**1.0577E+00(9.59E-02)**
LSMOP7	5	500	9.4612E-01(2.81E-02) −	5.9449E-01(8.13E-02)−	1.1059E+00(7.14E-01) −	1.0231E+00(4.28E-01) −	1.3051E+00(7.71E-01)−	**5.4243E-01(3.36E-02)**
		1000	9.0375E-01(5.27E-02) −	5.8573E-01(7.21E-02) +	1.1081E+00(3.64E-02) −	1.0804E+00(4.25E-03) −	1.1633E+00(4.63E-01)−	**5.0526E-01(2.29E-02)**
	10	500	8.7059E+00(3.24E-02) −	2.0985E+00(2.97E-01) −	1.5531E+00(6.27E-01) −	**1.2995E+00**(2.74E-01) +	1.8473E+00(4.03E-01) −	1.5009E+00(1.29E-01)
		1000	4.7198e+00(1.26E-02) −	**1.2239E+00(2.83E-01)** +	1.3323E+00(4.21E-02) +	1.5247E+00(2.35E-02) +	1.4792E+00(5.25E-01) +	2.7476E+00(9.93E-01)
LSMOP8	5	500	2.5087E-01(6.19E-02) +	2.4855E-01(2.81E-03)+	1.1126E+00(6.52E-01) −	3.3993E-01(5.31E-02) −	**2.4556E-01(4.28E-02)**+	3.0531E-01(5.61E-02)
		1000	2.7103E-01(5.51E-03) +	2.9536E-01(1.36E-02)+	1.1083E+00(2.24E-02) −	3.1812E-01(2.35E-02) −	**2.3091E-01(3.56E-02)**+	3.0242E-01(4.01E-02)
	10	500	6.4933E-01(2.69E-03) −	4.5501E-01(6.21E-03) −	7.1305E-01(3.67E-02) −	5.0994E-01(6.21E-02) −	4.9633E-01(4.76E-02) −	**4.3972E-01(2.65E-02)**
		1000	6.3258E-01(4.16E-02) −	7.1259E-01(8.09E-02) −	6.8827E-01(1.47E-01) −	5.3161E-01(3.24E-02) ≈	**5.2694E-01(9.93E-02)** +	5.4022E-01(4.26E-02)
LSMOP9	5	500	**4.6537E-01(1.13E-01)** +	7.6733E-01(1.96E-02) +	9.7673E-01(6.45E-02) +	5.7187E-01(3.57E-02) +	6.9072E-01(2.55E-02)+	1.1930E+00(4.38E-01)
		1000	**4.2847E-01(3.89E-02)** +	7.7786E-01(1.22E-02) +	1.0288E+00(2.34E-01)+	5.7985E-01(2.35E-02) +	1.1022E+00(3.30E-01)+	2.2414E+00(5.05E-01)
	10	500	2.9863E+00(2.64E-02) −	4.1357E+00(9.19E-01) −	4.8774E+00(3.45E-01) −	3.7467E+00(3.61E-01) −	1.7948E+00(6.59E-02) −	**1.5546E+00(2.10E-01)**
		1000	1.7574e+00(3.14E-02) −	3.0989E+00(2.47E-01) −	2.5066E+00(2.14E-02) −	1.1247E+00(1.25E-02) −	2.2605E+00(1.66E-01) −	**9.5812E-01(3.42E-02)**
+/−/≈			**11 / 24 / 1**	**13 / 23 / 0**	**12 / 24 / 0**	**13 / 21 / 2**	**14 / 20 / 2**	/
AvgRank			**3.7778**	**3.7778**	**4.1944**	**3.2500**	**3.1944**	**2.8056**

"+" means the result outperforms EAGOA, "−" means the opposite and "≈" indicates no significant difference.

