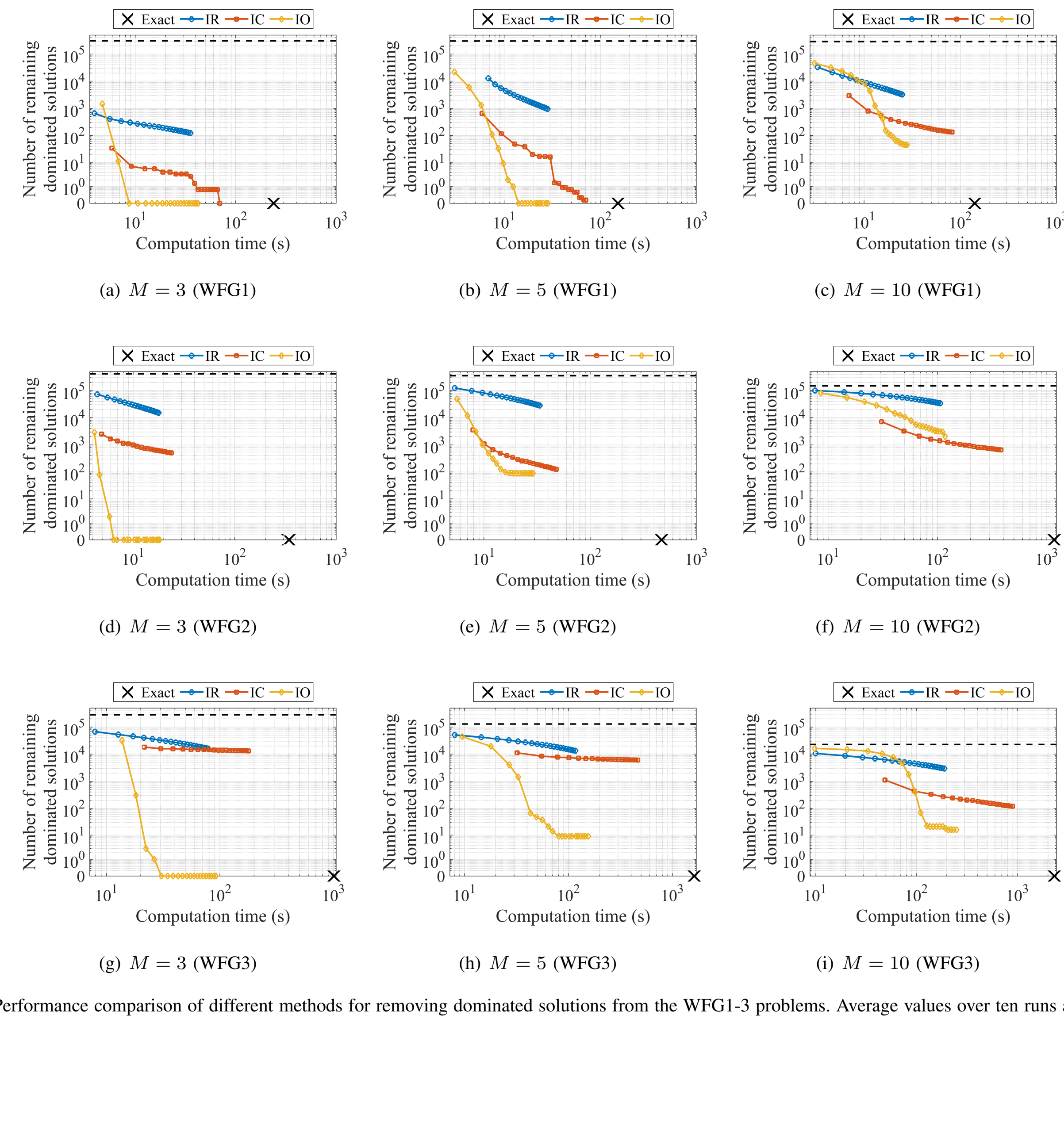
Supplementary Material of “Analysis of Partition Methods for Dominated Solution Removal from Large Solution Sets”

S1. Performance Comparison among Different Methods

Figure 1-3, Figure 4 and Figure 5 show the performance comparison of different methods in terms of the average computation time and the average number of remaining dominated solutions on the WFG, DTLZ and RE problems, respectively.

