Supplementary File of "Matching-Based Selection with Incomplete Lists for Decomposition Multi-Objective Optimization"

Mengyuan Wu, Ke Li, Sam Kwong, *Fellow, IEEE*, Yu Zhou and Qingfu Zhang, *Fellow, IEEE*

Abstract

The paper entitled "Matching-Based Selection with Incomplete Lists for Decomposition Multi-Objective Optimization" introduces the concept of incomplete preference lists into the stable matching model to remedy the loss of population diversity. In particular, each solution is only allowed to maintain a partial preference list consisting of its favorite subproblems. We implement two versions of stable matching-based selection mechanisms with incomplete preference lists: one achieves a two-level one-one matching and the other obtains a many-one matching. Furthermore, an adaptive mechanism is developed to automatically set the length of the incomplete preference list for each solution according to its local competitiveness. Due to the page limits of the paper, we present some additional figures, tables and discussions in this supplementary file.

I. Comparative Algorithms

This section briefly describes the characteristics of the comparative algorithms used in our experiments.

- *MOEA/D-IR*: its selection process is guided by the interrelationship between subproblems and solutions, which is built upon their mutual preferences. In particular, to further emphasize the population diversity, the niche count is added into the solution's preference setting.
- *gMOEA/D-AGR*: by specifying a particular replacement neighborhood as a set of closest solutions, it finds the appropriate solution for each subproblem. In particular, it employs a sigmoid function to dynamically control the size of the replacement neighborhood for all subproblems.
- *MOEA/D-M2M*: it decomposes the original MOP into a number of simplified MOPs and uses several independent NSGA-II procedures to solve these simplified MOPs in a collaborative manner.
- *MOEA/D-DRA*: it improves the original MOEA/D with a dynamically allocation of computation resources. For the problems with complicated PSs, we use TCH aggregation function and the reproduction operators are the same as introduced in Section III-F2 of the paper. For problems with more than three objectives, we use the penalty-based boundary intersection (PBI) aggregation function and replace the DE operator with the SBX operator in view of reported competitive performance [1]–[3].
- *HypE*: it uses the individual hypervolume contribution to assign fitness to each solution. To speed up the hypervolume calculation, it uses the Monte Carlo simulation to estimate this contribution when $m > 5$.
- *NSGA-III*: it is an improved NSGA-II for solving problems with more than three objectives. Instead of the crowding distance used in NSGA-II, it uses a set of uniformly distributed reference points to adjust the population density. In particular, solutions associated with a less crowded reference point have a higher priority to survive to the next generation.
- *PICEA-g*: it co-evolves a family of decision maker's target vectors, sampled in the objective space, together with a population of candidate solutions for solving problems with more than three objectives. It assigns fitness values to the target vectors according to the number of satisfied solutions; while the fitness values of solutions are evaluated as the number of satisfied target vectors.
- *MOEA/DD*: considering the complementary effects of Pareto- and decomposition-based techniques, it combines them in a single paradigm. In addition, to further emphasize the population diversity, it gives the dominated solutions lying on a less crowded area a second chance to survive to the next generation.

II. Performance Comparisons on MOP Instances

Fig. 1 to Fig. 7 presents the plots of the final solution sets found by all nine algorithms with the best IGD values from the 51 runs on MOP test instances.

M. Wu, S. Kwong and Q. Zhang are with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong SAR (e-mail: mengyuan.wu@my.cityu.edu.hk, cssamk@cityu.edu.hk, qingfu.zhang@cityu.edu.hk).

K. Li is with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, EX4 4QF, UK (e-mail: k.li@exeter.ac.uk).

Y. Zhou is with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong SAR and with the College of Computer Science and Software Engneering, Shenzhen University, Shenzhen 518060, China (e-mail: yzhou57-c@my.cityu.edu.hk).

III. Performance Comparisons on UF Instances

Fig. 8 to Fig. 17 presents the plots of the final solution sets found by all nine algorithms with the best IGD values from the 51 runs on UF test instances.

IV. Performance Comparisons on Bi-Objective WFG Instances

Fig. 18 to Fig. 26 presents the plots of the final solution sets found by all nine algorithms with the best IGD values from the 51 runs on bi-objective WFG test instances. Table I and Table II shows the IGD and HV results of all nine algorithms.

V. Performance Comparisons on Many-Objective Problems

In order to further investigate the scalability of MOEA/D-AOOSTM and MOEA/D-AMOSTM, we choose the widely used WFG benchmark suite for empirical studies. Due to the lack of appropriate sample points from the true PF, we only consider the HV metric in this study. Table III to Table VI present the experimental results for 3-, 5-, 8- and 10-objective WFG test instances, respectively. For the 3-objective WFG instances, MOEA/D-AOOSTM performs significantly better in 34 out of 45 comparisons and loses in 2 comparisons to PICEA-g on WFG3 and WFG6. Meanwhile, MOEA/D-AMOSTM has significantly better results in 36 out of 45 comparisons and significantly worse results in 3 comparisons. In particular, MOEA/D-AOOSTM has obtained the best mean metric values on WFG4, WFG5, WFG7 and WFG9 and the second best mean metric values on WFG1, WFG2 and WFG8. Whereas, MOEA/D-AMOSTM ranks the first on WFG1, WFG2 and WFG8 and ranks the second on WFG3 to WFG7. NSGA-III obtains the third final rank on its overall performance. MOEA/D-DRA and HypE are the worst algorithms in this comparison, which might be due to the lack of normalization scheme. In contrast, PICEA-g and MOEA/DD, which does not apply normalization, perform moderately among the seven algorithms. For the WFG instances with 5 objectives, MOEA/D-ASTM gets the best mean HV values on WFG1, WFG4, WFG5, WFG7 to WFG9 and achieves the best overall ranking followed by MOEA/D-AMOSTM. In terms of the Wilcoxon's rank sum test, MOEA/D-AOOSTM wins 43 out of 45 comparisons and is only beaten by PICEA-g once. MOEA/D-AMOSTM keeps being the top two algorithms except for WFG9. PICEA-g, NSGA-III and MOEA/DD rank after the two proposed algorithms. When the number of objectives increases to 8 and 10, the performance of MOEA/D-AOOSTM and MOEA/D-AMOSTM are statistically better in 72 and 73 out of 90 comparisons respectively. They are always the best and second best algorithms on WFG1 and WFG4 to WFG9 except that NSGA-III significantly outperformed them on the 8-objective WFG5 and WFG6. It also worth noting that MOEA/D-AOOSTM is significantly better than MOEA/D-AMOSTM on all 8- and 10-objective WFG problems except for WFG2 and WFG3, where MOEA/D-AMOSTM performs better. This is probably because MOEA/D-AOOSTM maintains better diversity than MOEA/D-AMOSTM since MOEA/D-AOOSTM tries to meet the needs of subproblems that do not find promising solution in their local area in the second-level stable matching.

References

[1] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints," *IEEE Trans. Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[2] H. Ishibuchi, N. Akedo, and Y. Nojima, "Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems," *IEEE Trans. Evolutionary Computation*, vol. 19, no. 2, pp. 264–283, 2015.

[3] K. Li, K. Deb, Q. Zhang, and S. Kwong, "An evolutionary many-objective optimization algorithm based on dominance and decomposition," *IEEE Trans. Evolutionary Computation*, vol. 19, no. 5, pp. 694–716, 2015.

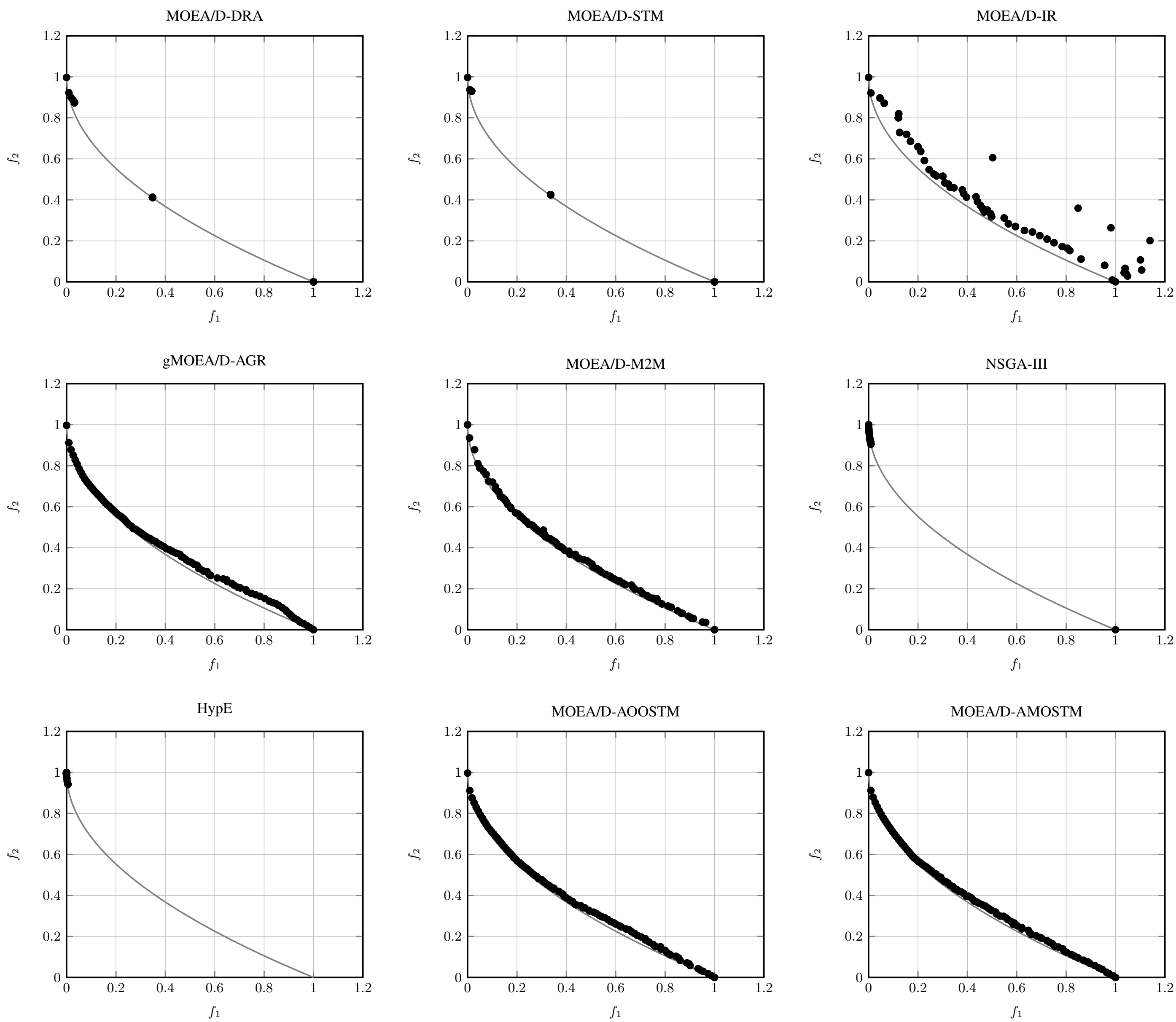

Fig. 1: Final solution sets with best IGD metric values found by 9 MOEAs on MOP1.

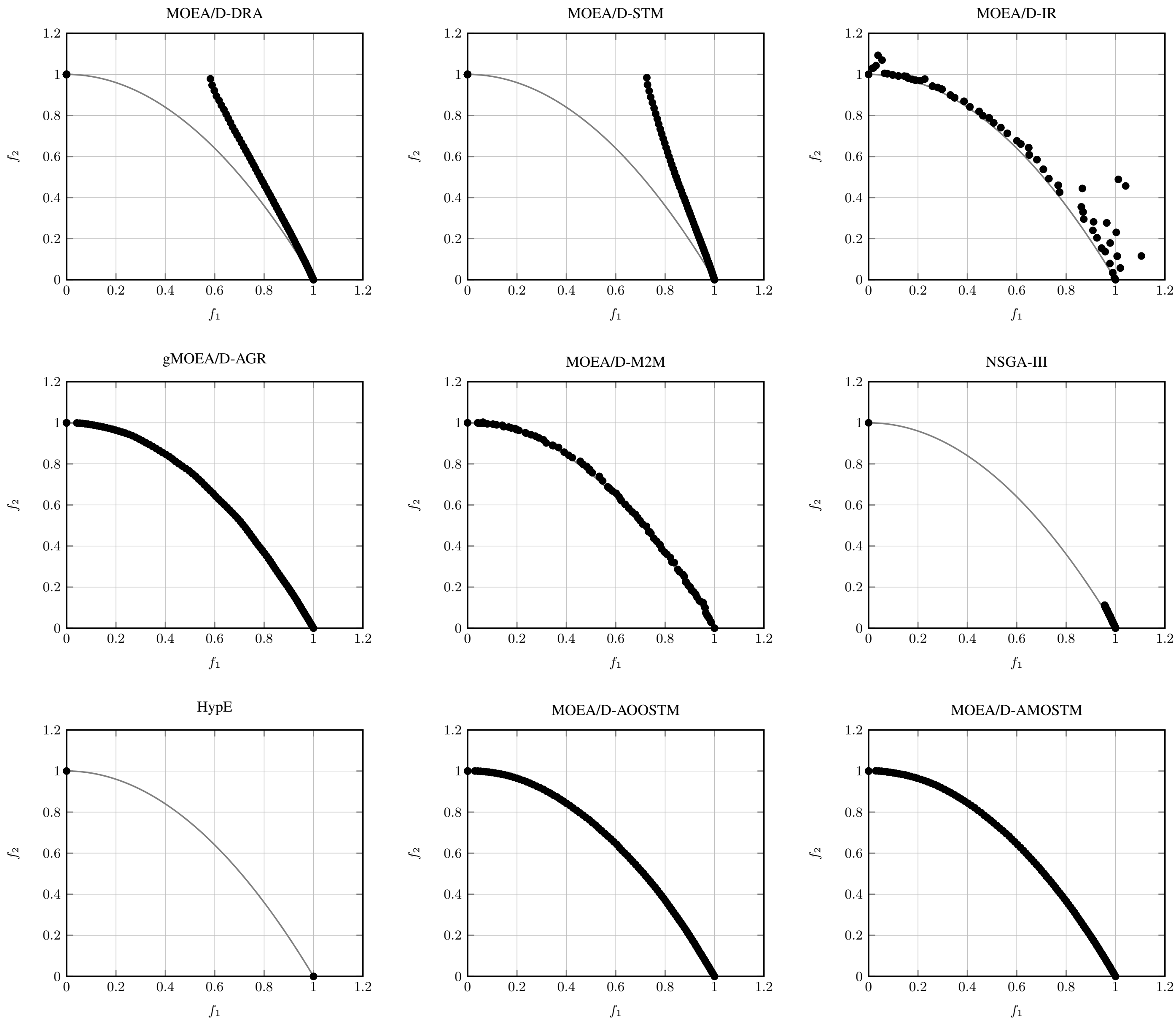

Fig. 2: Final solution sets with best IGD metric values found by 9 MOEAs on MOP2.

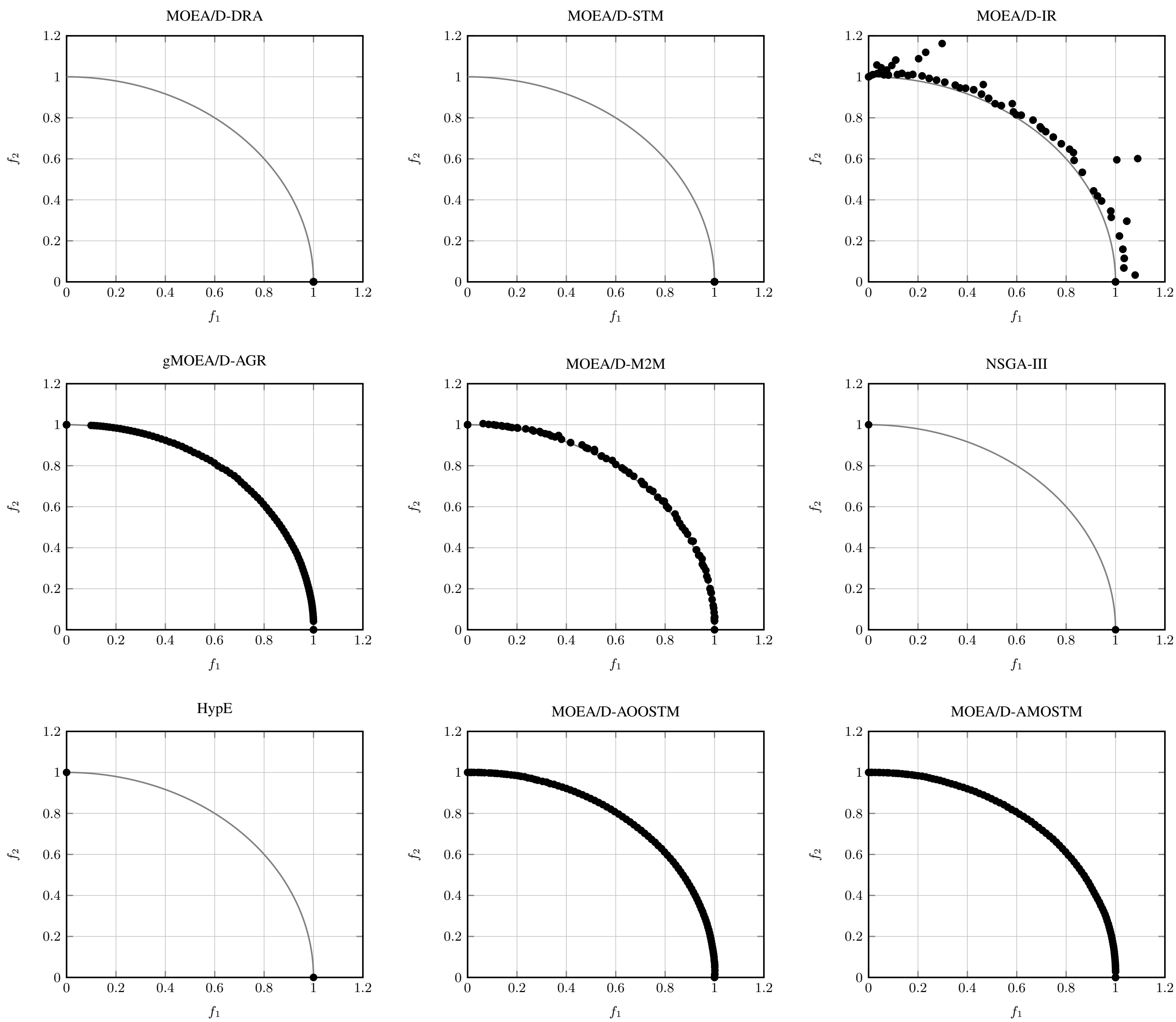

Fig. 3: Final solution sets with best IGD metric values found by 9 MOEAs on MOP3.