"AvgRank" represents the average order of algorithms according to the Friedman's test

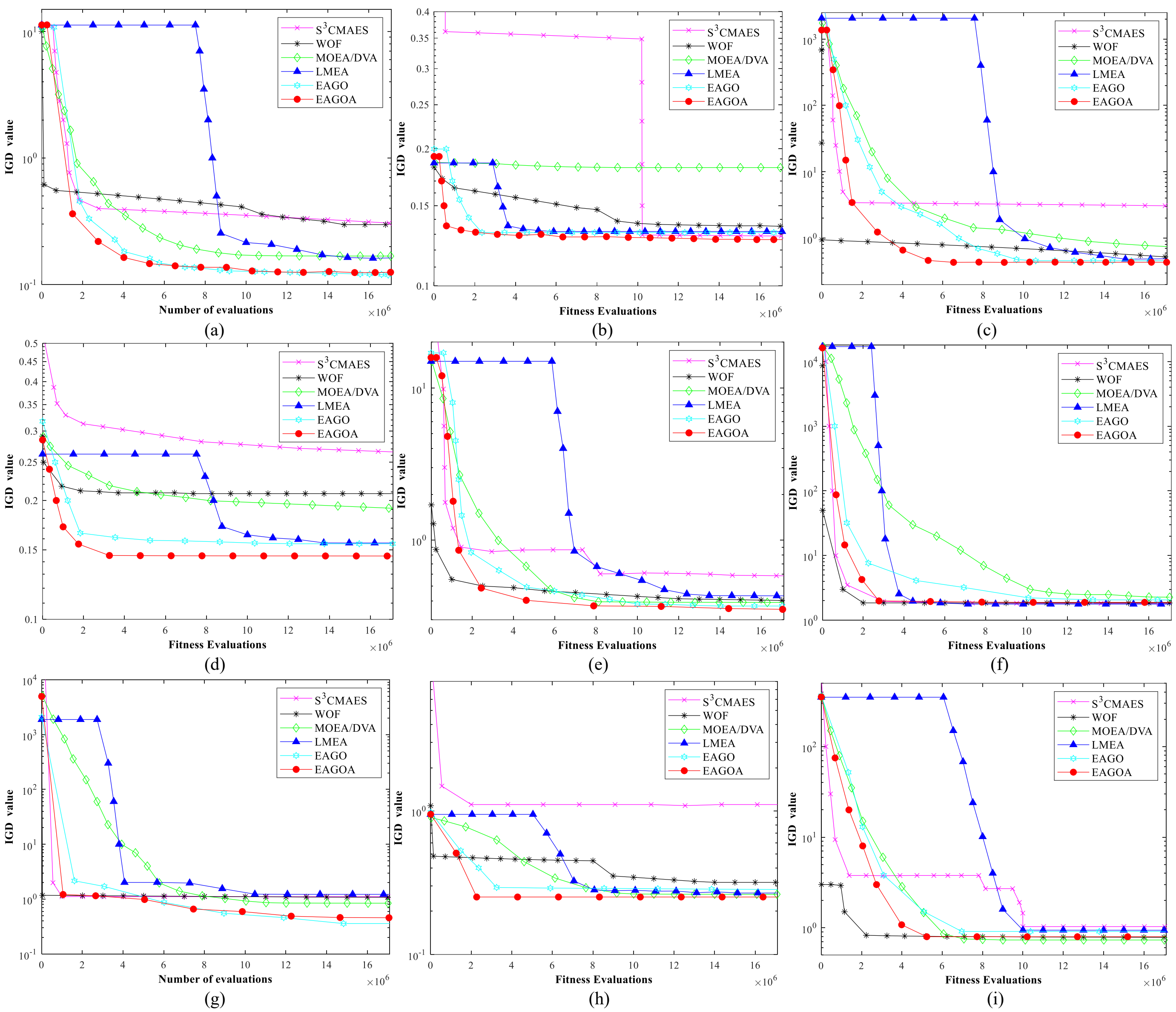

Fig. S3. The IGD convergence iterative process obtained by six algorithms on five-objective 1000-dimensional LSMOP1-9 problems. (a) LSMOP1. (b) LSMOP2. (c) LSMOP3. (d) LSMOP4. (e) LSMOP5. (f) LSMOP6. (g) LSMOP7. (h) LSMOP8. (g) LSMOP9.

B. More Performance Comparison

Table S-II shows the IGD results of six compared algorithms on LSMOP series problems. Fig. S3 shows the convergence process of IGD values of six algorithms under 5-objective and 1000-dimensional LSMOP1-9 problems. The abscissa is fitness evaluations (does not represent the algorithm time), and the ordinate is IGD. It can be clearly seen that EAGOA outperforms than its peers in most cases.

Table S-III shows the statistical results of the IGD values obtained by six algorithms on DTLZ1–DTLZ7 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 EAGOA. Friedman's test is adopted and "AvgRank" represents the average order of algorithms. When D = 100, fitness evaluation count is 1 000 000. It can be seen that the proposed EAGOA outperforms or at least reaches its peers' best level.