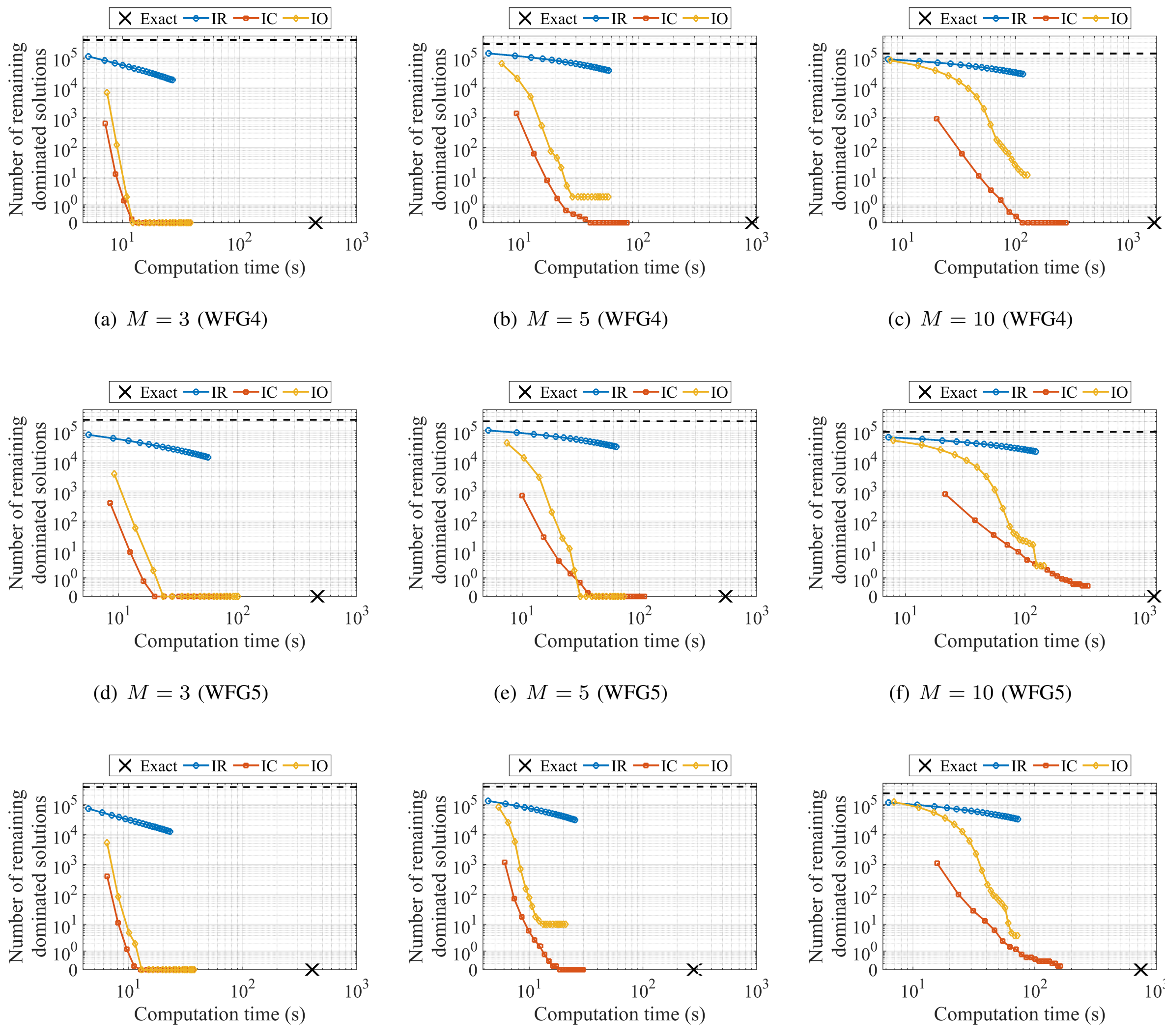


(a) $M = 3$ (WFG1) (b) $M = 5$ (WFG1) (c) $M = 10$ (WFG1)

(d) $M = 3$ (WFG2) (e) $M = 5$ (WFG2) (f) $M = 10$ (WFG2)

(g) $M = 3$ (WFG3) (h) $M = 5$ (WFG3) (i) $M = 10$ (WFG3)

Fig. 1. Performance comparison of different methods for removing dominated solutions from the WFG1-3 problems. Average values over ten runs are shown.