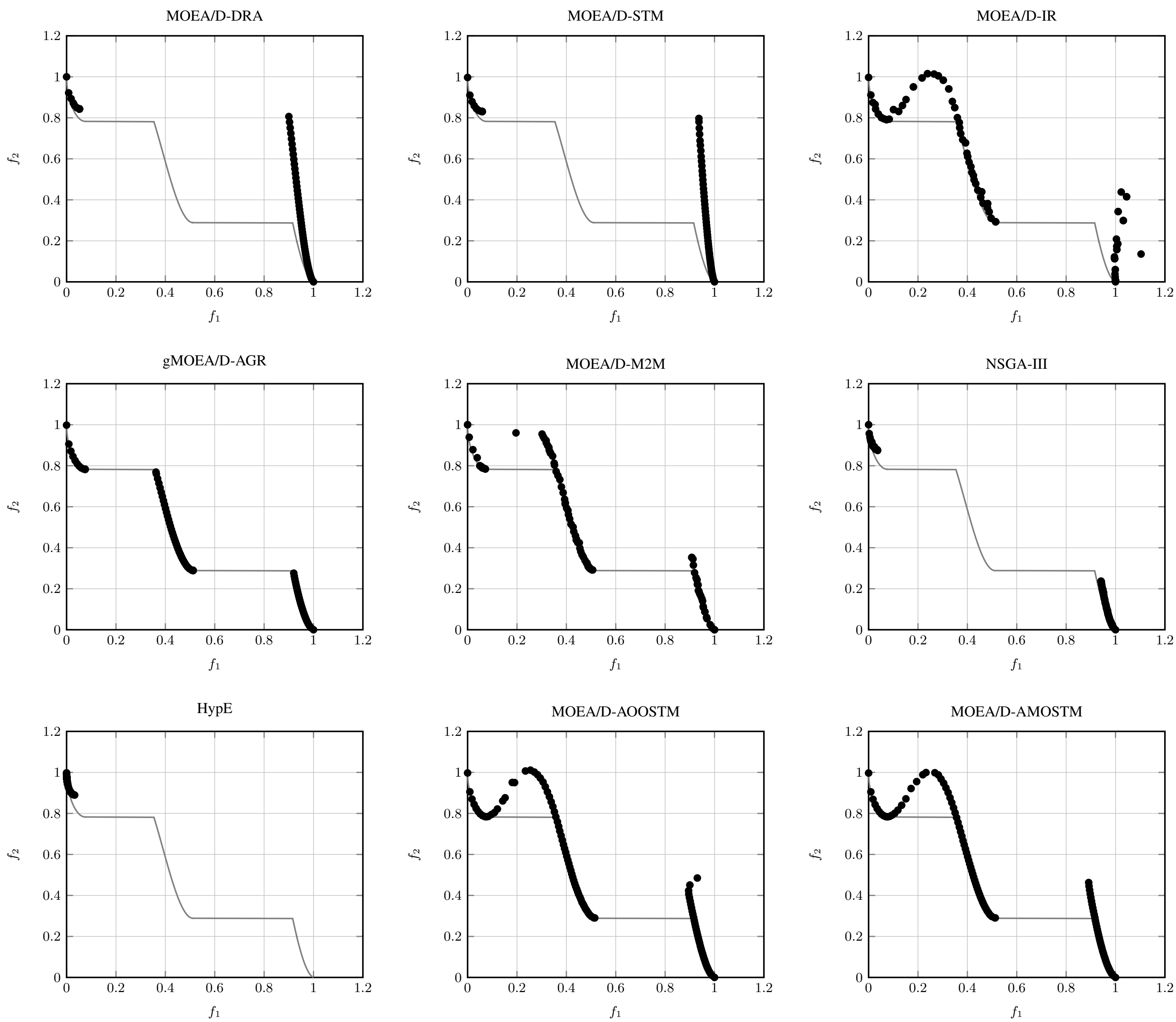

Fig. 4: Final solution sets with best IGD metric values found by 9 MOEAs on MOP4.

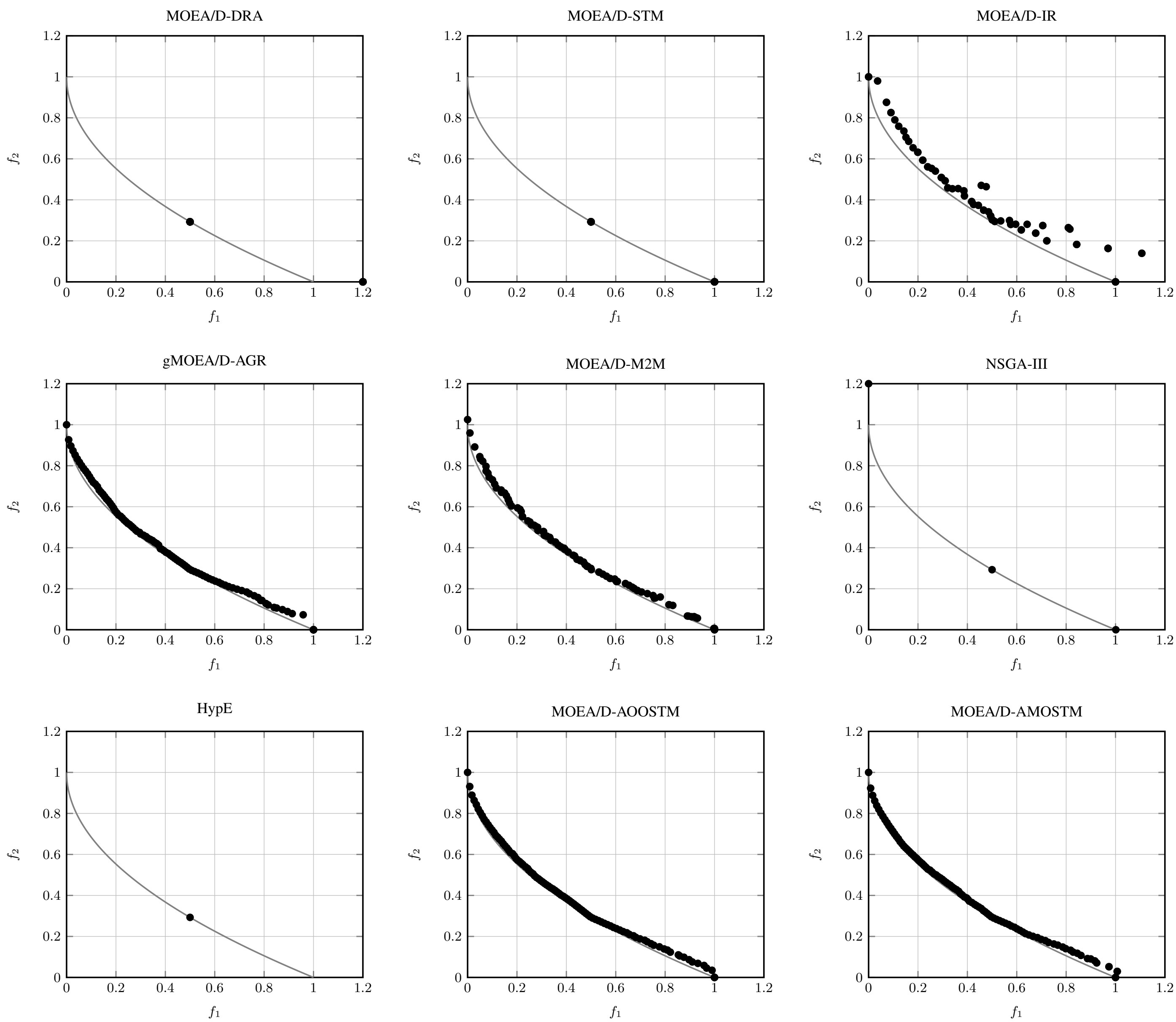

Fig. 5: Final solution sets with best IGD metric values found by 9 MOEAs on MOP5.

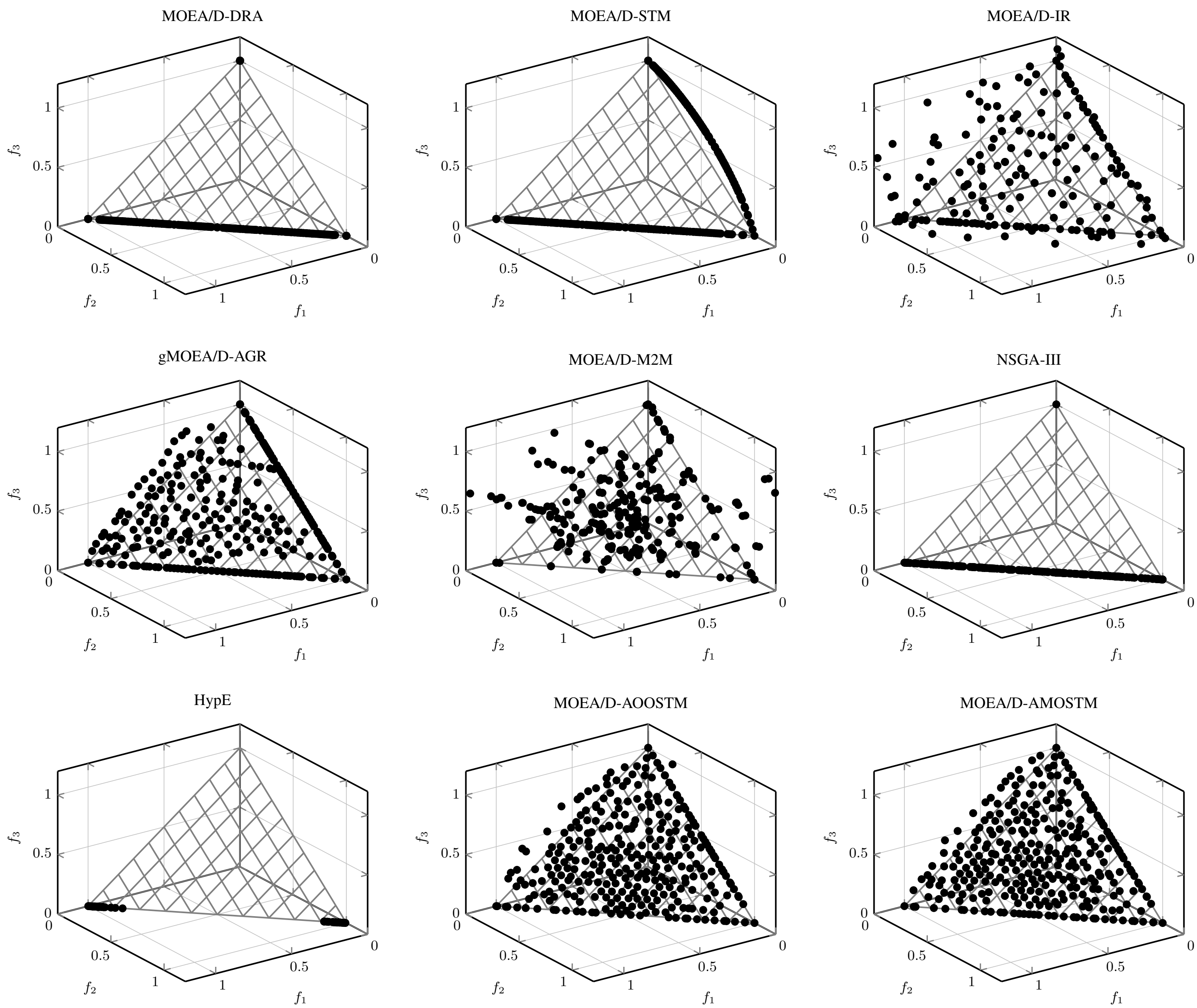

Fig. 6: Final solution sets with best IGD metric values found by 9 MOEAs on MOP6.

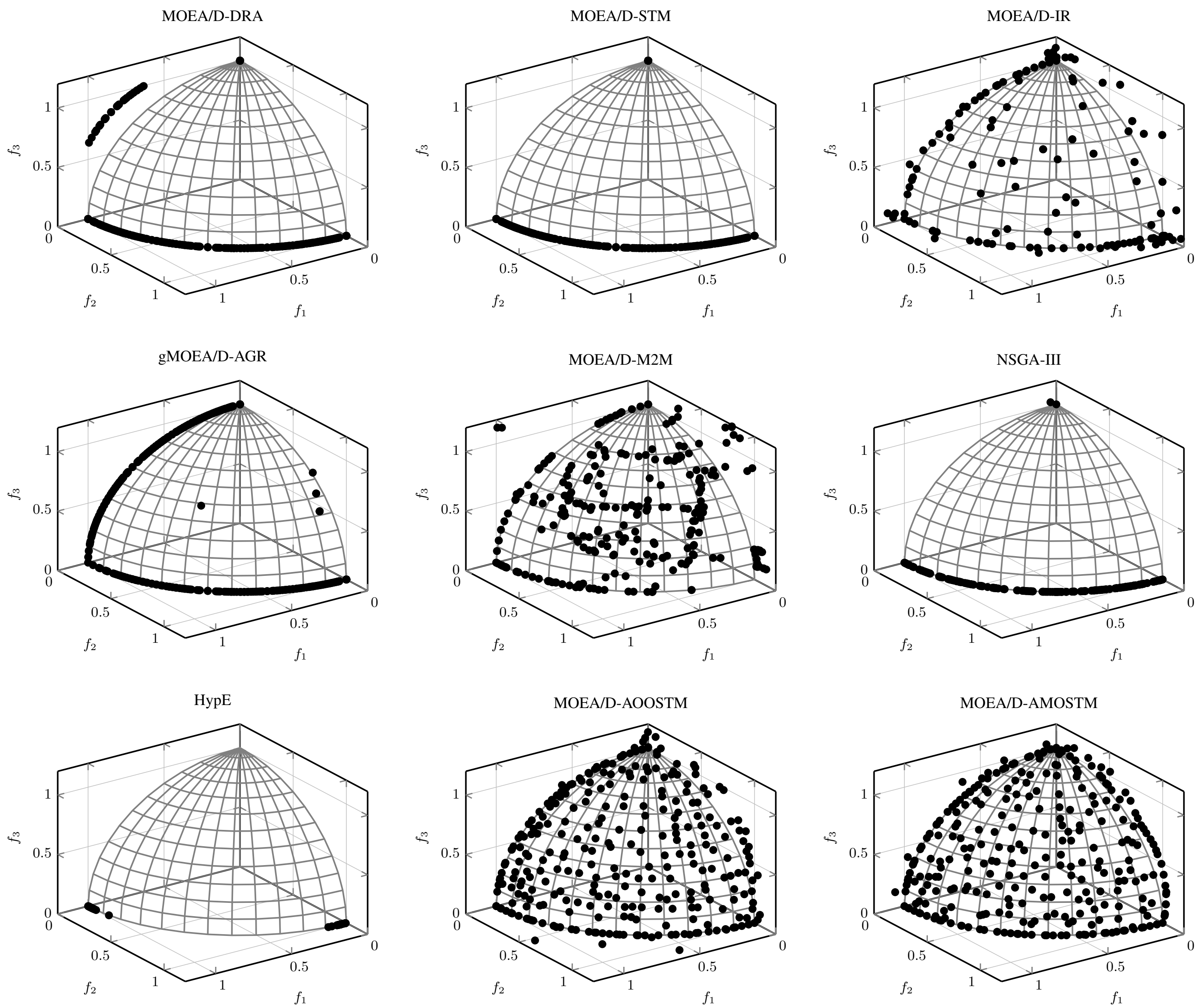

Fig. 7: Final solution sets with best IGD metric values found by 9 MOEAs on MOP7.