TABLE S-III
IGD METRIC VALUES OF SIX ALGORITHMS ON DTLZ1–DTLZ7 PROBLEMS, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	MOEA/DVA	LMEA	S^3-CMA-ES	WOF	EAGO	**EAGOA(ours)**
DTLZ1	5	100	1.0056E+01(5.72E-05) −	**6.2360E-02(4.22E-04)** ≈	5.5986E+02(3.24E+01) −	2.0477E-01(2.36E-02) −	2.8348E-01(4.04E-03) −	6.9006E-02(2.56E-03)
	10	100	3.0555E+01(1.79E-02) −	**1.0524E-01(4.87E-03)** +	4.3005E+02(4.68E+01) −	1.2431E+02(3.41E+01) −	1.3804E-01(7.03E-03) +	1.6688E-01(2.37E-02)
DTLZ2	5	100	2.9042E-01(9.26E-08) −	1.9126E-01(2.14E-03) −	1.9844E-01(3.31E-02) −	1.9499E-01(3.24E-02) −	1.9291E-01(7.50E-05) −	**1.8747E-01(1.71E-02)**
	10	100	6.1598E-01(5.13E-02) −	3.9818E-01(1.48E-02) −	4.1839E-01(2.43E-02) −	4.2554E-01(2.87E-02) −	3.9480E-01(1.30E-04) −	**3.7656E-01(4.51E-02)**
DTLZ3	5	100	2.9242E+01(5.92E-05) −	1.9540E-01(2.14E-03) −	2.3358E+03(3.75E+02) −	2.1673E+00(6.21E-01) −	3.6499E-01(6.54E-03) −	**1.8151E-01(3.78E-02)**
	10	100	1.2187E+02(3.77E-02) −	3.9735E-01(3.77E-02) −	2.2424E+03(2.75E+02) −	4.6357E+02(4.83E+01) −	4.0103E-01(1.91E-03) ≈	**3.8359E-01(3.40E-02)**
DTLZ4	5	100	6.3366E-01(1.29E-01) −	**2.5338E-01(1.55E-02)** +	5.8834E-01(2.74E-02) −	1.9509E-01(8.16E-02) +	3.3175E-01(2.42E-02) +	3.6427E-01(2.52E-02)
	10	100	7.1535E-01(3.33E-02) −	6.7294E-01(2.47e-02) −	6.5566E-01(3.45E-02) −	**4.2389E-01(4.68E-02)** +	6.2721E-01(4.65E-03) −	6.1183E-01(2.34E-02)
DTLZ5	5	100	4.8810E-01(5.06E-04) −	4.3544E-03(1.44E-04) −	4.1256E-03(3.54E-04) −	1.4902E-01(7.22E-02) −	4.3206E-03(7.06E-05) −	**3.6412E-03(1.68E-03)**
	10	100	3.1085E+00(1.87E-04) −	2.5612E-03(6.95E-05) +	3.6875E-02(2.58E-01) −	6.2533E-01(1.69E-02) −	**2.4812E-03(8.00E-05)** +	1.8172E-02(1.45E-03)
DTLZ6	5	100	5.9687E-01(2.43E-06) −	3.9526E-03(2.14E-04) +	6.5184E+01(5.71E+00) −	5.0083E-01(1.82E-02) −	**3.9479E-03(3.68E-04)** +	1.1485E-02(2.34E-03)
	10	100	2.1714E+00(4.09E-02) −	2.2089E-03(5.11E-04) +	6.0823E+01(3.43E+00) −	1.0390E+01(2.86E+00) −	**1.7114E-03(1.40E-05)** +	9.3941E-03(1.46E-03)
DTLZ7	5	100	3.8346E-01(2.51E-06) −	3.2417E-01(1.10E-02) −	3.1684E-01(3.34E-02) −	3.2324E-01(4.15E-02) −	3.2803E-01(2.45E-03) −	**3.1242E-01(1.37E-02)**
	10	100	1.0258E+00(7.84E-02) −	8.9741E-01(6.40E-03) −	1.0308E+00(4.78E-01) −	1.0207E+00(2.37E-01) −	**8.3919E-01(3.02E-02)** +	8.5510E-01(1.46E-02)
+/−/≈			**0 / 14 / 0**	**5 / 8 / 1**	**0 / 14 / 0**	**2 / 12 / 0**	**6 / 7 / 1**	**/**
AvgRank			**5.2857**	**2.5000**	**4.8571**	**3.8571**	**2.5000**	**2**