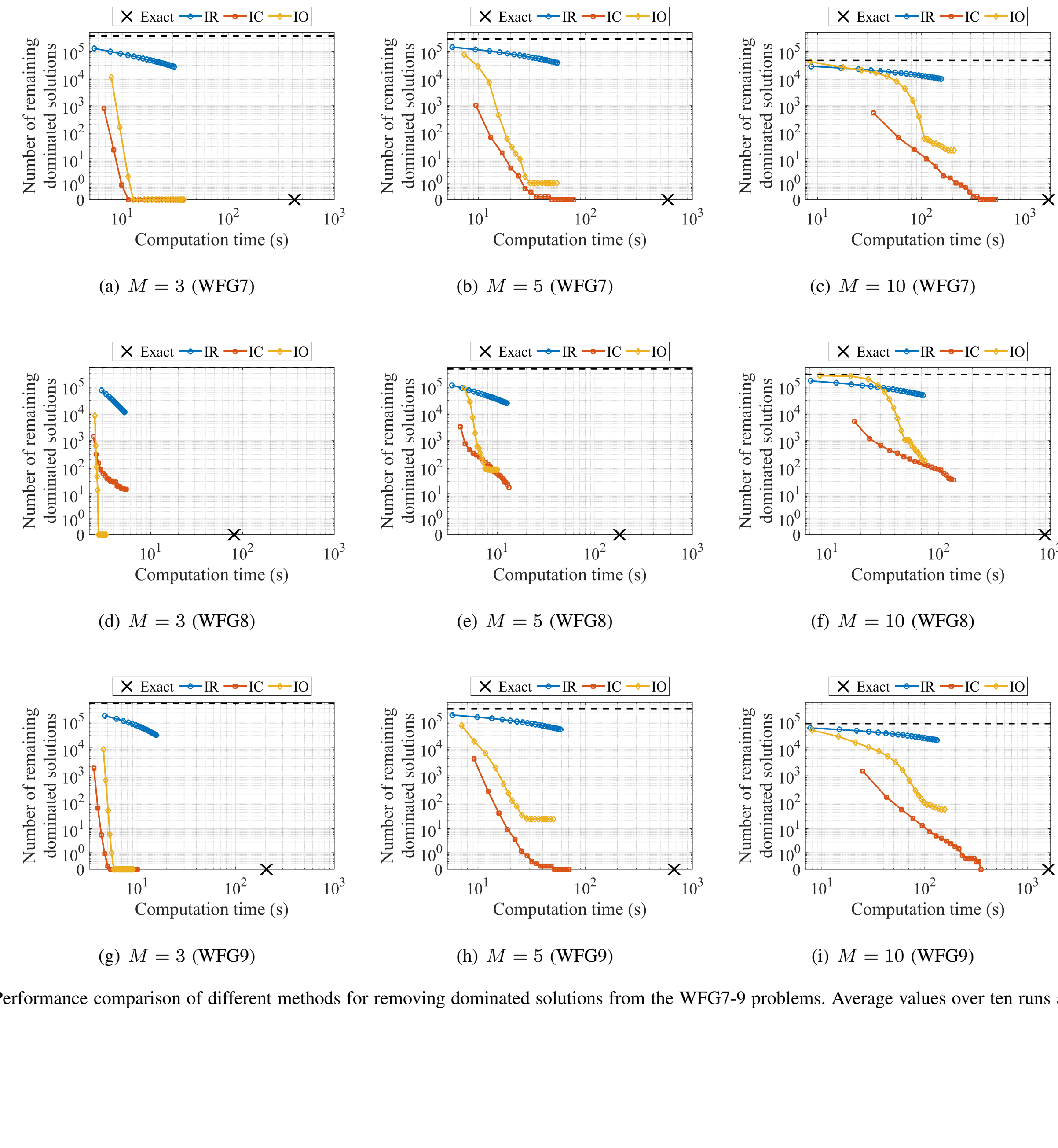


(a) $M = 3$ (WFG4) (b) $M = 5$ (WFG4) (c) $M = 10$ (WFG4)

(d) $M = 3$ (WFG5) (e) $M = 5$ (WFG5) (f) $M = 10$ (WFG5)

(g) $M = 3$ (WFG6) (h) $M = 5$ (WFG6) (i) $M = 10$ (WFG6)

Fig. 2. Performance comparison of different methods for removing dominated solutions from the WFG4-6 problems. Average values over ten runs are shown.