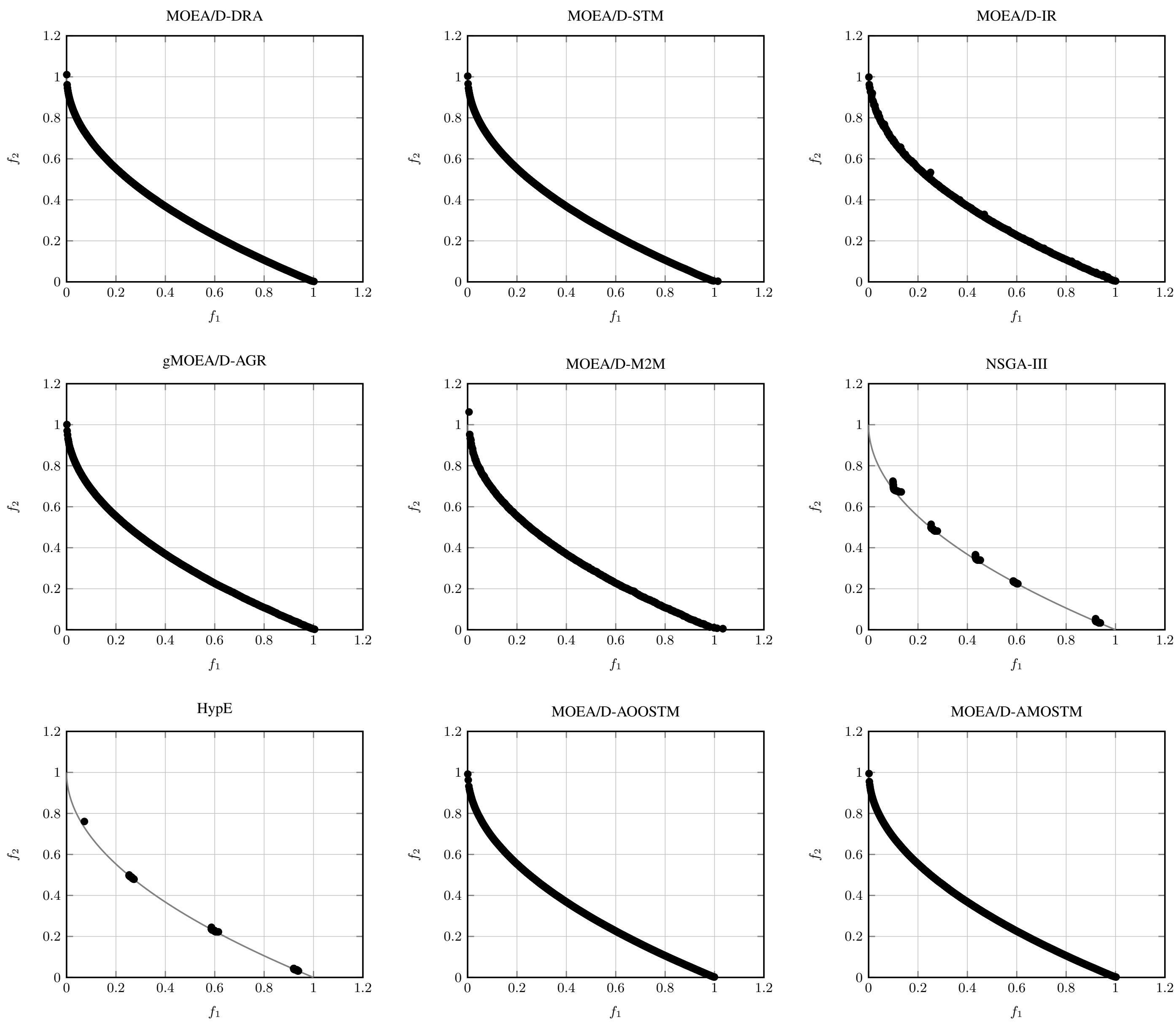

Fig. 8: Final solution sets with best IGD metric values found by 9 MOEAs on UF1.

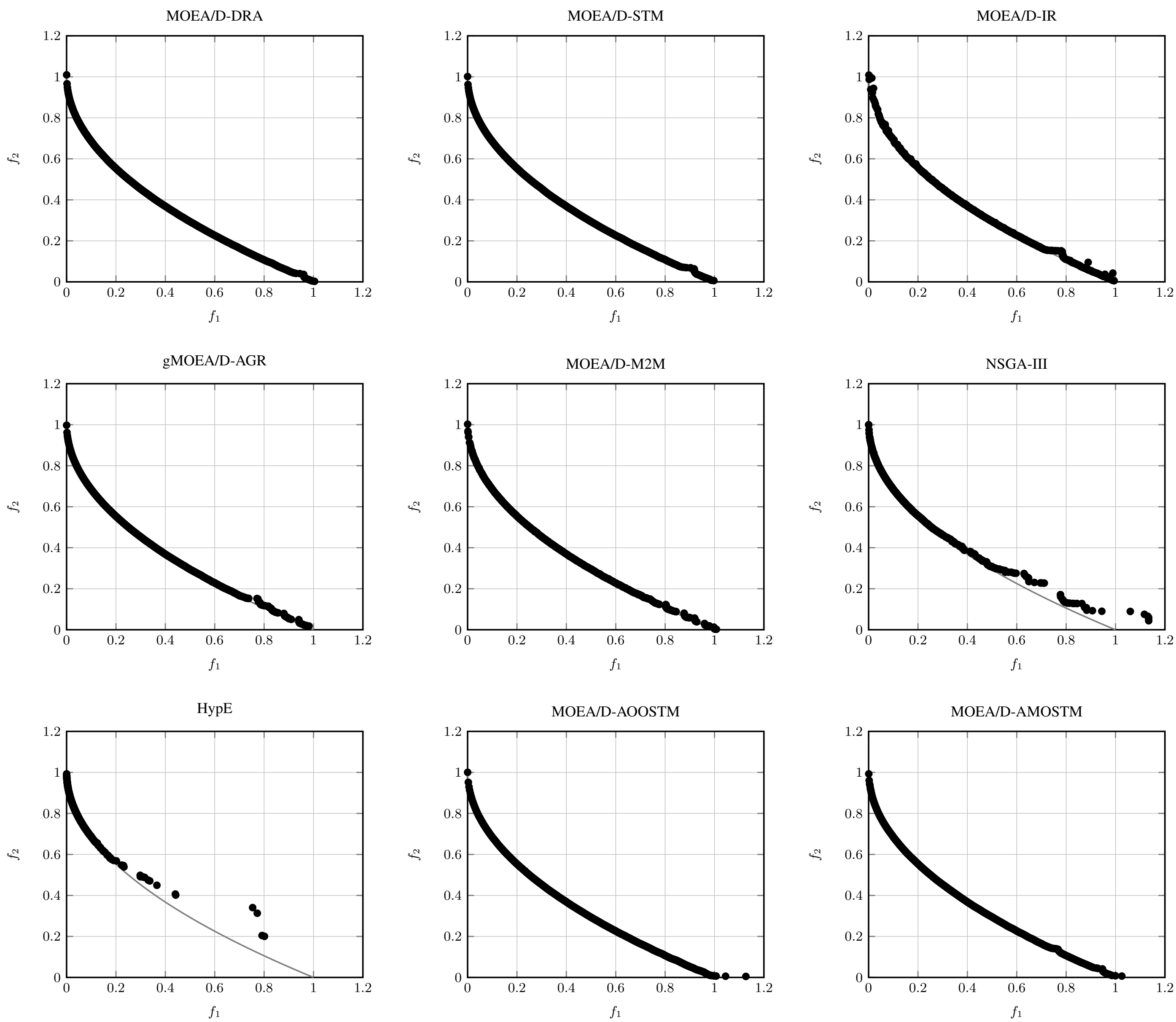

Fig. 9: Final solution sets with best IGD metric values found by 9 MOEAs on UF2.

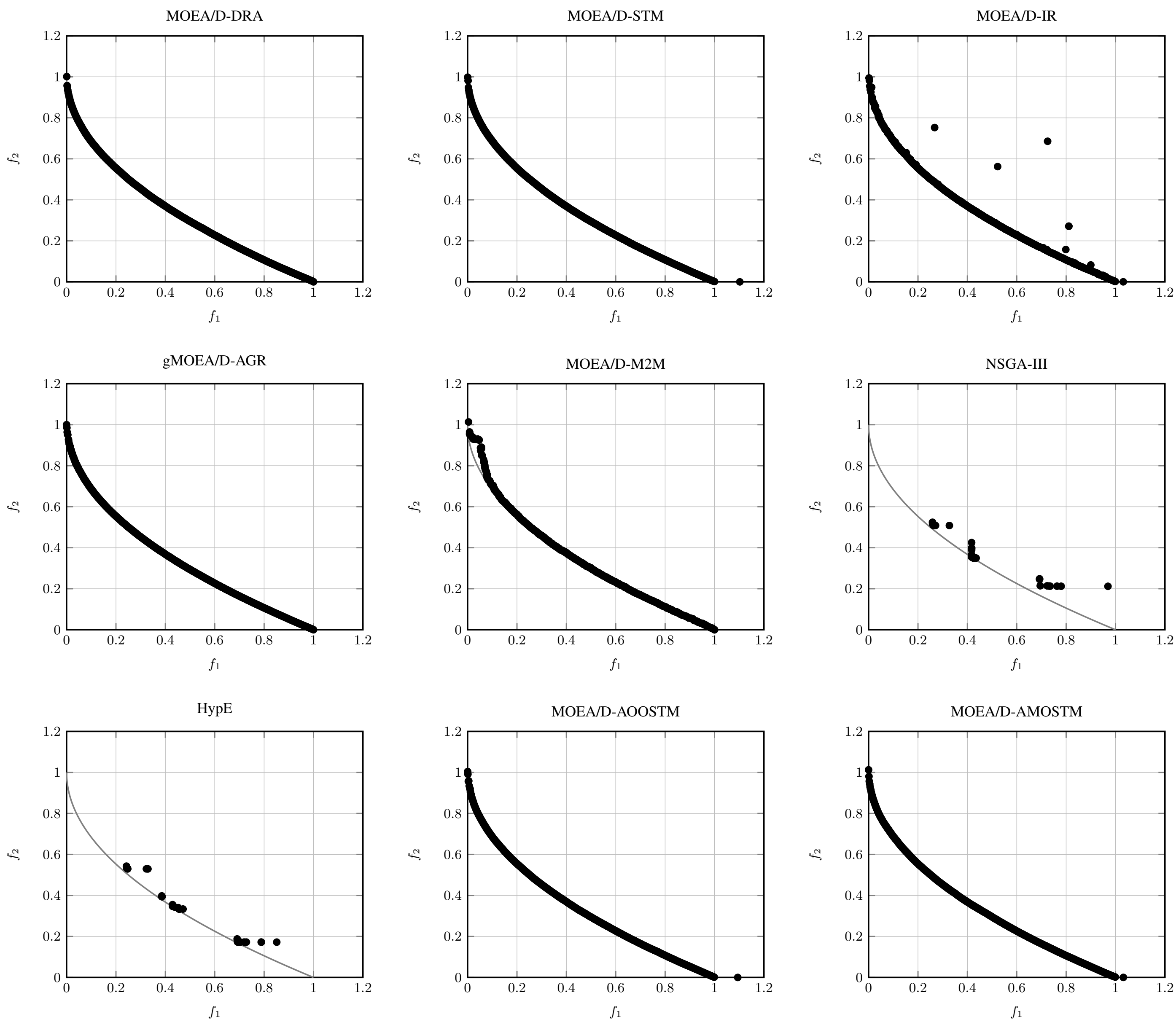

Fig. 10: Final solution sets with best IGD metric values found by 9 MOEAs on UF3.

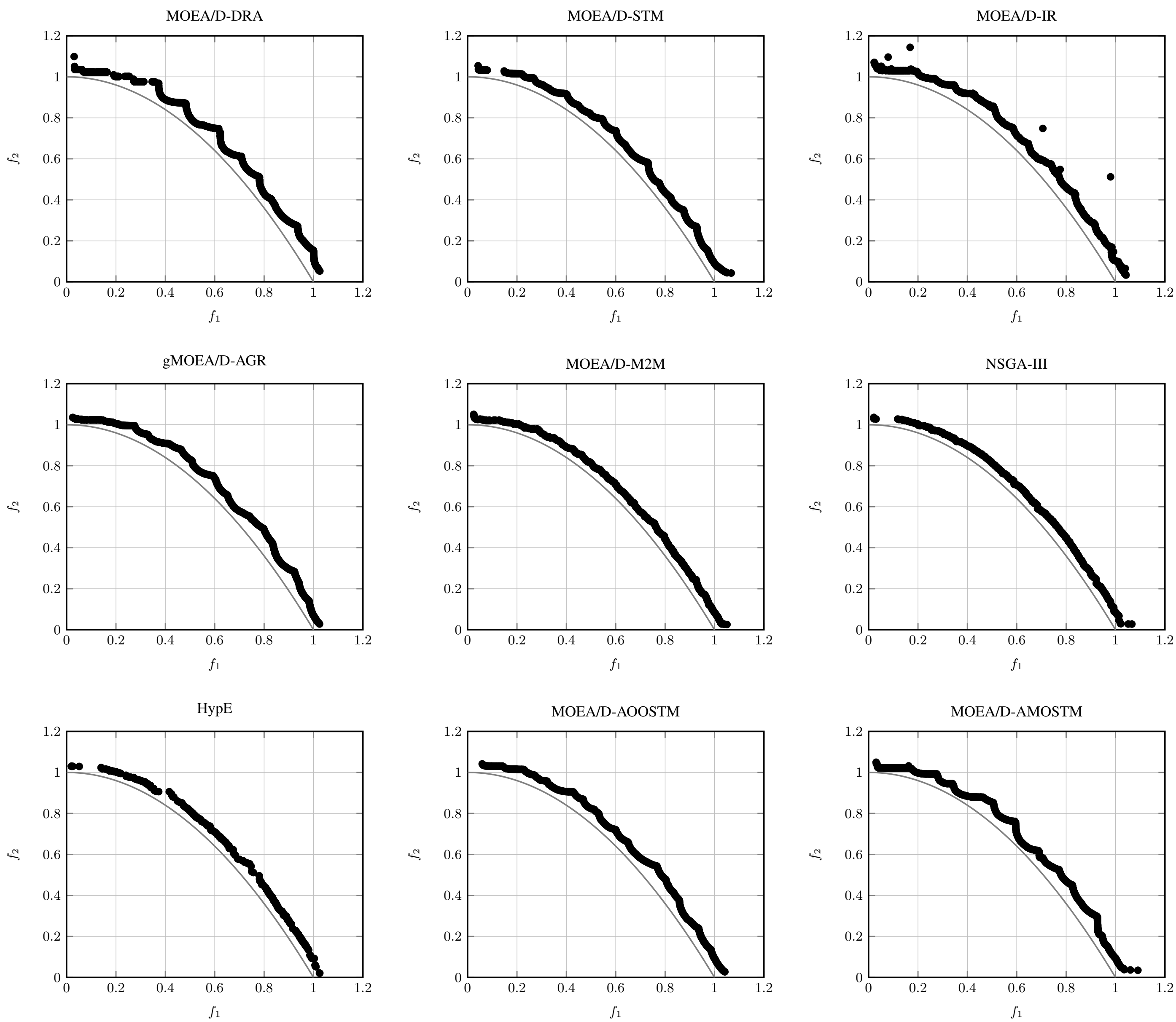

Fig. 11: Final solution sets with best IGD metric values found by 9 MOEAs on UF4.