TABLE S-IV
HV METRIC VALUES OF TWO ALGORITHMS ON LSMOP1–9 PROBLEMS, WHERE THE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	EAGO	**EAGOA(ours)**
LSMOP1	5	1000	**9.1514E-01(7.12E-02)** +	8.6372E-01(3.85E-02)
LSMOP2	5	1000	9.2673E-01(3.48E-02) −	**9.5726E-01(1.45E-02)**
LSMOP3	5	1000	3.2359E-01(5.67E-02) −	**3.8226E-01(2.78E-02)**
LSMOP4	5	1000	9.3651E-01(2.78E-02) ≈	**9.6378E-01(4.77E-02)**
LSMOP5	5	1000	6.7088E-01(2.26E-02) −	**6.7727E-01(4.25E-02)**
LSMOP6	5	1000	**2.0392E-05(9.12E-05)** +	3.1247E-06(1.39E-05)
LSMOP7	5	1000	**5.8637E-01(4.39E-02)** +	4.2541E-01(1.78E-02)
LSMOP8	5	1000	6.2281E-01(3.78E-02) ≈	**6.8961E-01(4.25E-02)**
LSMOP9	5	1000	3.1228E-03(6.48E-02) −	**2.3657E-01(5.18E-02)**
+/−/≈			**3 / 4 / 2**	/

Besides, we present HV values on 5-objective 1000-dimensional LSMOP1-9 problems for a comprehensive comparison. The fitness evaluation count is 17,000,000. Table S-IV shows that EAGOA outperforms EAGO in most test problems.

Next, we use some common activation functions, i.e., *Tanh*, *Sigmoid*, *ReLU*, to replace activation functions we have defined, thus resulting EAGOA-T, EAGOA-S and EAGOA-R, respectively. A numerical comparison has been performed to study the influence of different activation functions. The IGD values for four algorithms are given in Table S-V. It is clear that activation functions we have defined for objective space mapping can help to generate a better offspring population than common activation functions, e.g., *Tanh*, *Sigmoid* and *Relu*.

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

Problem	*M*	*D*	EAGOA-S	EAGOA-T	EAGOA-R	EAGOA
LSMOP1	5	1000	1.3929E-01(3.47E-02) ≈	1.8437E-01(6.12E-02) −	1.4912E-01(2.34E-02) −	**1.3591E-01(1.31E-01)**
LSMOP2	5	1000	1.3851E-01(4.85E-02) ≈	1.4173E-01(1.69E-02) −	1.3712E-01(3.57E-02) ≈	**1.3445E-01(2.36E-03)**
LSMOP3	5	1000	**5.1712E-01(5.02E-02)** +	9.5883E-01(2.69E-02) +	5.5584E-01(2.41E-02) +	1.3886E+00(9.44E-01)
LSMOP4	5	1000	**1.2906E-01(1.67E-02)** +	1.3485E-01(4.35E-02) ≈	1.3072E-01(1.49E-02) +	1.5213E-01(3.11E-02)
LSMOP5	5	1000	8.1522E-01(49.04E-02) −	4.5565E-01(3.79E-02) −	9.5013E-01(4.72E-02) −	**4.1443E-01(7.50E-02)**
LSMOP6	5	1000	2.5593E+00(3.76E-01) −	2.9522E+00(5.17E-01) −	2.0304E+00(4.87E-01) ≈	**1.9572E+00(9.37E-01)**
LSMOP7	5	1000	5.9473E-01(6.48E-02) ≈	**5.7464E-01(6.98E-02)** +	1.4121E+00(8.47E-01) −	6.1441E-01(5.17E-02)
LSMOP8	5	1000	4.0417E-01(3.45E-02) −	2.9505E-01(2.17E-02) −	3.4344E-01(3.65E-02) −	**2.8727E-01(1.02E-02)**
LSMOP9	5	1000	1.1915E+00(3.45E-01) −	1.2085E+00(6.46E-01) −	1.3787E+00(2.45E-01) −	**9.4756E-01(1.23E-01)**
+/−/≈			**2 / 4 / 3**	**2 / 6 / 1**	**2 / 5 / 2**	/
AvgRank			**2.3333**	**2.8889**	**2.8889**	**1.8889**