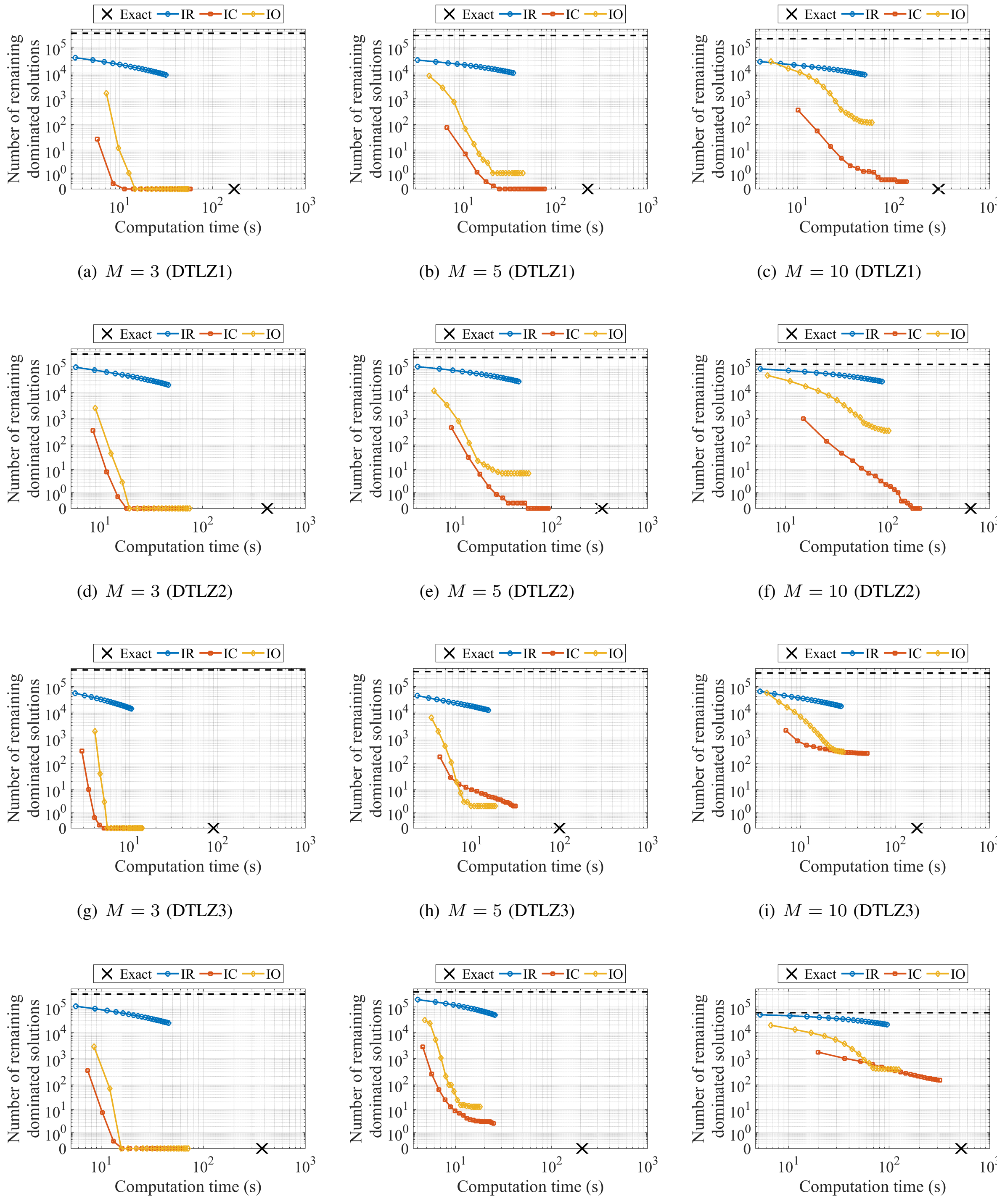


(a) $M = 3$ (WFG7) (b) $M = 5$ (WFG7) (c) $M = 10$ (WFG7)

(d) $M = 3$ (WFG8) (e) $M = 5$ (WFG8) (f) $M = 10$ (WFG8)

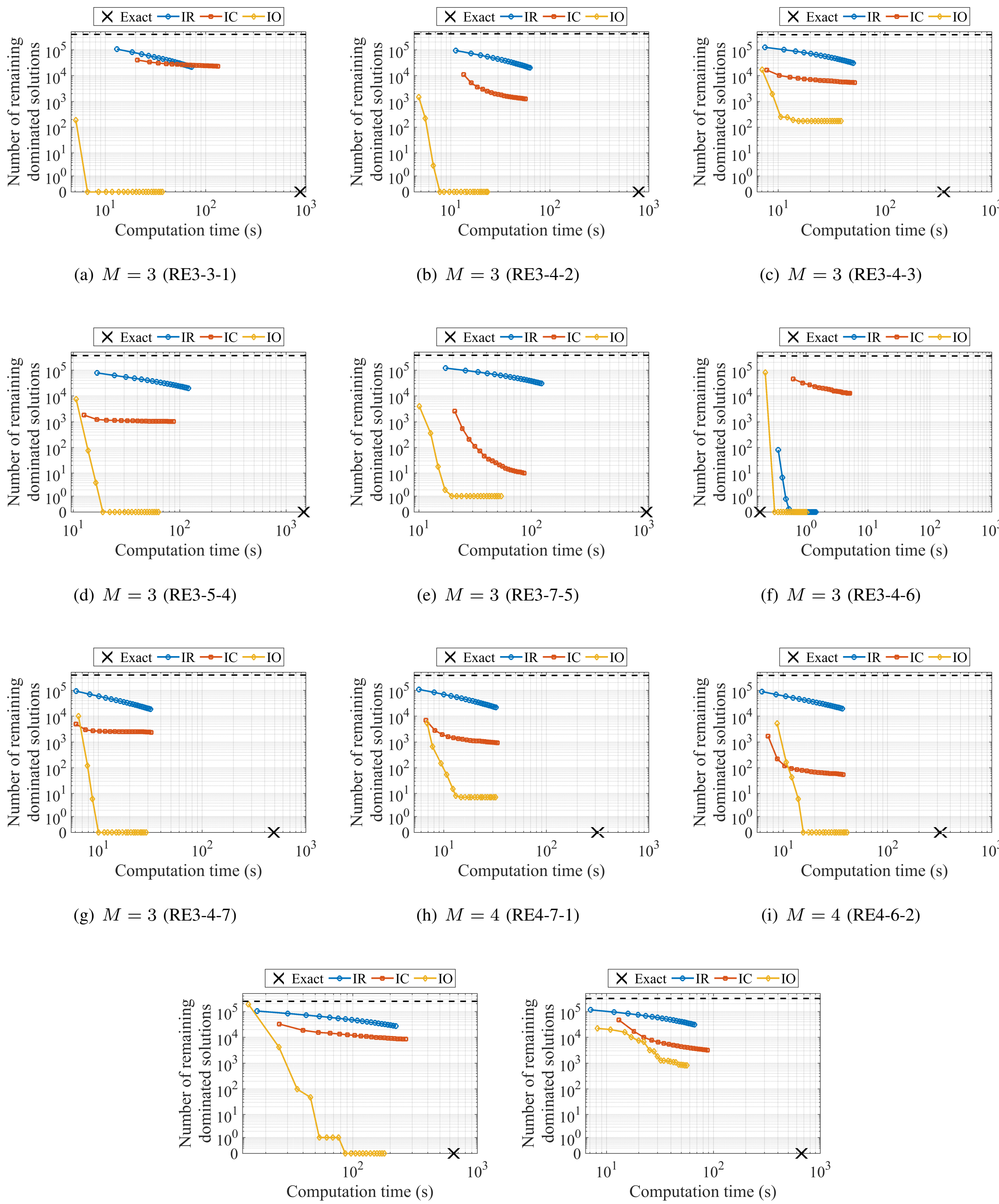
(g) $M = 3$ (WFG9) (h) $M = 5$ (WFG9) (i) $M = 10$ (WFG9)