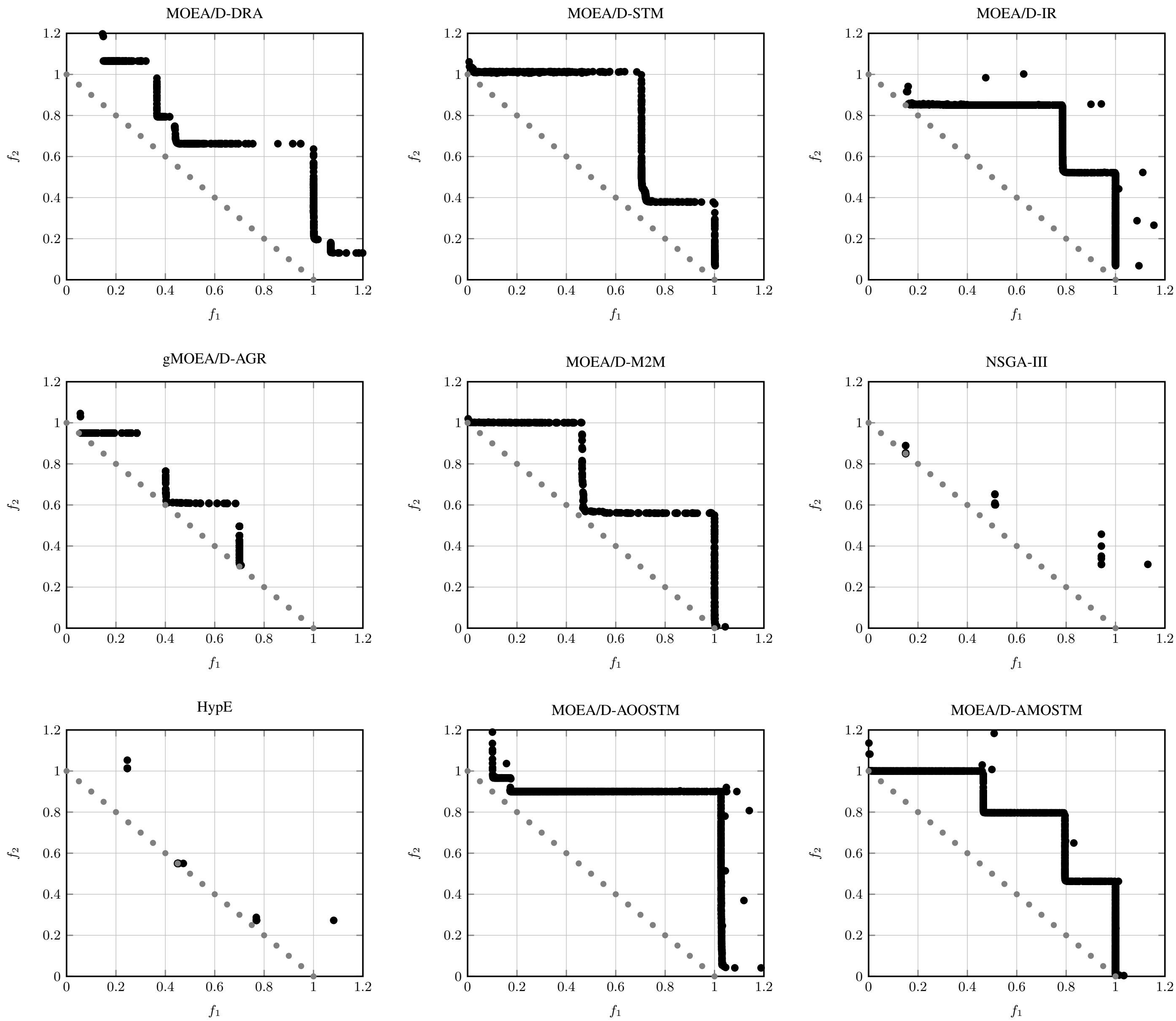

Fig. 12: Final solution sets with best IGD metric values found by 9 MOEAs on UF5.

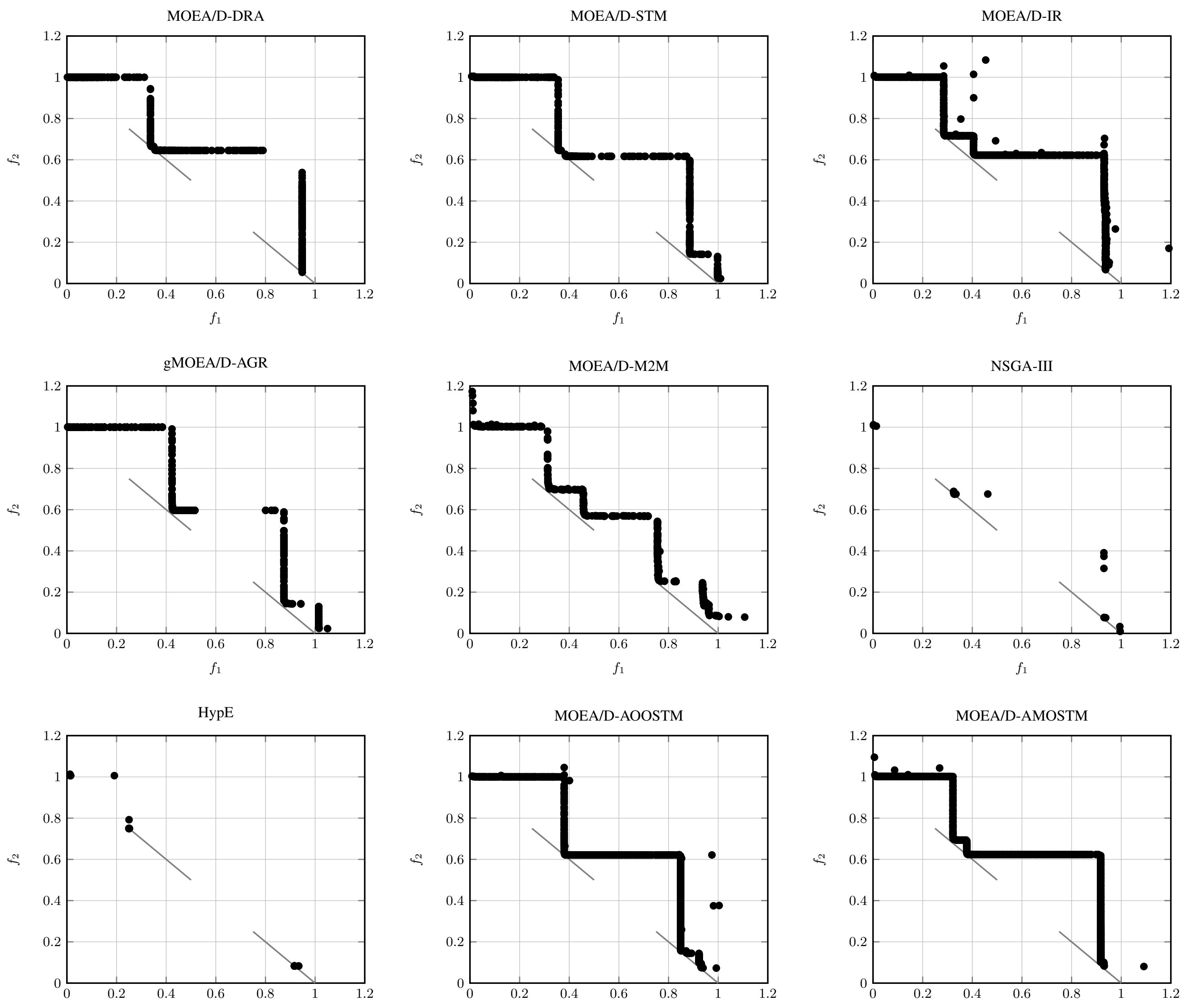

Fig. 13: Final solution sets with best IGD metric values found by 9 MOEAs on UF6.

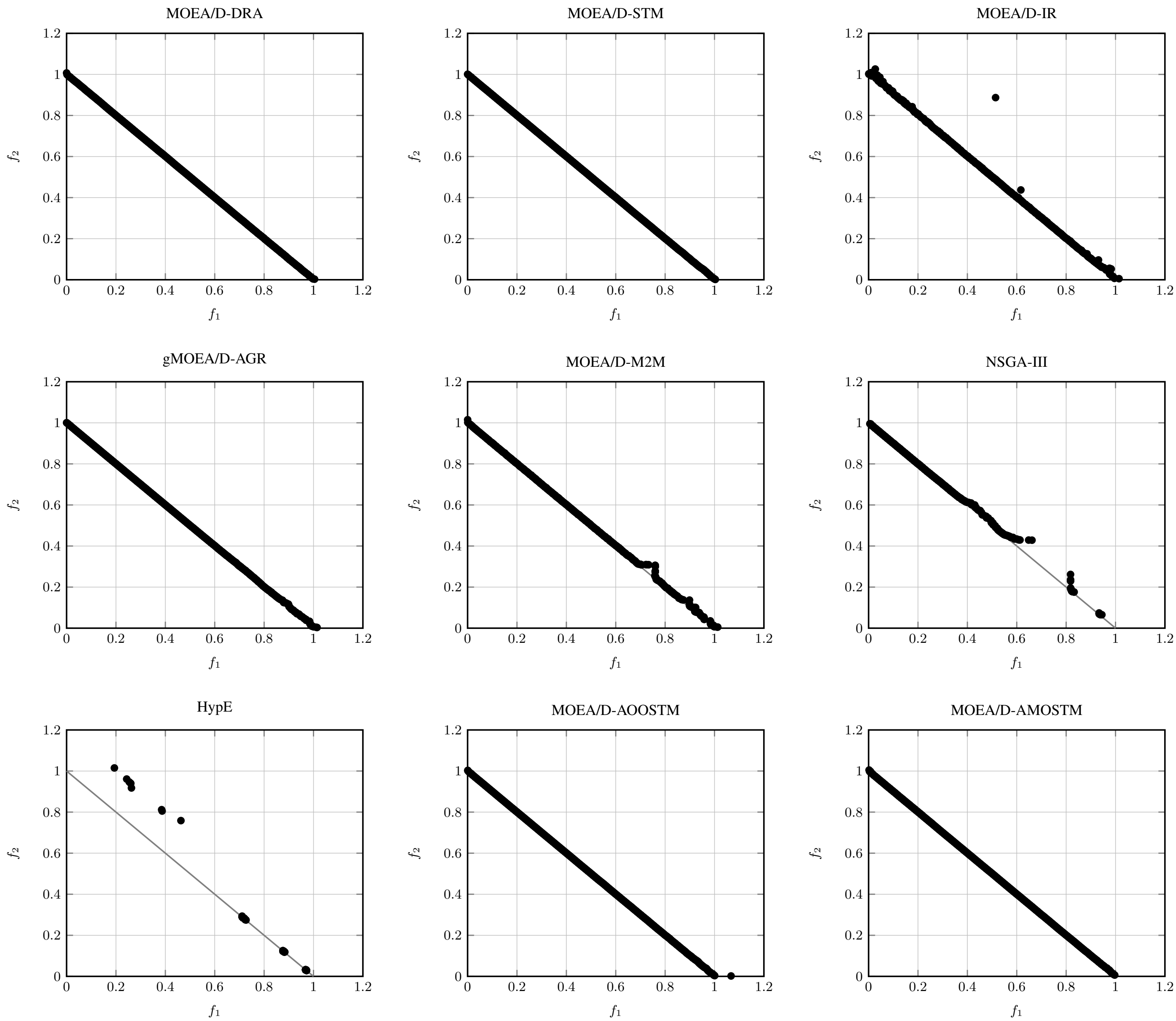

Fig. 14: Final solution sets with best IGD metric values found by 9 MOEAs on UF7.

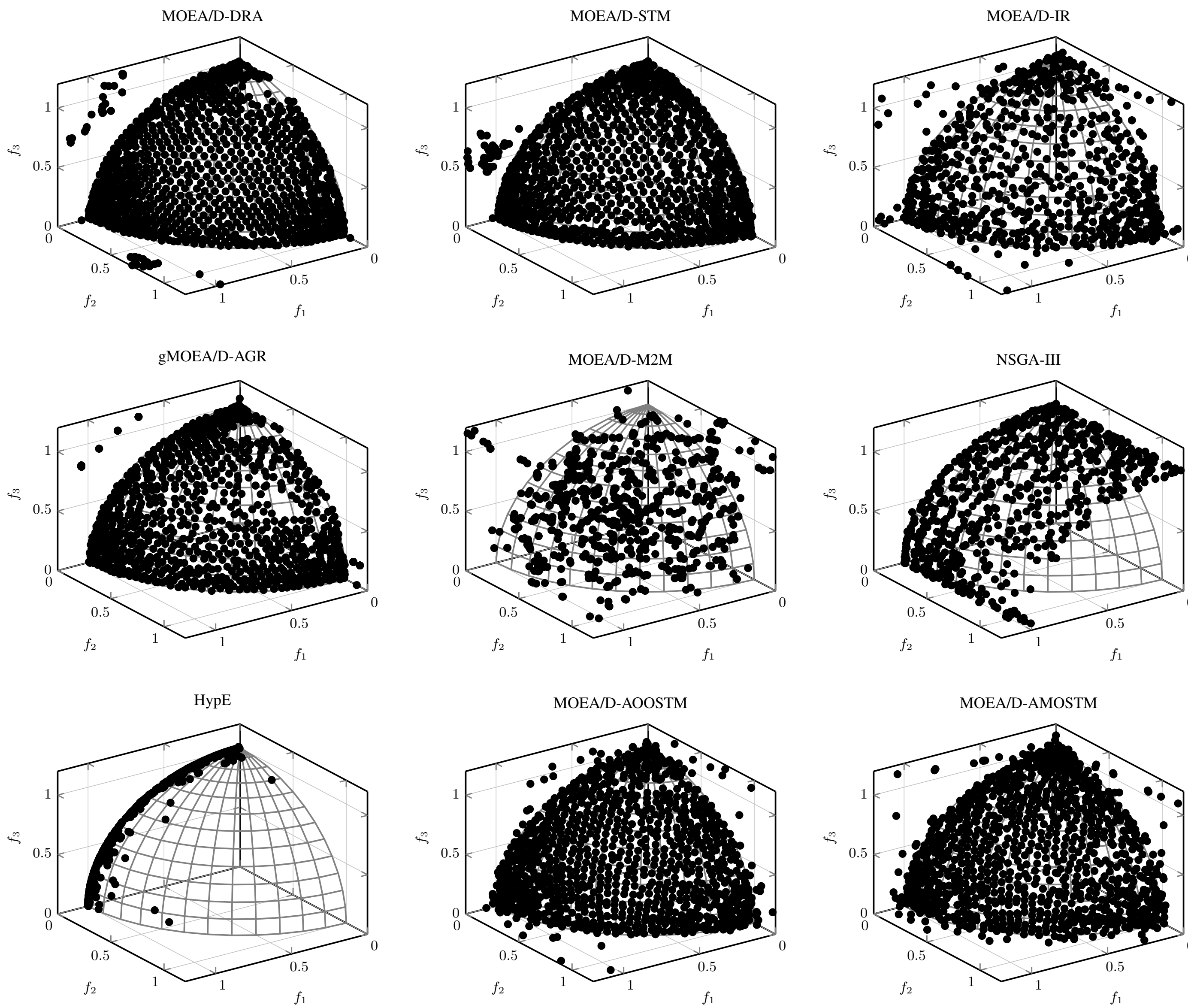

Fig. 15: Final solution sets with best IGD metric values found by 9 MOEAs on UF8.

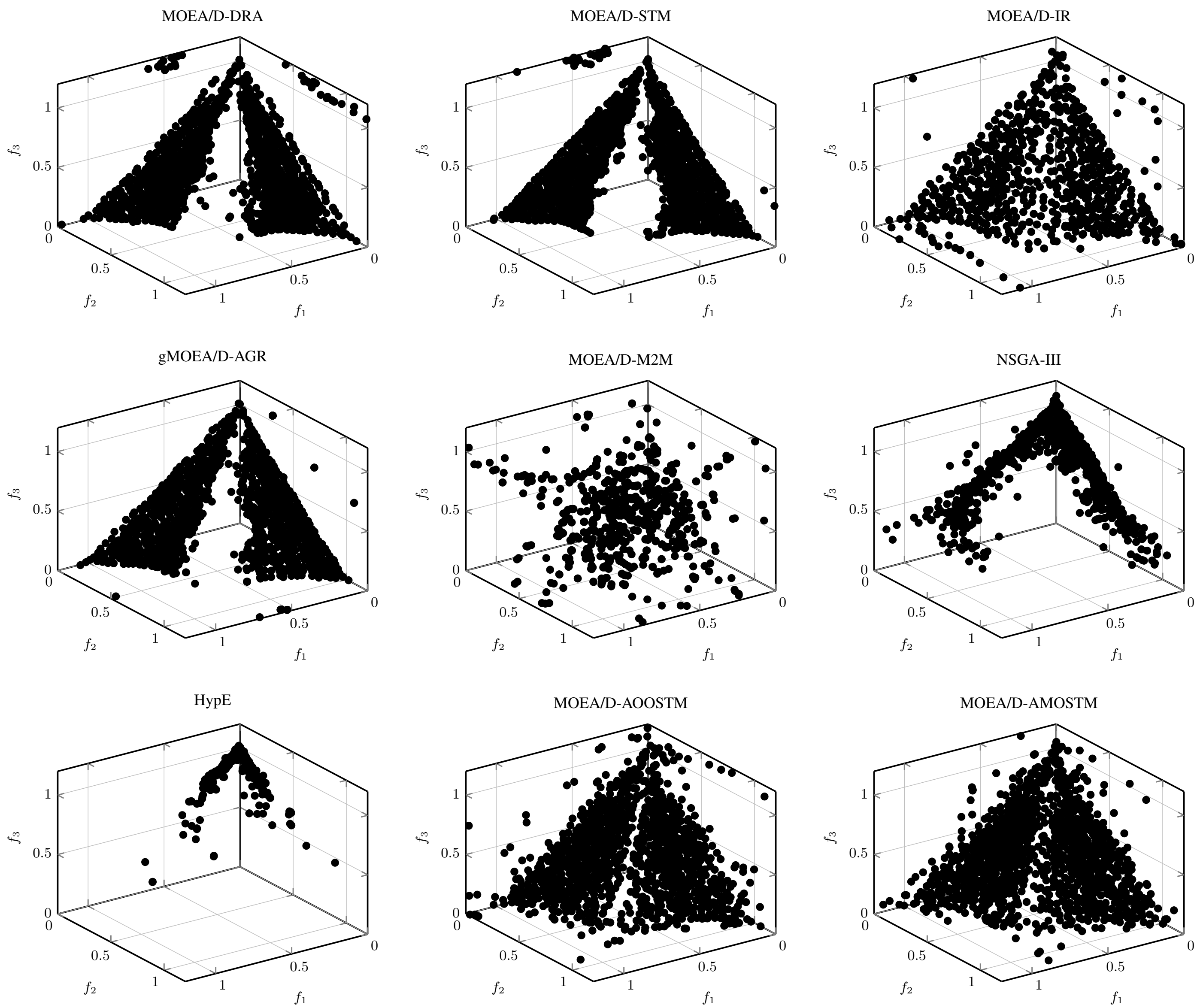

Fig. 16: Final solution sets with best IGD metric values found by 9 MOEAs on UF9.