C. Ablation Experiments

The transition matrix T remains unchanged all the time in EAGOA and it depends on the initially generated population. To explore what will happen if we update T by using the later population, we add some new experiments and try updating T in each iteration. This new algorithm can be named as EAGOA-T. Experimental results in terms of IGD value of both EAGOA and EAGOA-T on 5-objective 500D DTLZ1-7 problems are shown in Table S-VI. It can be seen from Table S-VI that after updating T in each iteration, the overall performance of EAGOA-T is worse than EAGOA, especially on DTLZ1 and DTLZ3 (almost an order of magnitude difference). It means updating T does have a negative effect on performance improvement in our EAGOA.

The matrix T indeed depends on the initially generated population because the first population is randomly initialized and the distribution in the decision space is uniform. In other words, the solutions generated by T are more uniform or diverse. If we update T by using the later population, these population are more centered near the Pareto front and our mapping may be limited into a thin area, the possibility of generating a more diverse population will be smaller. Our main idea is to generate high-quality offspring more efficiently by using nonlinear inverse mapping model. So, we don't update T in our EAGOA.

TBLE S-VI
IGD VALUES OF TWO ALGORITHMS ON DTLZ1–DTLZ7 PROBLEMS, WHERE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	EAGOA-T	EAGOA
DTLZ1	5	500	4.1252E+01(2.78E+00) –	**4.4887E+00(5.83E-01)**
DTLZ2	5	500	**2.0775E-01(1.53E-02)** ≈	2.0914E-01(5.17E-02)
DTLZ3	5	500	1.2459E+02(5.78E+01) –	**1.6017E-01(2.65E-02)**
DTLZ4	5	500	2.9591E-01(1.95E-02) –	**2.7113E-01(1.29E-02)**
DTLZ5	5	500	4.9462E-02(3.78E-03) –	**1.1472E-02(7.01E-03)**
DTLZ6	5	500	2.1523E-02(1.67E-03) –	**1.3302E-02(5.48E-03)**
DTLZ7	5	500	**3.0724E-01(8.57E-02)** +	3.1101E-01(4.57E-02)
+/−/≈			1 / 5 / 1	/

Next, we remove T of EAGOA and use a traditional crossover-mutation to generate offsprings. This new algorithm is named as EAGOA-CM. Experimental results in terms of IGD value of both EAGOA and EAGOA-CM on 5-objective 500D DTLZ1-7 problems are shown in Table S-VII.

From Table S-VII, it is very clear that the overall performance of EAGOA-CM is much worse than that of EAGOA and EAGOA almost have an advantage of an order of magnitude difference in each case. These experimental results can verify the effectiveness of our activation Function-assisted nonlinear mapping.

TABLE S-VII
IGD VALUES OF TWO ALGORITHMS ON DTLZ1–DTLZ7 PROBLEMS, WHERE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	EAGOA-CM	EAGOA
DTLZ1	5	500	1.0021E+04(4.58E+03) –	**4.4887E+00(5.83E-01)**
DTLZ2	5	500	1.9692E+01(1.83E+00) –	**2.0914E-01(5.17E-02)**
DTLZ3	5	500	2.5097E+03(4.58E+02) –	**1.6017E-01(2.65E-02)**
DTLZ4	5	500	1.1142E+00(1.68E-01) –	**2.7113E-01(1.29E-02)**
DTLZ5	5	500	9.6827E-01(3.67E-02) –	**1.1472E-02(7.01E-03)**
DTLZ6	5	500	4.1899E+02(1.58E+01) –	**1.3302E-02(5.48E-03)**
DTLZ7	5	500	1.4413E+00(2.47E-01) –	**3.1101E-01(4.57E-02)**
+/−/≈			0 / 7 / 0	/

We also try removing the decision variable analysis of EAGOA and name the new algorithm as EAGOA-U. Experimental results in terms of IGD values of EAGOA and EAGOA-U on 5-objective 500D LSMOP1-9 problems are shown in Table S-VIII.

After removing the decision variable analysis, the overall performance of EAGOA-U is worse than EAGOA. It means the decision variable analysis does have an effect on performance improvement in our EAGOA.