Fig. 3. Performance comparison of different methods for removing dominated solutions from the WFG7-9 problems. Average values over ten runs are shown.



(a) $M = 3$ (DTLZ1)
(b) $M = 5$ (DTLZ1)
(c) $M = 10$ (DTLZ1)
(d) $M = 3$ (DTLZ2)
(e) $M = 5$ (DTLZ2)
(f) $M = 10$ (DTLZ2)
(g) $M = 3$ (DTLZ3)
(h) $M = 5$ (DTLZ3)
(i) $M = 10$ (DTLZ3)
(j) $M = 3$ (DTLZ4)
(k) $M = 5$ (DTLZ4)
(l) $M = 10$ (DTLZ4)

Fig. 4. Performance comparison of different methods for removing dominated solutions from the DTLZ1-4 problems. Average values over ten runs are shown.

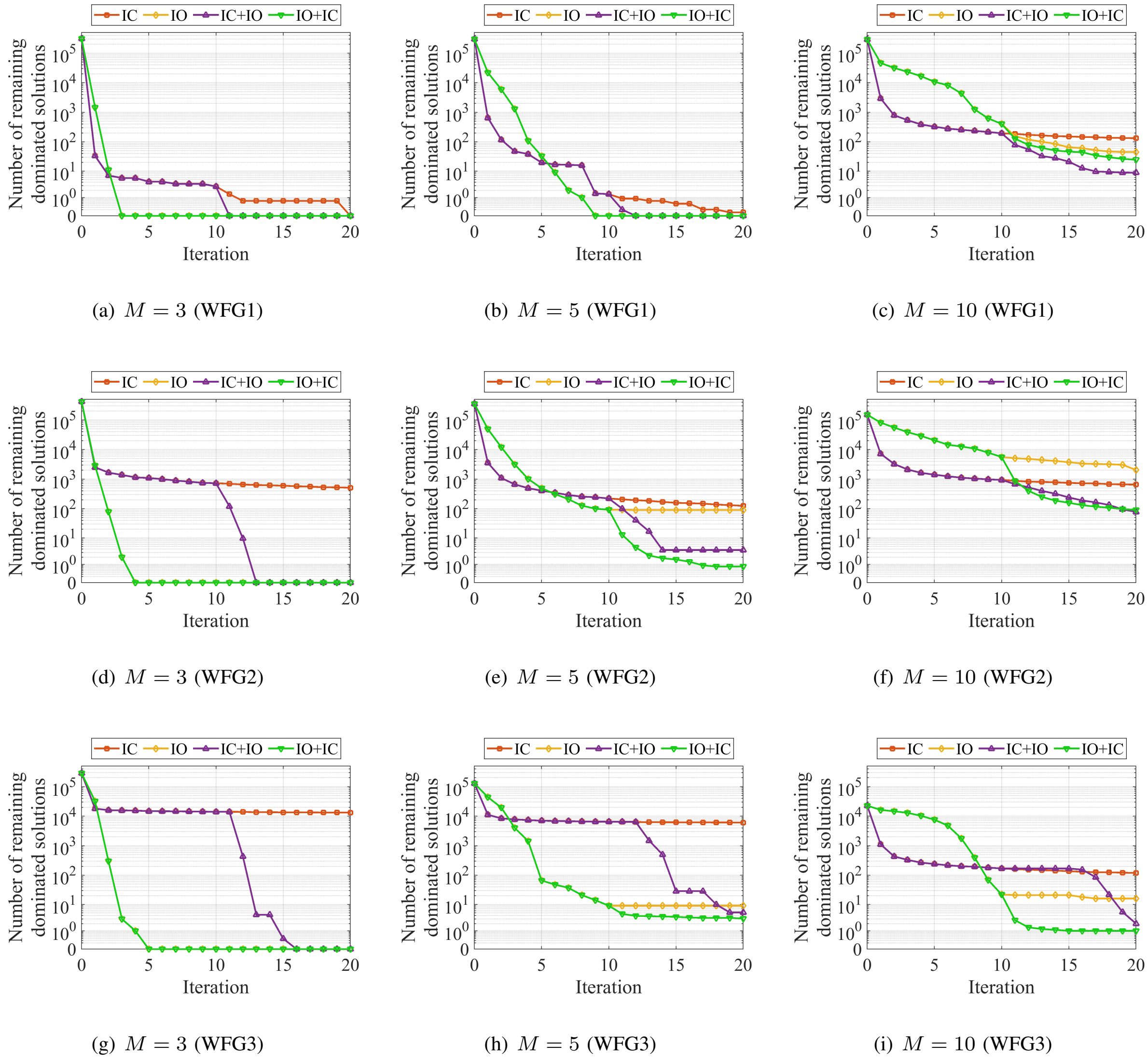


(a) $M = 3$ (RE3-3-1)
(b) $M = 3$ (RE3-4-2)
(c) $M = 3$ (RE3-4-3)
(d) $M = 3$ (RE3-5-4)
(e) $M = 3$ (RE3-7-5)
(f) $M = 3$ (RE3-4-6)
(g) $M = 3$ (RE3-4-7)
(h) $M = 4$ (RE4-7-1)
(i) $M = 4$ (RE4-6-2)
(j) $M = 6$ (RE6-3-1)
(k) $M = 9$ (RE9-7-1)

Fig. 5. Performance comparison of different methods for removing dominated solutions from the RE problems. Average values over ten runs are shown.

S2. Comparison between Hybrid Versions and Original Versions

Figure 6-8, Figure 9 and Figure 10 compare the hybrid partition methods and the original partition methods for the WFG, DTLZ and RE problems, respectively.