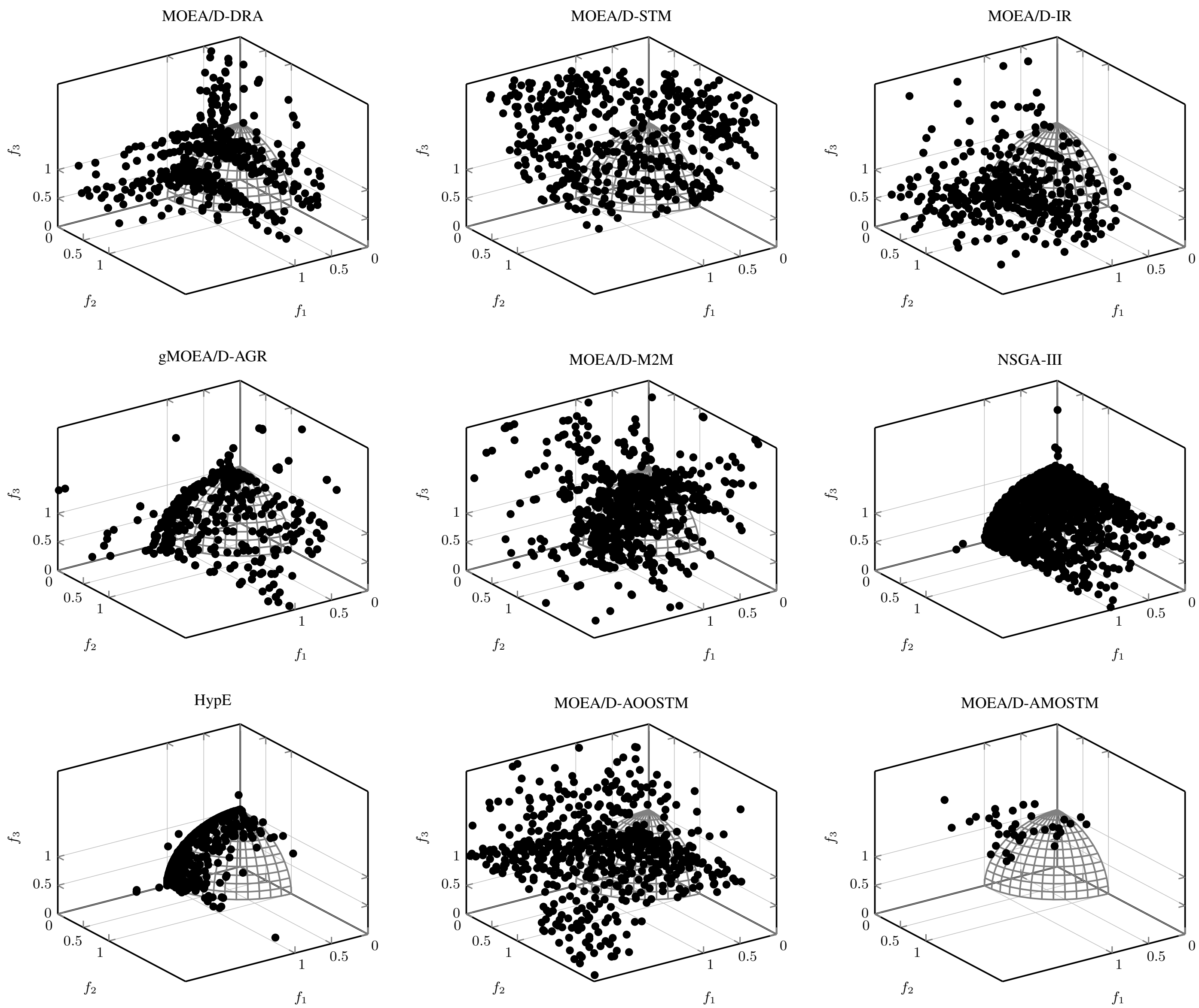

Fig. 17: Final solution sets with best IGD metric values found by 9 MOEAs on UF10.

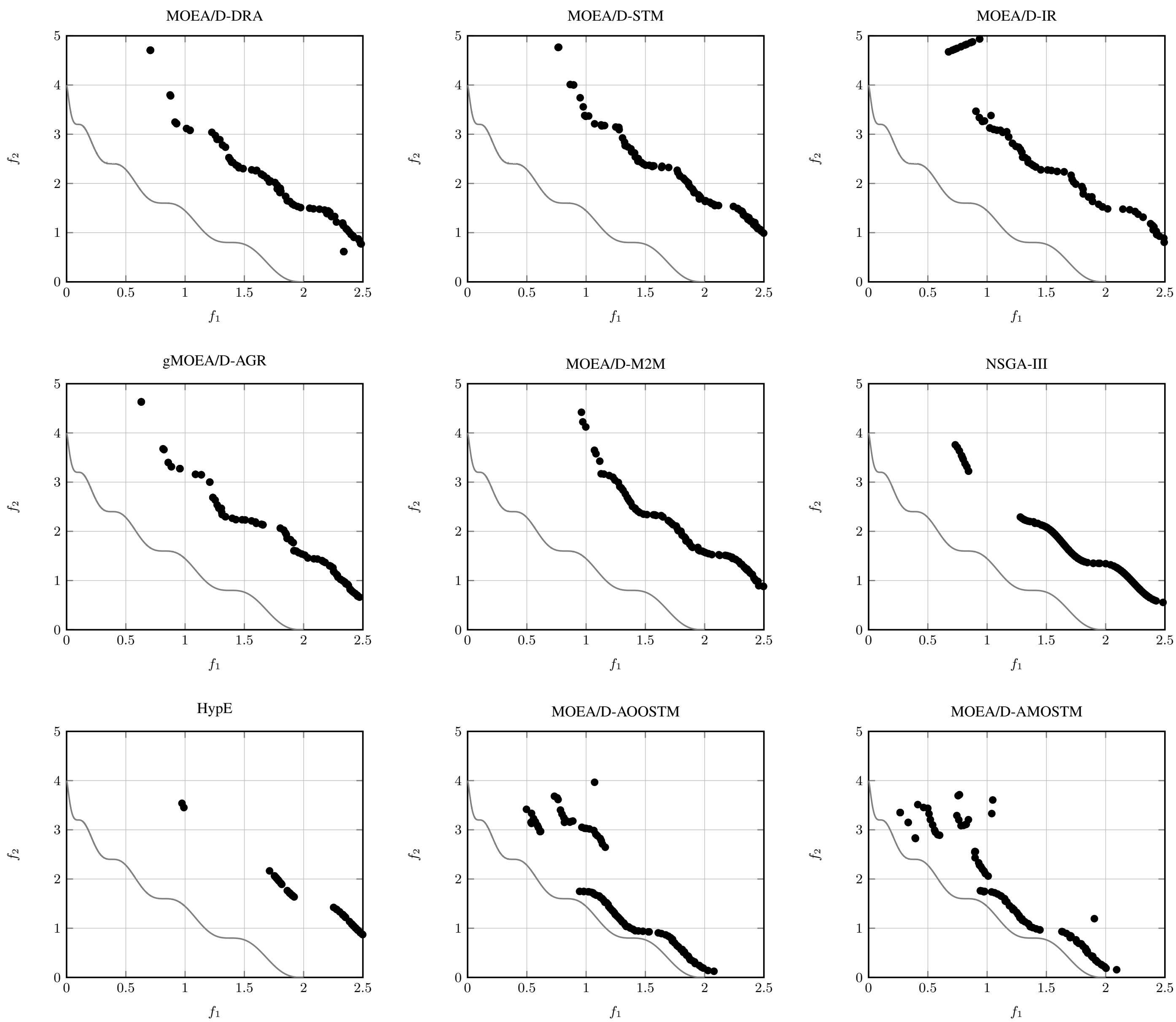

Fig. 18: Final solution sets with best IGD metric values found by 9 MOEAs on WFG1.

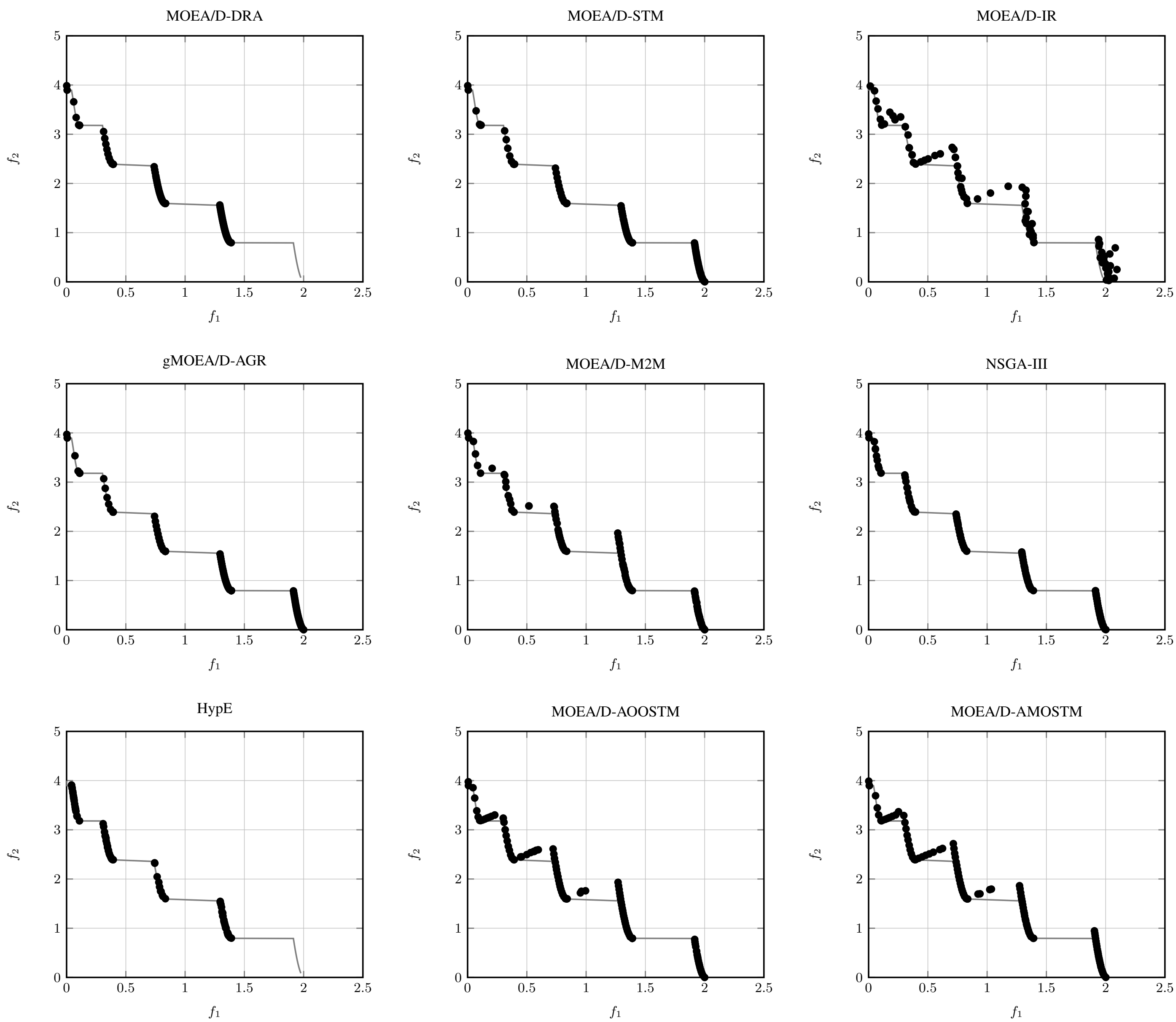

Fig. 19: Final solution sets with best IGD metric values found by 9 MOEAs on WFG2.

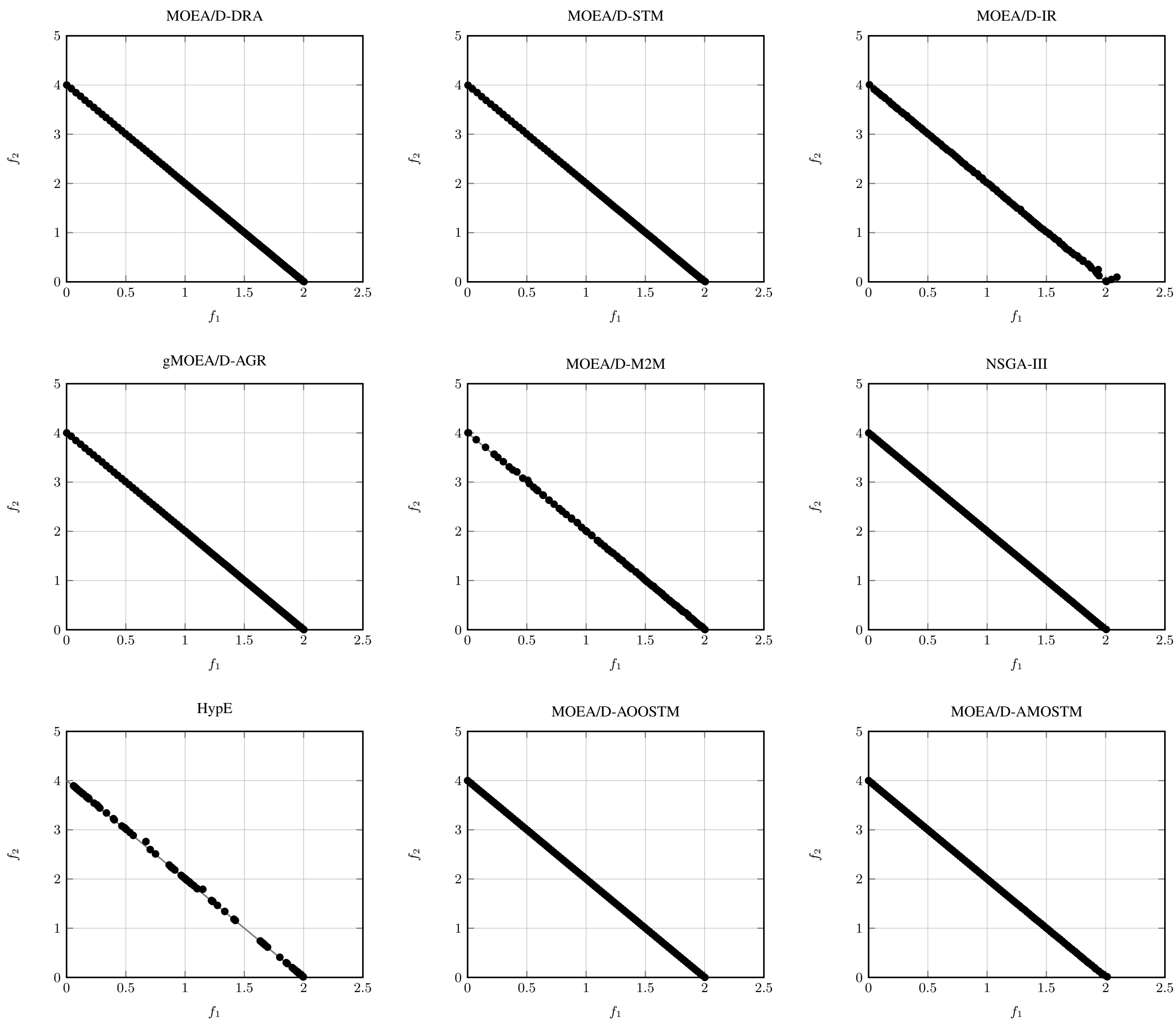

Fig. 20: Final solution sets with best IGD metric values found by 9 MOEAs on WFG3.