Our main purpose is to improve inverse mapping's function expression ability and thus generate high-quality offspring more efficiently on the premise of ensuring a rapid construction of feasible bridge. To further improve performance, we introduce decision variable analysis of LMEA.

TABLE S-VIII
IGD VALUES OF TWO ALGORITHMS ON LSMOP1–LSMOP9 PROBLEMS, WHERE BEST RESULTS ON EACH TEST INSTANCE ARE BOLDED

Problem	*M*	*D*	EAGOA-U	EAGOA
LSMOP1	5	500	8.9387E-01(3.58E-02) –	**1.5102E-01(2.33E-02)**
LSMOP2	5	500	1.5347E-01(2.91E-02) –	**1.3832E-01(1.78E-02)**
LSMOP3	5	500	9.5873E-01(5.43E-02) –	**3.6545E-01(9.96E-02)**
LSMOP4	5	500	2.6955E-01(6.18E-02) –	**1.7955E-01(1.96E-02)**
LSMOP5	5	500	8.0322E-01(3.06E-02) –	**4.0823E-01(5.94E-03)**
LSMOP6	5	500	3.2029E+00(2.90E-01) –	**1.2912E+00(4.25E-01)**
LSMOP7	5	500	1.1662E+00(6.49E-01) –	**6.7163E-01(8.56E-02)**
LSMOP8	5	500	7.0018E-01(1.65E-02) –	**2.4564E-01(2.49E-02)**
LSMOP9	5	500	3.1205E+00(2.39E-01) –	**1.2367E+00(3.26E-01)**
+/−/≈			0 / 9 / 0	/

TABLE S-IX
HV VALUES OF THREE ALGORITHMS ON VEHICLESAFETY PROBLEM, WHERE BEST RESULTS ON EACH TEST INSTANCE ARE IN BOLD

Problem	*M*	*D*	LERD	FLEA	LSMaODE	EAGO	EAGOA(ours)
FEWN	5	567	6.0296E-01 (6.37E-02)–	5.6842E-01 (5.10E-02)–	8.0911E-01 (3.94E-02)–	6.6219E-01 (7.85E-02)–	**8.6239E-01 (5.11E-02)**

The HV value is normalized to a range of 0-1, with larger values being better.

III. Real-World Application

We have now performed experiments on a Large Scale Many-Objective Problem in Food–Energy–Water Nexus (FEWN) [1]-[3].

The Food–Energy–Water Nexus (FEWN) problem includes water (surface water, groundwater, desalinated water, wastewater reuse, recycled water reuse), energy (petroleum, hydro, wind, solar, biofuels) and food (cereals, tubers, vegetables, fruits, food products.) resources as well final demand for these resources. Optimizing FEWN is achieved by minimizing resource consumption and at the same time maximizing resource production. Minimizing the intensity coefficients implies a simultaneous reduction in the consumption of a resource and increasing the production of another resource: (A low resource intensity implies using less to produce more)

(1) Minimize f1 (water for energy intensity),
(2) Minimize f2 (water for food intensity),
(3) Minimize f3 (energy for water intensity),
(4) Minimize f4 (energy for food intensity),
(5) Minimize f5 (food for energy intensity).

FEWN consumption data [6] is about resource consumption within the nexus and the final demand that includes households, government, rest of the economy, losses, storage and exports, including 315 variables. We can increase the number of resources in each category from 5 to 7 and obtain 567 variables. So this FEWN problem has 5 objectives and 567 decision variables. Its result is shown in Table S-IX, which shows that our algorithm has an advantage in performance and we can design a more efficient food–Energy–Water Nexus system and make better policies to promote sustainable consumption and production of resources in human livelihood.

References

[1] I. Okola, E. Omulo, D. Ochieng, G. Ouma, "A comparison of evolutionary algorithms on a Large Scale Many-Objective Problem in Food–Energy–Water Nexus," *Results in Control and Optimization*, vol. 10, no. 100195, 2023.

[2] A. Karnib, "Bridging science and policy in water-energy-food nexus: Using the Q-nexus model for informing policy making," *Water Resources Management*, vol. 32, pp. 4895–4909, 2018.

[3] S. Du, G. Liu, H. Li, W. Zhang, R. Santagata, "System dynamic analysis of urban household food-energy-water nexus in Melbourne (Australia)," *Journal of Cleaner Production*, vol. 379, no. 134675, 2022.