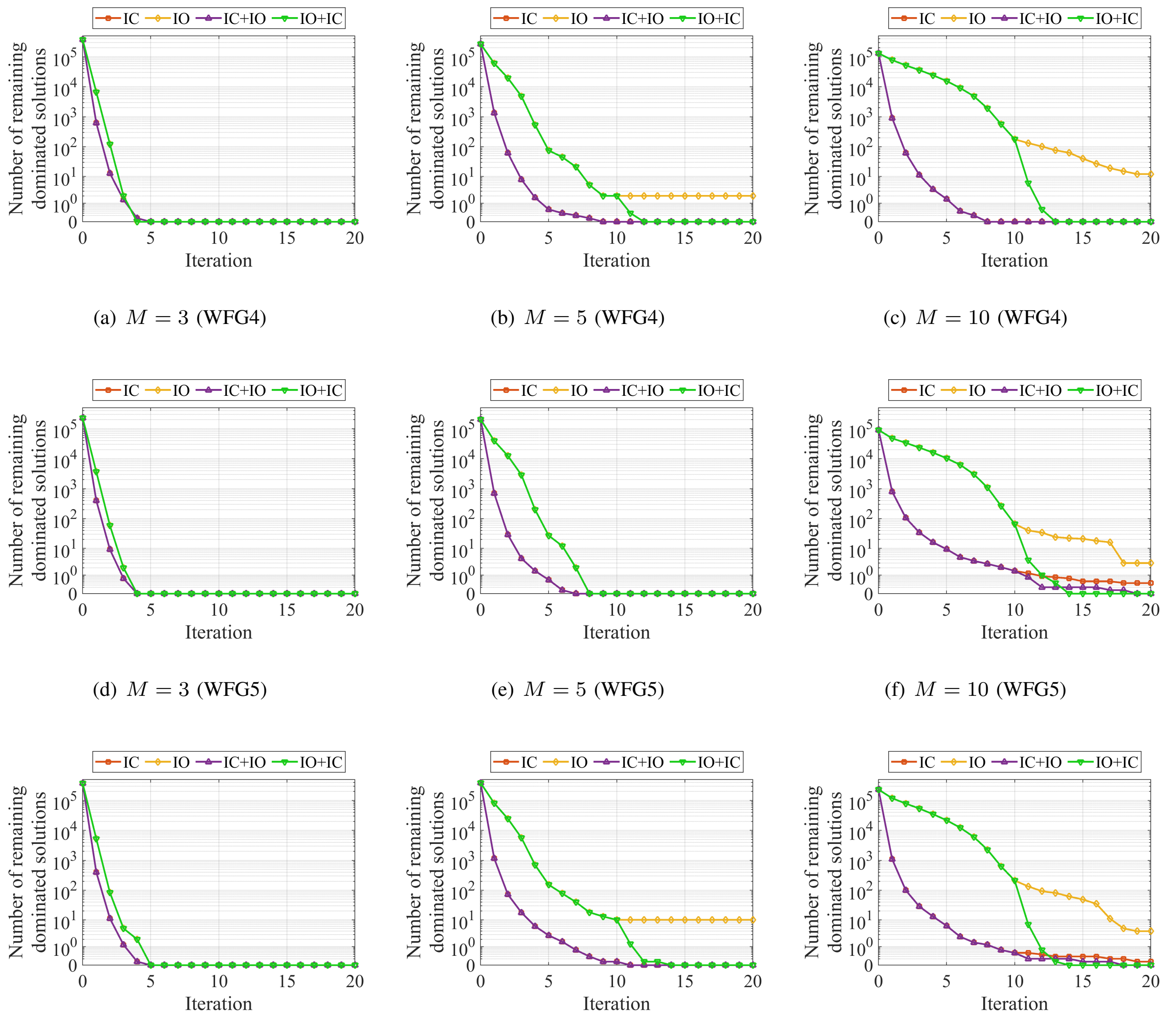


(a) $M = 3$ (WFG1) (b) $M = 5$ (WFG1) (c) $M = 10$ (WFG1)

(d) $M = 3$ (WFG2) (e) $M = 5$ (WFG2) (f) $M = 10$ (WFG2)

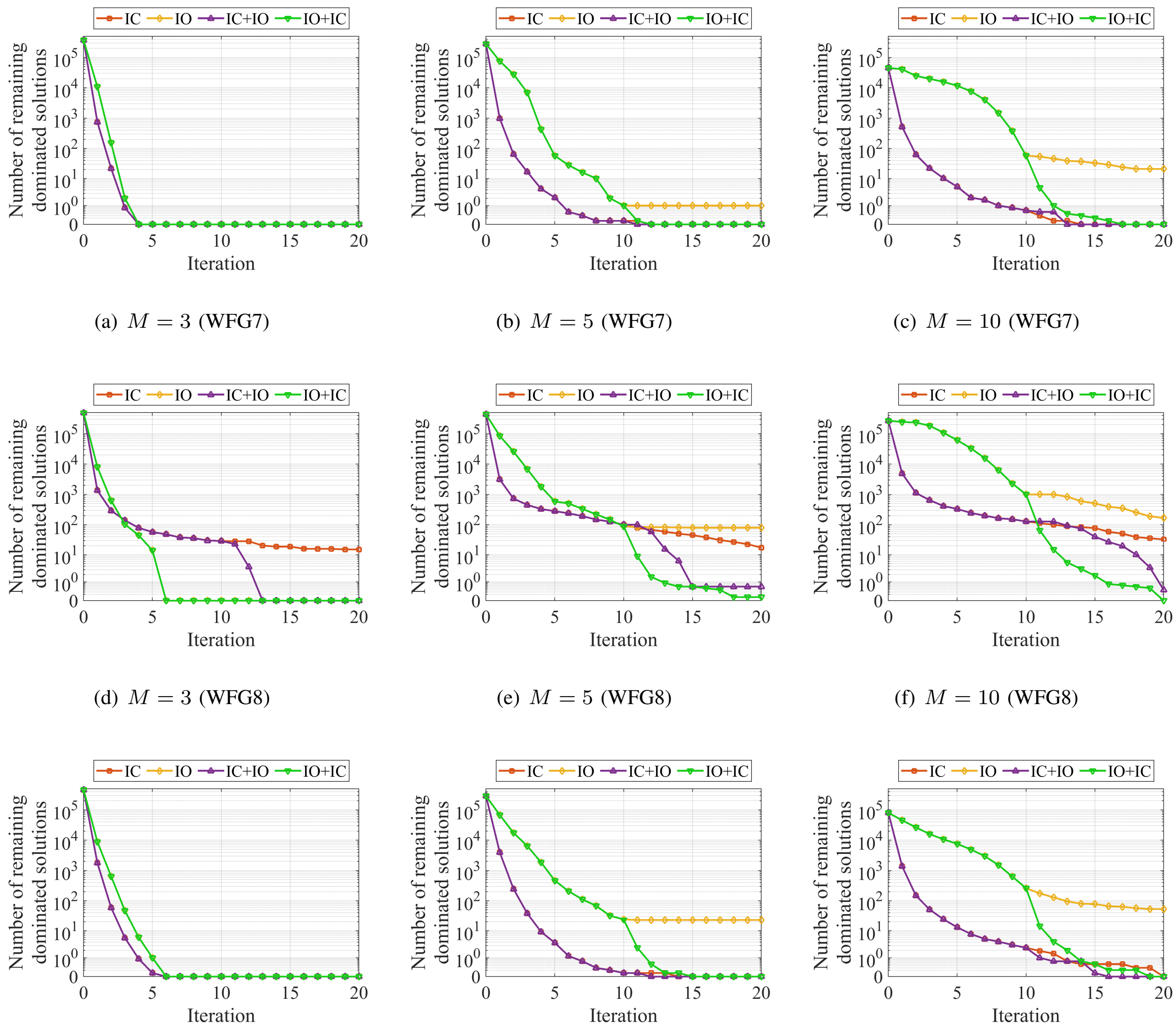
(g) $M = 3$ (WFG3) (h) $M = 5$ (WFG3) (i) $M = 10$ (WFG3)

Fig. 6. Comparison between the hybrid partition methods and the original partition methods for the WFG1-3 problems.