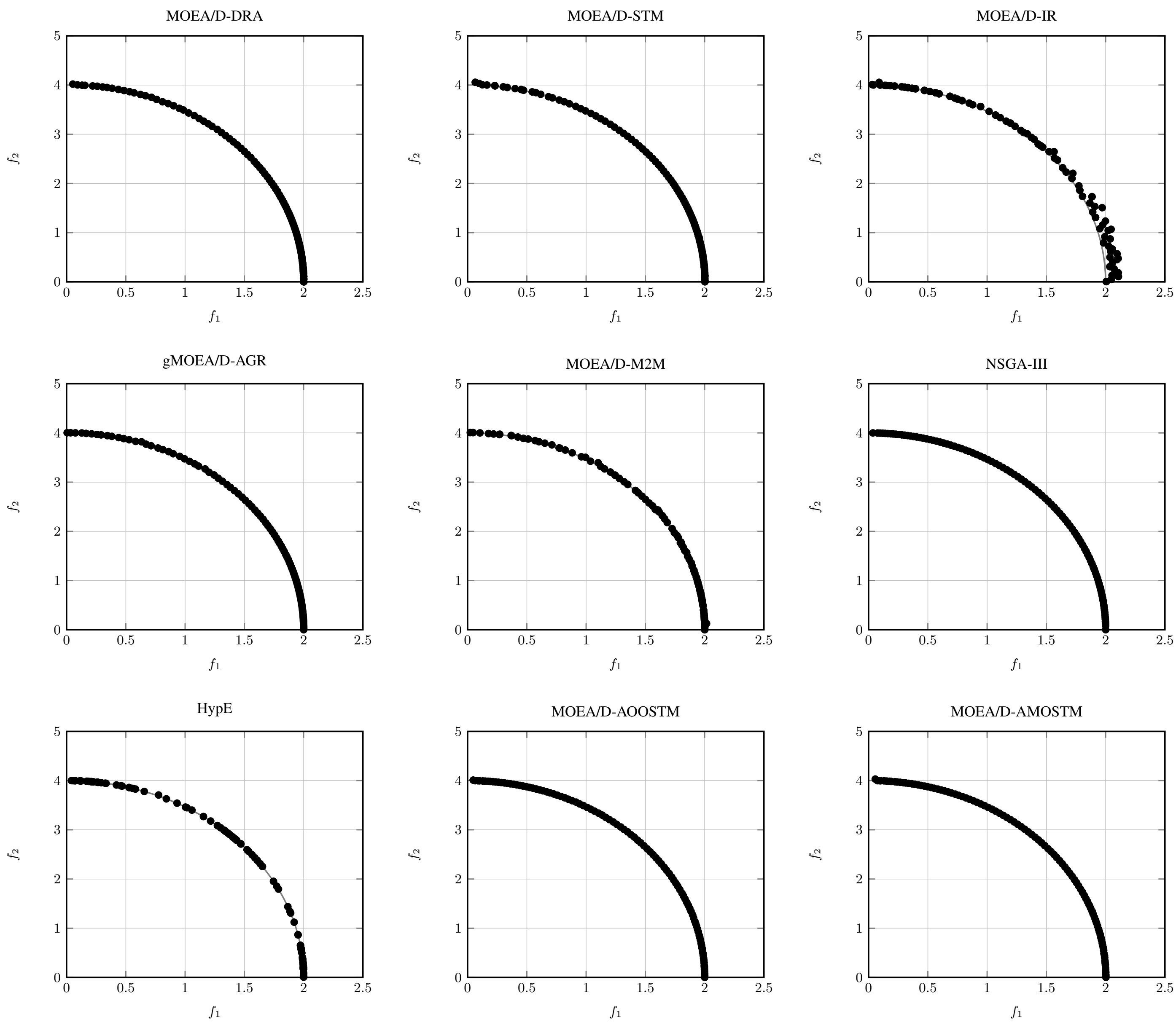

Fig. 21: Final solution sets with best IGD metric values found by 9 MOEAs on WFG4.

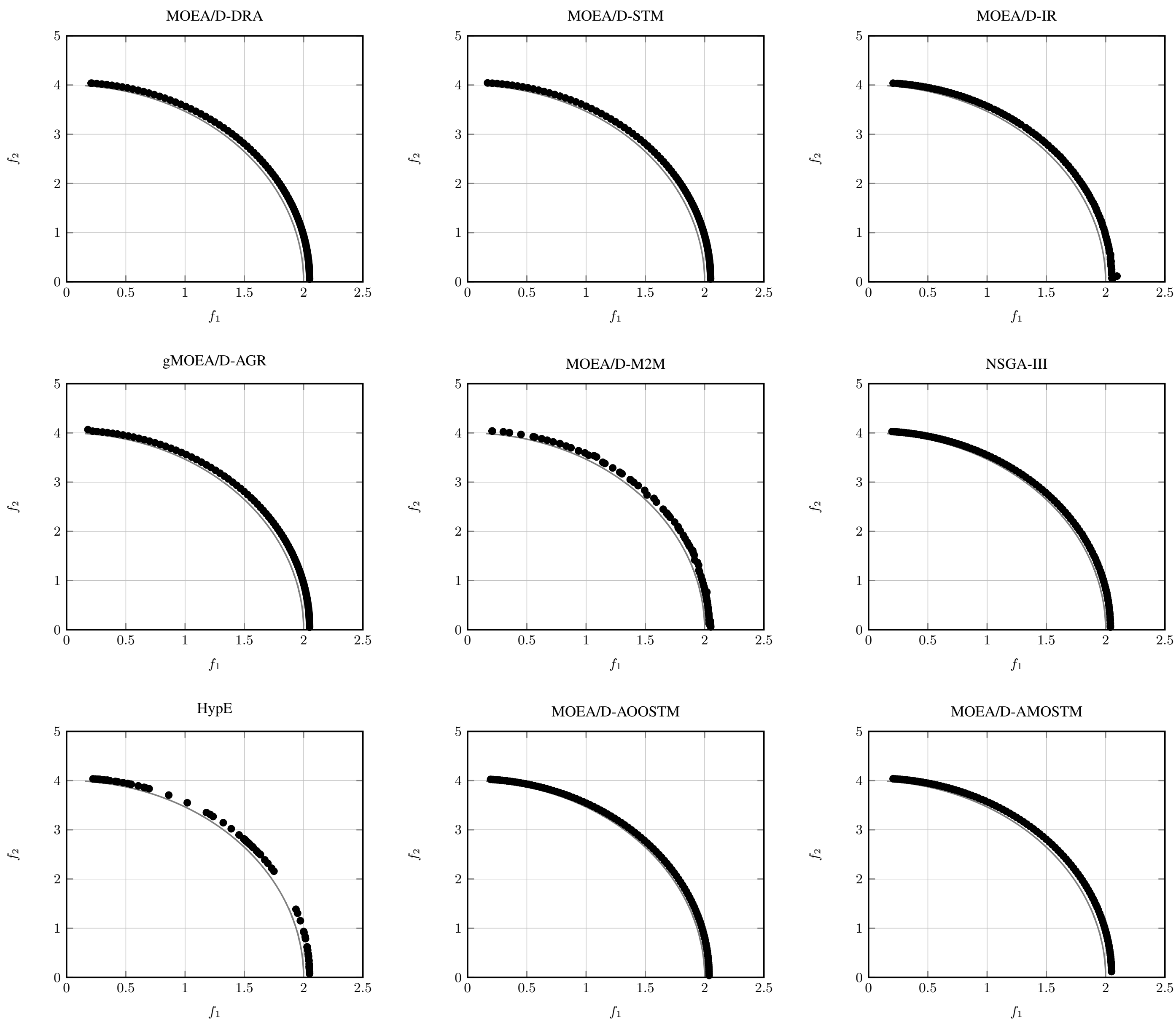

Fig. 22: Final solution sets with best IGD metric values found by 9 MOEAs on WFG5.

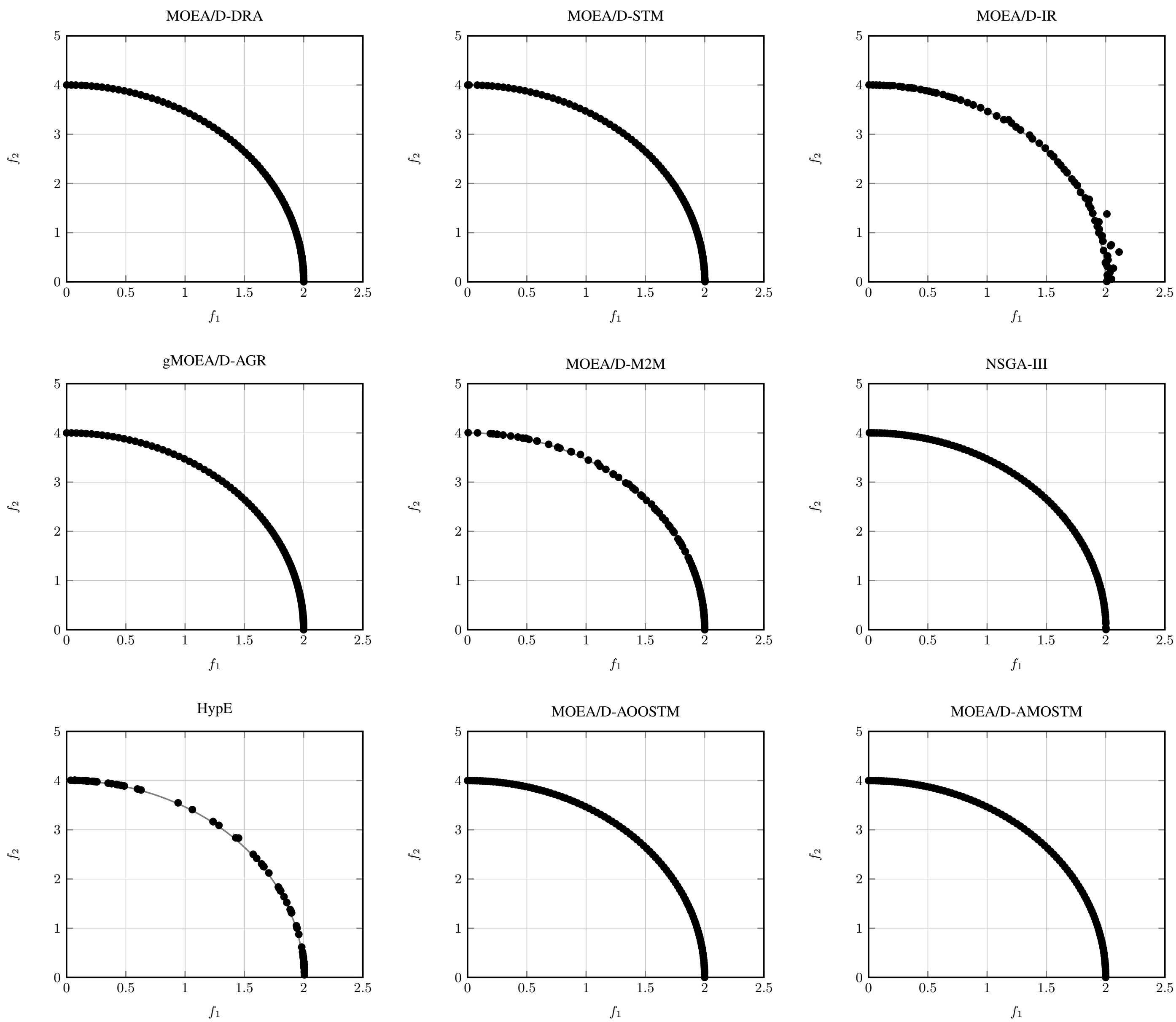

Fig. 23: Final solution sets with best IGD metric values found by 9 MOEAs on WFG6.

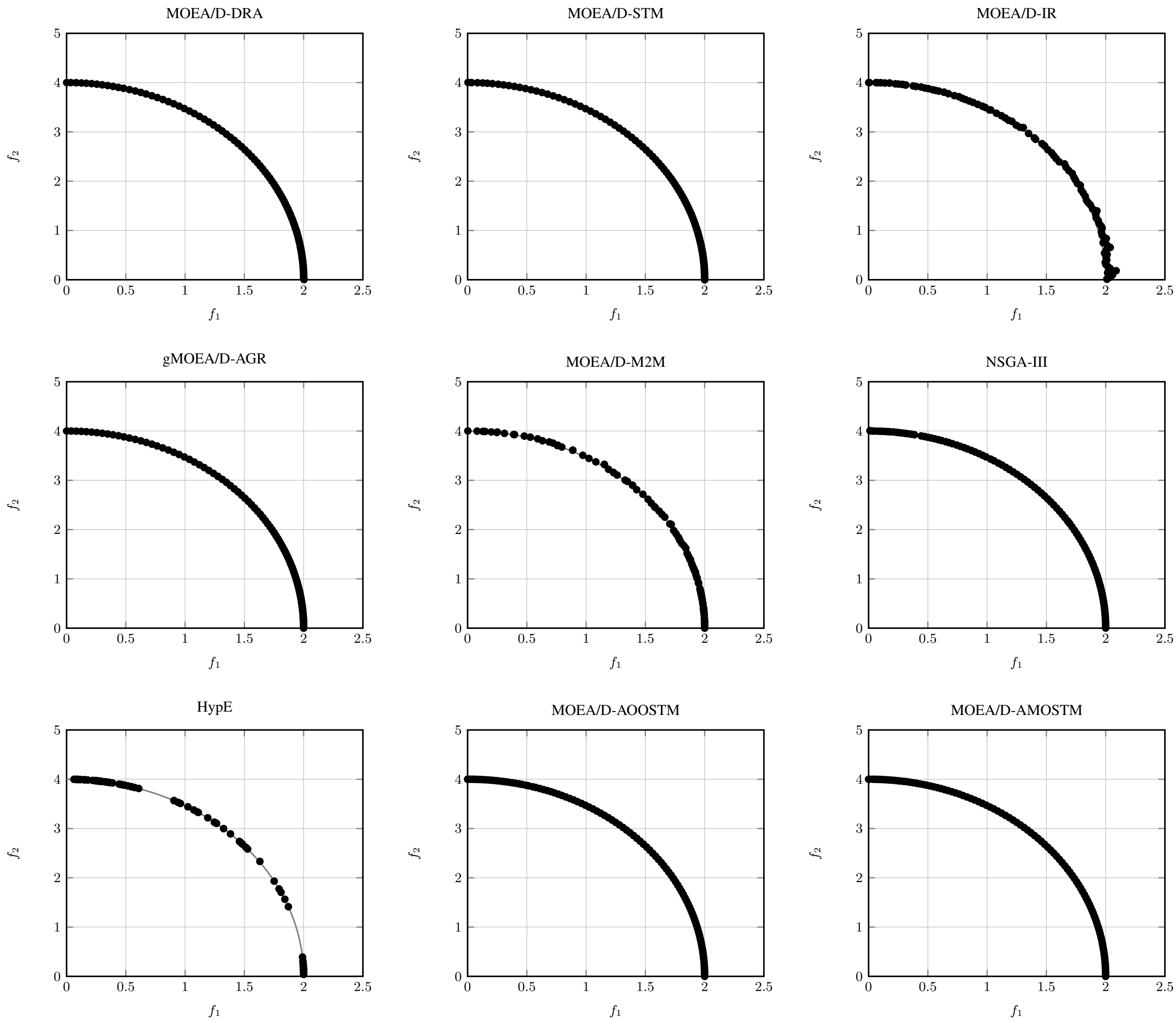

Fig. 24: Final solution sets with best IGD metric values found by 9 MOEAs on WFG7.

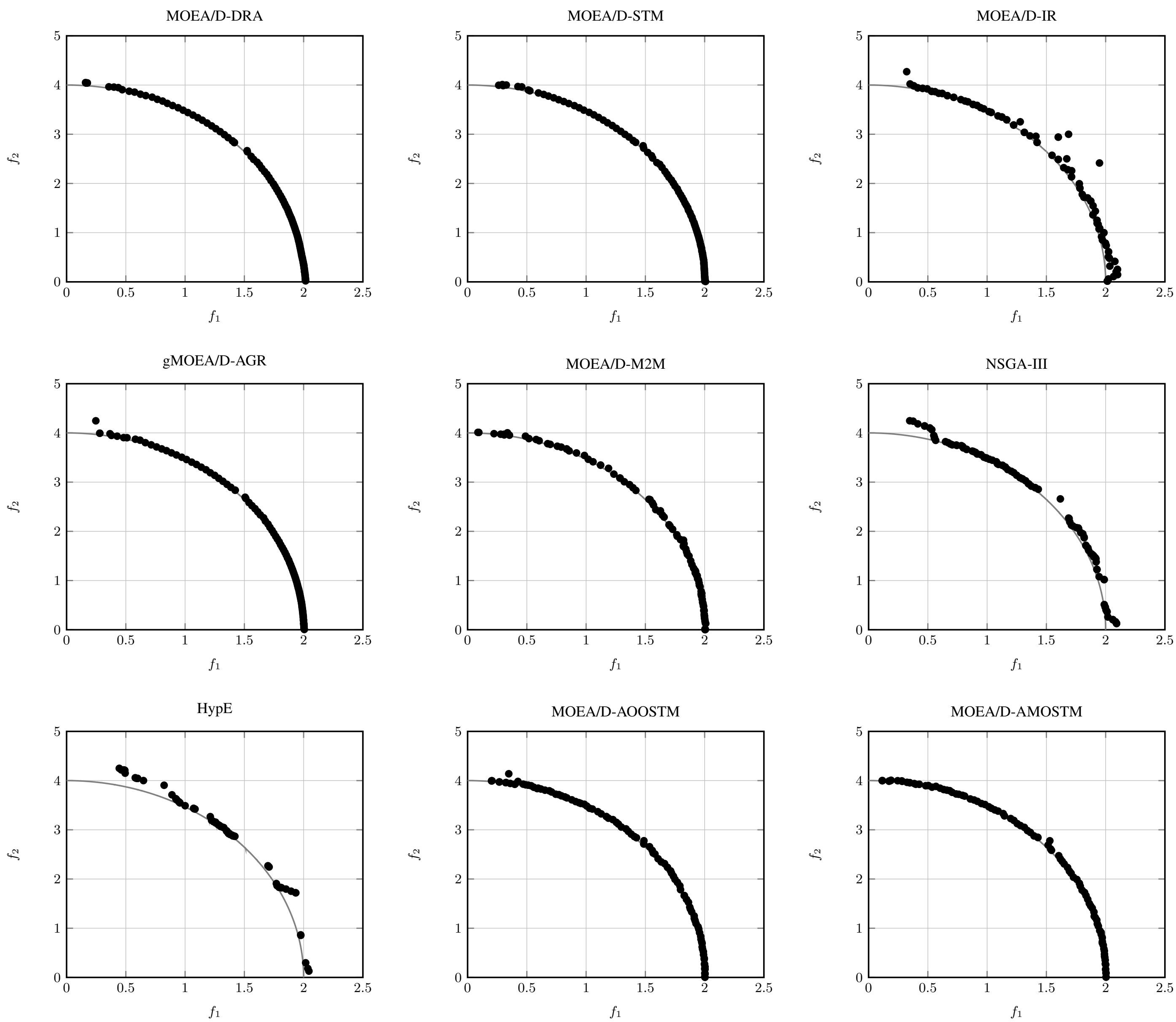

Fig. 25: Final solution sets with best IGD metric values found by 9 MOEAs on WFG8.

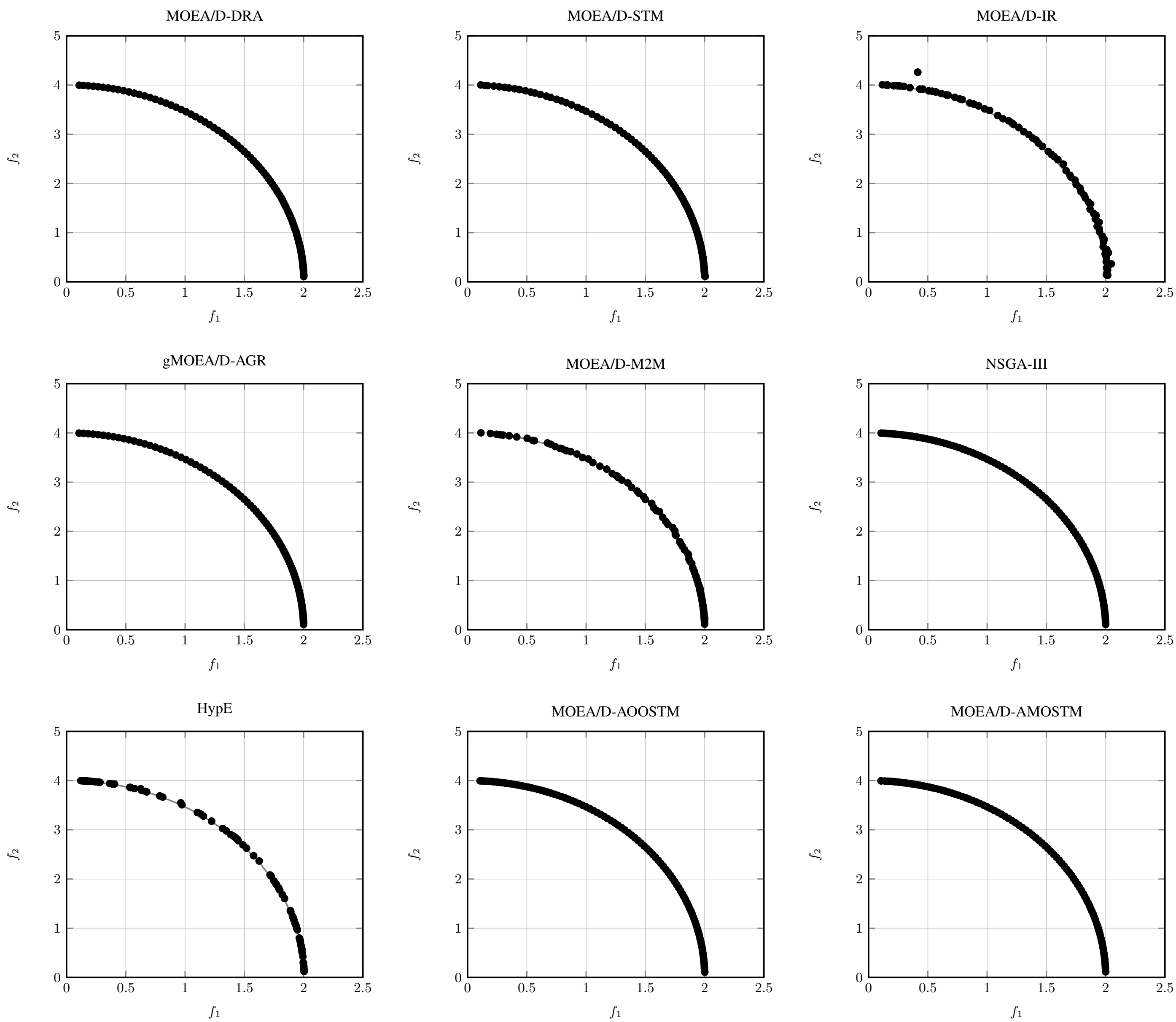

Fig. 26: Final solution sets with best IGD metric values found by 9 MOEAs on WFG9.

TABLE I: IGD Results on Bi-Objective WFG Test Instances.

Problem	IGD	DRA	STM	IR	AGR	M2M	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	3.962E-1	4.083E-1	4.153E-1	3.872E-1	4.289E-1	4.773E-1	5.658E-1	2.118E-1	**1.892E-1**
	Std	1.435E-2	7.929E-3	3.090E-2	1.522E-2	8.467E-3	7.015E-2	6.697E-2	4.290E-2	**3.664E-2**
	Rank	$4 - \downarrow$	$5 - \downarrow$	$6 - \downarrow$	$3 - \downarrow$	$7 - \downarrow$	$8 - \downarrow$	$9 - \downarrow$	$2 \downarrow$	1
WFG2	Mean	1.018E-2	1.015E-2	1.339E-2	1.019E-2	8.300E-3	**5.374E-3**	1.582E-2	6.295E-3	6.408E-3
	Std	4.110E-4	2.421E-4	1.317E-3	2.826E-4	6.039E-4	**3.877E-4**	2.010E-2	3.531E-4	3.815E-4
	Rank	$6 - \downarrow$	$5 - \downarrow$	$8 - \downarrow$	$7 - \downarrow$	$4 - \downarrow$	$1 + \uparrow$	$9 - \downarrow$	$2 \parallel$	3
WFG3	Mean	4.166E-2	4.166E-2	4.370E-2	4.167E-2	4.281E-2	4.178E-2	5.247E-2	4.166E-2	**4.164E-2**
	Std	2.207E-5	3.211E-5	8.565E-4	1.692E-5	1.421E-4	1.017E-4	4.759E-3	9.011E-6	**7.755E-5**
	Rank	$4 \approx \downarrow$	$3 \approx \downarrow$	$8 - \downarrow$	$5 - \downarrow$	$7 - \downarrow$	$6 - \downarrow$	$9 - \downarrow$	$2 \downarrow$	1
WFG4	Mean	7.175E-3	1.015E-2	1.236E-2	6.971E-3	8.917E-3	**4.545E-3**	1.932E-2	4.680E-3	4.634E-3
	Std	1.265E-3	1.187E-3	2.089E-3	5.640E-4	6.980E-4	**3.878E-5**	7.527E-3	1.531E-4	1.061E-4
	Rank	$5 - \downarrow$	$7 - \downarrow$	$8 - \downarrow$	$4 - \downarrow$	$6 - \downarrow$	$1 + \uparrow$	$9 - \downarrow$	$3 \parallel$	2
WFG5	Mean	2.769E-2	2.771E-2	2.921E-2	2.768E-2	2.927E-2	2.726E-2	6.672E-2	**2.722E-2**	2.734E-2
	Std	5.091E-5	7.895E-5	7.705E-4	2.657E-5	3.320E-4	7.174E-4	1.510E-2	**9.134E-4**	9.092E-6
	Rank	$5 - \downarrow$	$6 - \downarrow$	$7 - \downarrow$	$4 - \downarrow$	$8 - \downarrow$	$2 - \downarrow$	$9 - \downarrow$	$1 \downarrow$	3
WFG6	Mean	4.951E-3	4.948E-3	1.265E-2	4.956E-3	7.681E-3	1.285E-2	3.537E-2	4.522E-3	**4.514E-3**
	Std	5.318E-5	3.016E-5	1.647E-3	2.757E-5	3.140E-4	9.979E-3	1.795E-2	2.446E-5	**2.253E-5**
	Rank	$4 - \downarrow$	$3 - \downarrow$	$7 - \downarrow$	$5 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$9 - \downarrow$	$2 \downarrow$	1
WFG7	Mean	5.235E-3	5.235E-3	1.187E-2	5.266E-3	9.086E-3	4.771E-3	3.129E-2	**4.760E-3**	4.768E-3
	Std	2.716E-5	2.453E-5	1.354E-3	2.396E-5	6.262E-4	1.353E-5	1.057E-2	**9.460E-6**	1.186E-5
	Rank	$4 - \downarrow$	$5 - \downarrow$	$8 - \downarrow$	$6 - \downarrow$	$7 - \downarrow$	$3 - \parallel$	$9 - \downarrow$	$1 \uparrow$	2
WFG8	Mean	4.902E-2	2.941E-2	4.266E-2	3.514E-2	**1.225E-2**	6.713E-2	8.833E-2	3.582E-2	3.954E-2
	Std	2.465E-2	2.602E-2	2.522E-2	2.697E-2	**9.297E-4**	1.588E-2	1.946E-2	2.702E-2	2.665E-2
	Rank	$7 - \downarrow$	$2 + \uparrow$	$6 - \downarrow$	$3 \approx \parallel$	$1 \approx \parallel$	$8 - \downarrow$	$9 - \downarrow$	$4 \parallel$	5
WFG9	Mean	5.961E-3	6.635E-3	1.367E-2	5.659E-3	8.546E-3	5.323E-3	1.819E-2	4.964E-3	**4.957E-3**
	Std	4.685E-4	4.045E-4	9.049E-4	3.487E-4	4.817E-4	4.273E-4	4.329E-3	1.996E-4	**1.852E-4**
	Rank	$5 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$4 - \downarrow$	$7 - \downarrow$	$3 - \downarrow$	$9 - \downarrow$	$2 \parallel$	1
	Total Rank	44	42	66	41	53	40	81	19	19
	Final Rank	6	5	8	4	7	3	9	1	1

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while $\uparrow$, $\downarrow$ and $\parallel$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE II: HV Results on Bi-Objective WFG Test Instances.

Problem	HV	DRA	STM	IR	AGR	M2M	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	0.446	0.433	0.424	0.459	0.425	0.380	0.298	0.734	**0.766**
	Std	1.594E-2	9.531E-3	3.464E-2	1.770E-2	8.541E-3	6.637E-2	5.720E-2	5.757E-2	**5.233E-2**
	Rank	$4 - \downarrow$	$5 - \downarrow$	$7 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$9 - \downarrow$	$2 \downarrow$	1
WFG2	Mean	1.004	1.005	0.991	**1.006**	1.004	0.986	0.966	1.000	1.002
	Std	7.184E-3	6.805E-4	7.788E-3	**6.580E-4**	6.663E-4	2.132E-2	2.362E-2	1.468E-2	1.266E-2
	Rank	$4 - \downarrow$	$2 - \downarrow$	$7 - \downarrow$	$1 - \downarrow$	$3 - \downarrow$	$8 - \downarrow$	$9 - \downarrow$	$6 \parallel$	5
WFG3	Mean	0.814	0.814	0.810	0.814	0.812	0.815	0.799	**0.815**	0.815
	Std	1.715E-4	1.890E-4	1.189E-3	1.752E-4	4.258E-4	3.572E-4	4.862E-3	**1.477E-4**	3.064E-4
	Rank	$4 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$5 - \downarrow$	$7 - \downarrow$	$3 - \downarrow$	$9 - \downarrow$	$1 \uparrow$	2
WFG4	Mean	0.658	0.648	0.641	0.660	0.658	**0.666**	0.646	0.664	0.664
	Std	2.579E-3	3.381E-3	7.654E-3	2.578E-3	1.472E-3	**7.581E-4**	4.952E-3	1.393E-3	1.366E-3
	Rank	$6 - \downarrow$	$7 - \downarrow$	$9 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$1 + \uparrow$	$8 - \downarrow$	$2 \uparrow$	3
WFG5	Mean	0.625	0.626	0.622	0.626	0.623	**0.629**	0.577	0.627	0.626
	Std	2.772E-4	1.176E-3	1.910E-3	1.813E-3	1.422E-3	**3.232E-3**	1.405E-2	2.049E-3	1.850E-3
	Rank	$6 - \downarrow$	$5 - \downarrow$	$8 - \downarrow$	$4 - \downarrow$	$7 - \downarrow$	$1 + \uparrow$	$9 - \downarrow$	$2 \uparrow$	3
WFG6	Mean	0.648	0.648	0.636	0.649	0.644	0.630	0.604	**0.649**	0.649
	Std	3.693E-4	3.029E-4	2.621E-3	2.232E-4	7.462E-4	1.853E-2	1.929E-2	**3.172E-4**	4.704E-4
	Rank	$4 - \downarrow$	$5 - \downarrow$	$7 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$9 - \downarrow$	$1 \uparrow$	2
WFG7	Mean	0.649	0.648	0.636	0.649	0.643	0.649	0.615	**0.650**	0.649
	Std	1.942E-4	3.140E-4	2.022E-3	1.307E-4	6.411E-4	4.571E-4	8.051E-3	**3.603E-4**	6.449E-4
	Rank	$5 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$4 - \downarrow$	$7 - \downarrow$	$3 - \downarrow$	$9 - \downarrow$	$1 \uparrow$	2
WFG8	Mean	0.543	0.577	0.543	0.566	**0.626**	0.484	0.439	0.564	0.557
	Std	5.132E-2	5.668E-2	5.553E-2	5.753E-2	**4.234E-3**	3.627E-2	4.001E-2	5.892E-2	5.580E-2
	Rank	$6 \approx \parallel$	$2 \approx \parallel$	$7 - \downarrow$	$3 \approx \parallel$	$1 + \uparrow$	$8 - \downarrow$	$9 - \downarrow$	$4 \parallel$	5
WFG9	Mean	0.693	0.691	0.679	0.695	0.690	0.694	0.676	**0.696**	0.695
	Std	1.498E-3	1.316E-3	1.472E-3	1.344E-3	1.180E-3	1.602E-3	4.366E-3	**1.269E-3**	1.205E-3
	Rank	$5 - \downarrow$	$6 - \downarrow$	$8 - \downarrow$	$3 - \downarrow$	$7 - \downarrow$	$4 - \downarrow$	$9 - \downarrow$	$1 \parallel$	2
	Total Rank	44	44	69	30	49	44	80	20	25
	Final Rank	5	5	8	3	7	4	9	1	2

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while $\uparrow$, $\downarrow$ and $\parallel$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE III: HV Results on 3-Objective WFG Test Instances.

Problem	HV	DRA	MOEA/DD	PICEA-g	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	0.845	0.808	0.773	0.535	0.325	1.157	**1.170**
	Std	6.300E-2	7.173E-2	1.139E-2	2.820E-2	2.464E-2	4.869E-2	**6.357E-2**
	Rank	3 − ↓	4 − ↓	5 − ↓	6 − ↓	7 − ↓	2 ‖	1
WFG2	Mean	1.303	1.505	1.513	1.513	1.426	1.517	**1.532**
	Std	7.798E-2	1.145E-1	1.103E-1	1.120E-1	8.529E-2	1.064E-1	**1.041E-1**
	Rank	7 − ↓	5 − ↓	4 ≈ ‖	3 ≈ ↓	6 − ↓	2 ‖	1
WFG3	Mean	1.042	1.087	**1.180**	1.147	1.035	1.151	1.155
	Std	4.702E-2	1.179E-2	**1.168E-2**	1.055E-2	2.777E-2	8.816E-3	1.387E-2
	Rank	6 − ↓	5 − ↓	1 + ↑	4 ≈ ↓	7 − ↓	3 ‖	2
WFG4	Mean	1.050	1.099	1.101	1.103	1.026	**1.120**	1.119
	Std	9.972E-3	3.250E-3	5.749E-3	4.315E-3	1.997E-2	**4.623E-3**	4.693E-3
	Rank	6 − ↓	5 − ↓	4 − ↓	3 − ↓	7 − ↓	1 ‖	2
WFG5	Mean	1.019	1.039	1.030	1.048	0.965	**1.066**	1.060
	Std	5.412E-3	4.574E-3	9.741E-3	6.817E-3	2.027E-2	**4.680E-3**	5.117E-3
	Rank	6 − ↓	4 − ↓	5 − ↓	3 − ↓	7 − ↓	1 ↑	2
WFG6	Mean	1.005	1.053	**1.061**	1.058	0.960	1.055	1.058
	Std	2.312E-2	1.051E-2	**8.650E-3**	1.045E-2	2.577E-2	1.330E-2	1.123E-2
	Rank	6 − ↓	5 ≈ ↓	1 + ‖	3 ≈ ‖	7 − ↓	4 ‖	2
WFG7	Mean	1.012	1.108	1.116	1.117	0.958	**1.117**	1.117
	Std	4.076E-2	1.859E-2	3.002E-3	3.810E-3	4.883E-2	**4.343E-3**	5.100E-3
	Rank	6 − ↓	5 − ↓	4 − ‖	3 ≈ ‖	7 − ↓	1 ‖	2
WFG8	Mean	0.949	0.985	0.971	0.987	0.834	1.008	**1.011**
	Std	1.268E-2	6.719E-3	9.310E-3	7.673E-3	2.112E-2	3.573E-3	**3.667E-3**
	Rank	6 − ↓	4 − ↓	5 − ↓	3 − ↓	7 − ↓	2 ↓	1
WFG9	Mean	0.917	1.003	0.978	1.001	0.803	**1.003**	0.984
	Std	3.594E-2	3.437E-2	2.380E-2	3.237E-2	8.772E-2	**4.336E-2**	3.376E-2
	Rank	6 − ↓	2 ≈ ‖	5 ≈ ↑	3 ≈ ↑	7 − ↓	1 ‖	4
	Total Rank	52	39	34	31	62	17	17
	Final Rank	6	5	4	3	7	1	1

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while ↑, ↓ and ‖ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE IV: HV Results on 5-Objective WFG Test Instances.

Problem	HV	DRA	MOEA/DD	PICEA-g	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	1.966	1.536	1.568	0.811	0.647	**2.274**	2.200
	Std	1.069E-1	7.823E-2	6.283E-2	3.766E-2	1.933E-2	**2.652E-2**	4.473E-2
	Rank	3 − ↓	5 − ↓	4 − ↓	6 − ↓	7 − ↓	1 ↑	2
WFG2	Mean	2.002	2.369	2.422	2.419	2.262	2.430	**2.440**
	Std	1.119E-1	1.017E-1	1.415E-1	1.081E-1	1.525E-1	1.330E-1	**1.226E-1**
	Rank	7 − ↓	5 − ↓	3 ≈ ↓	4 − ↓	6 − ↓	2 ↓	1
WFG3	Mean	1.485	1.556	**1.777**	1.668	1.545	1.729	1.734
	Std	7.256E-2	3.026E-2	**1.338E-2**	1.641E-2	2.751E-2	1.535E-2	1.325E-2
	Rank	7 − ↓	5 − ↓	1 + ↑	4 − ↓	6 − ↓	3 ‖	2
WFG4	Mean	1.912	2.098	2.095	2.089	1.595	**2.160**	2.159
	Std	4.472E-2	7.420E-3	9.276E-2	8.472E-3	7.605E-2	**7.260E-3**	9.300E-3
	Rank	6 − ↓	3 − ↓	4 − ↓	5 − ↓	7 − ↓	1 ‖	2
WFG5	Mean	1.904	1.977	1.999	2.020	1.680	**2.061**	2.058
	Std	3.214E-2	7.590E-3	9.957E-3	6.294E-3	8.360E-2	**6.076E-3**	5.817E-3
	Rank	6 − ↓	5 − ↓	4 − ↓	3 − ↓	7 − ↓	1 ↑	2
WFG6	Mean	1.649	2.014	2.052	2.035	1.716	2.053	**2.054**
	Std	1.636E-1	1.459E-2	1.217E-2	1.316E-2	6.289E-2	2.992E-2	**2.779E-2**
	Rank	7 − ↓	5 − ↓	3 − ↓	4 − ↓	6 − ↓	2 ‖	1
WFG7	Mean	1.834	2.120	2.143	2.140	1.559	**2.171**	2.171
	Std	2.918E-2	5.074E-3	4.537E-3	5.125E-3	1.123E-1	**5.127E-3**	4.546E-3
	Rank	6 − ↓	5 − ↓	3 − ↓	4 − ↓	7 − ↓	1 ‖	2
WFG8	Mean	1.607	1.901	1.892	1.897	1.401	**1.915**	1.910
	Std	2.804E-1	2.260E-2	1.312E-2	6.639E-3	7.253E-2	**2.013E-2**	1.546E-2
	Rank	6 − ↓	3 − ↓	5 − ↓	4 − ↓	7 − ↓	1 ↑	2
WFG9	Mean	1.738	1.851	1.828	1.839	1.293	**1.912**	1.828
	Std	1.316E-1	4.496E-2	6.152E-3	3.898E-2	1.776E-1	**8.252E-2**	3.414E-2
	Rank	6 − ↓	2 − ↑	5 − ↑	3 − ↑	7 − ↓	1 ↑	4
	Total Rank	54	38	32	37	60	13	18
	Final Rank	6	5	3	4	7	1	2

According to Wilcoxon's rank sum test, +, − and ≈ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while ↑, ↓ and ‖ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE V: HV Results on 8-Objective WFG Test Instances.

Problem	HV	DRA	MOEA/DD	PICEA-g	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	3.375	3.086	3.308	1.190	1.077	**3.965**	3.920
	Std	2.653E-1	1.978E-1	1.494E-1	6.373E-2	2.814E-2	**8.430E-2**	8.167E-2
	Rank	$3 - \downarrow$	$5 - \downarrow$	$4 - \downarrow$	$6 - \downarrow$	$7 - \downarrow$	$1 \uparrow$	2
WFG2	Mean	3.411	4.039	**4.096**	4.064	3.804	3.869	3.999
	Std	1.860E-1	1.249E-1	**3.059E-1**	3.135E-1	2.809E-1	3.498E-1	3.514E-1
	Rank	$7 - \downarrow$	$3 \approx \downarrow$	$1 + \uparrow$	$2 \approx \downarrow$	$6 \approx \downarrow$	$5 \downarrow$	4
WFG3	Mean	1.743	2.416	**2.948**	2.485	2.626	1.516	1.780
	Std	9.377E-2	5.578E-2	**7.445E-2**	1.686E-1	6.623E-2	9.549E-2	9.581E-2
	Rank	$6 + \downarrow$	$4 + \uparrow$	$1 + \uparrow$	$3 + \uparrow$	$2 + \uparrow$	$7 \downarrow$	5
WFG4	Mean	2.360	3.864	2.934	3.977	2.358	**4.084**	4.044
	Std	1.828E-1	3.513E-2	1.565E-1	1.723E-2	2.108E-1	**2.233E-2**	1.950E-2
	Rank	$6 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \downarrow$	$7 - \downarrow$	$1 \uparrow$	2
WFG5	Mean	2.671	3.494	3.228	**3.844**	2.217	3.832	3.765
	Std	1.411E-1	5.859E-2	1.114E-1	**9.012E-3**	2.904E-1	3.490E-2	2.932E-2
	Rank	$6 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$1 + \uparrow$	$7 - \downarrow$	$2 \uparrow$	3
WFG6	Mean	1.688	3.725	3.761	**3.879**	2.752	3.848	3.826
	Std	2.637E-1	5.744E-2	1.503E-1	**3.334E-2**	1.478E-1	4.808E-2	4.707E-2
	Rank	$7 - \downarrow$	$5 - \downarrow$	$4 - \Vert$	$1 + \uparrow$	$6 - \downarrow$	$2 \uparrow$	3
WFG7	Mean	1.996	3.966	3.469	4.067	2.352	**4.075**	4.068
	Std	1.170E-1	1.339E-2	1.330E-1	8.147E-3	1.517E-1	**1.526E-2**	1.410E-2
	Rank	$7 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \Vert$	$6 - \downarrow$	$1 \uparrow$	2
WFG8	Mean	0.683	3.422	3.409	3.579	2.381	**3.647**	3.617
	Std	3.251E-1	1.477E-1	8.796E-2	3.905E-2	2.145E-1	**5.071E-2**	3.528E-2
	Rank	$7 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$1 \uparrow$	2
WFG9	Mean	1.721	3.161	3.206	3.345	1.776	**3.511**	3.411
	Std	5.495E-1	1.022E-1	8.248E-2	8.070E-2	3.686E-1	**9.378E-2**	1.532E-1
	Rank	$7 - \downarrow$	$5 - \downarrow$	$4 - \downarrow$	$3 - \Vert$	$6 - \downarrow$	$1 \uparrow$	2
	Total Rank	56	38	34	25	53	21	25
	Final Rank	7	5	4	2	6	1	2

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while $\uparrow$, $\downarrow$ and $\Vert$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE VI: HV Results on 10-Objective WFG Test Instances.

Problem	HV	DRA	MOEA/DD	PICEA-g	NSGA-III	HypE	AOOSTM	AMOSTM
WFG1	Mean	4.346	5.213	5.630	1.855	1.577	**5.910**	5.903
	Std	4.018E-1	1.774E-1	2.251E-1	9.321E-2	4.056E-2	**4.689E-2**	2.921E-2
	Rank	$5 - \downarrow$	$4 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$7 - \downarrow$	$1 \uparrow$	2
WFG2	Mean	5.189	5.910	6.001	**6.108**	5.751	5.728	5.809
	Std	3.614E-1	2.511E-2	3.823E-1	**1.977E-1**	2.962E-1	5.098E-1	4.948E-1
	Rank	$7 - \downarrow$	$3 \approx \downarrow$	$2 + \uparrow$	$1 \approx \downarrow$	$5 \approx \downarrow$	$6 \Vert$	4
WFG3	Mean	1.614	3.433	**4.355**	4.037	3.932	2.273	3.369
	Std	2.424E-1	6.710E-2	**6.723E-2**	3.693E-1	1.008E-1	1.377E-1	3.381E-1
	Rank	$7 - \downarrow$	$4 + \Vert$	$1 + \uparrow$	$2 + \uparrow$	$3 + \uparrow$	$6 \downarrow$	5
WFG4	Mean	3.133	5.556	4.458	5.926	3.736	**6.102**	6.091
	Std	2.213E-1	6.423E-2	3.231E-1	2.114E-2	5.297E-1	**8.401E-3**	1.079E-2
	Rank	$7 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$1 \uparrow$	2
WFG5	Mean	3.478	5.052	4.740	5.691	3.837	**5.730**	5.693
	Std	2.671E-1	7.583E-2	1.795E-1	8.529E-3	3.430E-1	**1.287E-2**	1.163E-2
	Rank	$7 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \Vert$	$6 - \downarrow$	$1 \uparrow$	2
WFG6	Mean	2.419	5.334	5.385	5.742	4.209	**5.759**	5.745
	Std	3.566E-1	9.123E-2	1.777E-1	4.530E-2	2.374E-1	**8.171E-2**	4.168E-2
	Rank	$7 - \downarrow$	$5 - \downarrow$	$4 - \downarrow$	$3 - \Vert$	$6 - \downarrow$	$1 \uparrow$	2
WFG7	Mean	2.814	5.783	5.202	6.037	4.132	**6.099**	6.097
	Std	7.476E-2	1.415E-2	2.253E-1	6.758E-3	3.111E-1	**4.981E-3**	4.438E-3
	Rank	$7 - \downarrow$	$4 - \downarrow$	$5 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$1 \uparrow$	2
WFG8	Mean	0.922	5.009	5.108	5.481	4.013	**5.671**	5.593
	Std	3.005E-1	2.868E-1	2.155E-1	4.014E-2	3.093E-1	**7.056E-2**	2.317E-2
	Rank	$7 - \downarrow$	$5 - \downarrow$	$4 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$1 \uparrow$	2
WFG9	Mean	2.418	4.552	4.734	4.960	2.855	**5.360**	5.244
	Std	6.264E-1	1.182E-1	8.507E-2	1.174E-1	3.285E-1	**1.137E-1**	1.959E-1
	Rank	$7 - \downarrow$	$5 - \downarrow$	$4 - \downarrow$	$3 - \downarrow$	$6 - \downarrow$	$1 \uparrow$	2
	Total Rank	61	38	33	27	51	19	23
	Final Rank	7	5	4	3	6	1	2

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AOOSTM, while $\uparrow$, $\downarrow$ and $\Vert$ indicate that the corresponding EMO algorithm is significantly better than, worse than or similar to MOEA/D-AMOSTM.

TABLE VII: IGD Results of MOEA/D-AOOSTM-v, MOEA/D-AOOSTM, MOEA/D-AMOSTM-v and MOEA/D-AMOSTM on MOP and UF test instances.

IGD	MOEA/D-AOOSTM-v			MOEA/D-AOOSTM		MOEA/D-AMOSTM-v			MOEA/D-AMOSTM	
Problem	Mean	Std	Test	Mean	Std	Mean	Std	Test	Mean	Std
MOP1	2.409E-2	2.324E-3	$\approx$	**2.407E-2**	**2.907E-3**	**2.307E-2**	**2.003E-3**	$\uparrow$	2.390E-2	2.551E-3
MOP2	4.127E-2	7.034E-2	$-$	**2.034E-2**	**4.301E-2**	8.011E-2	1.003E-1	$\downarrow$	**3.115E-2**	**6.203E-2**
MOP3	4.713E-2	8.372E-2	$\approx$	**4.140E-2**	**7.378E-2**	5.429E-2	8.506E-2	$\|$	**3.203E-2**	**6.527E-2**
MOP4	2.154E-2	1.658E-2	$-$	**2.025E-2**	**3.284E-2**	2.844E-2	3.486E-2	$\downarrow$	**1.414E-2**	**1.155E-2**
MOP5	2.076E-2	2.951E-3	$\approx$	**2.035E-2**	**1.692E-3**	2.284E-2	1.878E-3	$\|$	**2.042E-2**	**1.803E-3**
MOP6	**5.286E-2**	**2.917E-3**	$+$	5.398E-2	3.094E-3	**5.258E-2**	**1.959E-3**	$\uparrow$	5.328E-2	2.917E-3
MOP7	**7.079E-2**	**2.798E-3**	$+$	8.186E-2	2.778E-3	**7.011E-2**	**2.095E-3**	$\|$	7.912E-2	2.619E-3
UF1	9.697E-4	6.294E-5	$\approx$	**9.631E-4**	**4.650E-5**	9.726E-4	6.898E-5	$\|$	**9.696E-4**	**5.158E-5**
UF2	2.805E-3	1.389E-3	$\approx$	**2.270E-3**	**5.587E-4**	2.650E-3	6.354E-4	$\|$	**2.577E-3**	**5.649E-4**
UF3	**5.910E-3**	**5.365E-3**	$-$	7.296E-3	8.380E-3	5.273E-3	6.229E-3	$\|$	**4.110E-3**	**3.128E-3**
UF4	5.306E-2	3.161E-3	$\approx$	**5.269E-2**	**3.523E-3**	5.120E-2	3.522E-3	$\|$	**5.043E-2**	**2.803E-3**
UF5	2.527E-1	2.541E-2	$\approx$	**2.514E-1**	**1.766E-2**	2.436E-1	2.454E-2	$\|$	**2.392E-1**	**2.220E-2**
UF6	8.293E-2	4.294E-2	$\approx$	**8.146E-2**	**4.048E-2**	**5.854E-2**	**2.408E-2**	$\|$	6.876E-2	3.300E-2
UF7	1.221E-3	3.194E-4	$\approx$	**1.150E-3**	**1.095E-4**	1.175E-3	1.244E-4	$\|$	**1.148E-3**	**1.481E-4**
UF8	**2.653E-2**	**1.122E-2**	$\approx$	2.921E-2	5.154E-3	**5.297E-2**	**7.652E-3**	$\|$	5.393E-2	9.528E-3
UF9	**3.275E-2**	**2.917E-3**	$\approx$	3.704E-2	3.125E-2	3.853E-2	4.579E-2	$\|$	**3.769E-2**	**4.306E-2**
UF10	**9.741E-1**	**3.765E-1**	$\approx$	1.028E+0	2.943E-1	**2.329E+0**	**1.980E-1**	$\|$	2.426E+0	1.868E-1

According to Wilcoxon's rank sum test, $+$, $-$ and $\approx$ indicate that MOEA/D-AOOSTM-v is significantly better than, worse than or similar to MOEA/D-AOOSTM. $\uparrow$, $\downarrow$ and $\|$ indicate that MOEA/D-AMOSTM-v is significantly better than, worse than or similar to MOEA/D-AMOSTM.