(a) $M = 3$ (WFG4)
(b) $M = 5$ (WFG4)
(c) $M = 10$ (WFG4)
(d) $M = 3$ (WFG5)
(e) $M = 5$ (WFG5)
(f) $M = 10$ (WFG5)
(g) $M = 3$ (WFG6)
(h) $M = 5$ (WFG6)
(i) $M = 10$ (WFG6)

Fig. 7. Comparison between the hybrid partition methods and the original partition methods for the WFG4-6 problems.

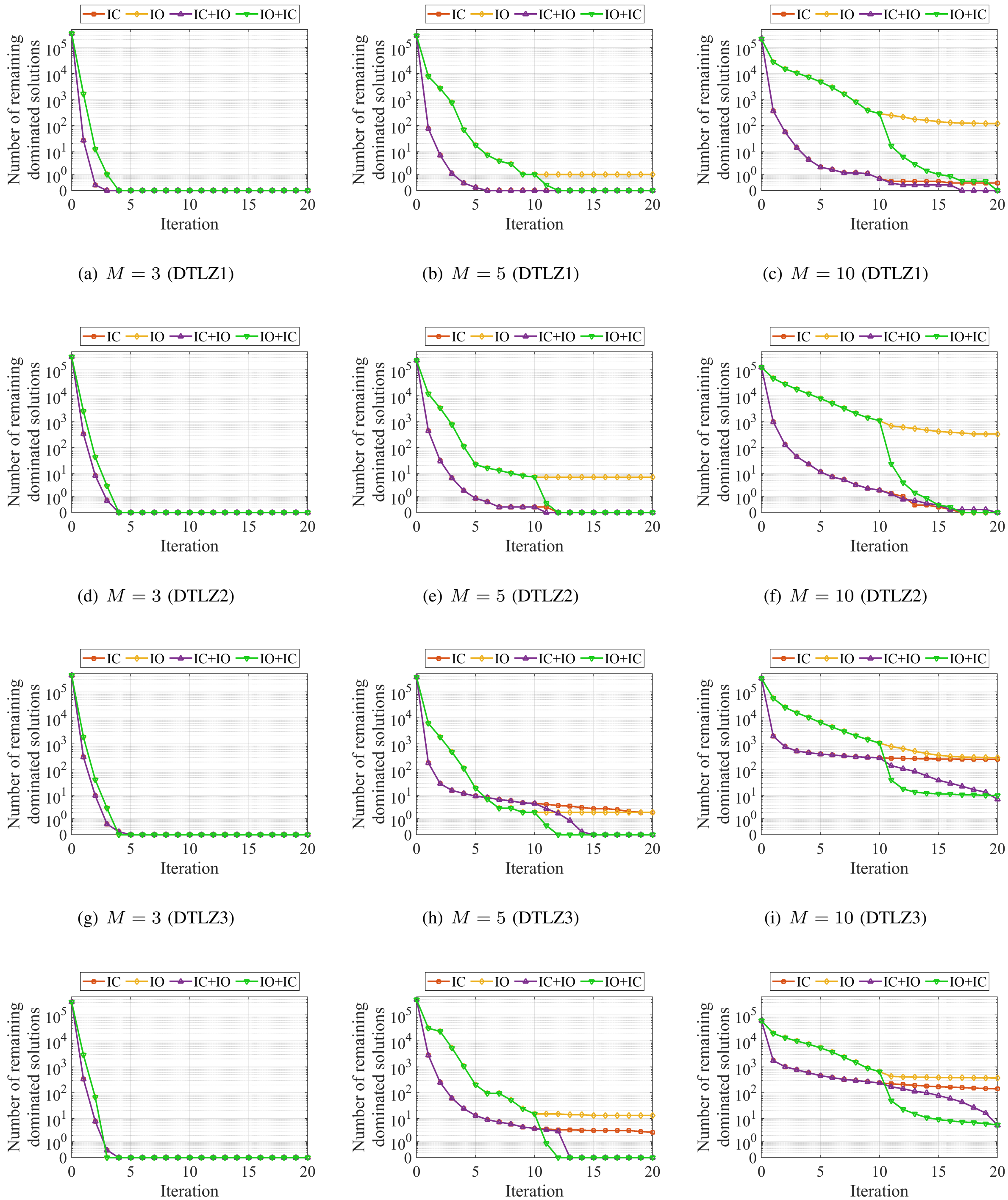


(a) $M = 3$ (WFG7) (b) $M = 5$ (WFG7) (c) $M = 10$ (WFG7)

(d) $M = 3$ (WFG8) (e) $M = 5$ (WFG8) (f) $M = 10$ (WFG8)

(g) $M = 3$ (WFG9) (h) $M = 5$ (WFG9) (i) $M = 10$ (WFG9)

Fig. 8. Comparison between the hybrid partition methods and the original partition methods for the WFG7-9 problems.



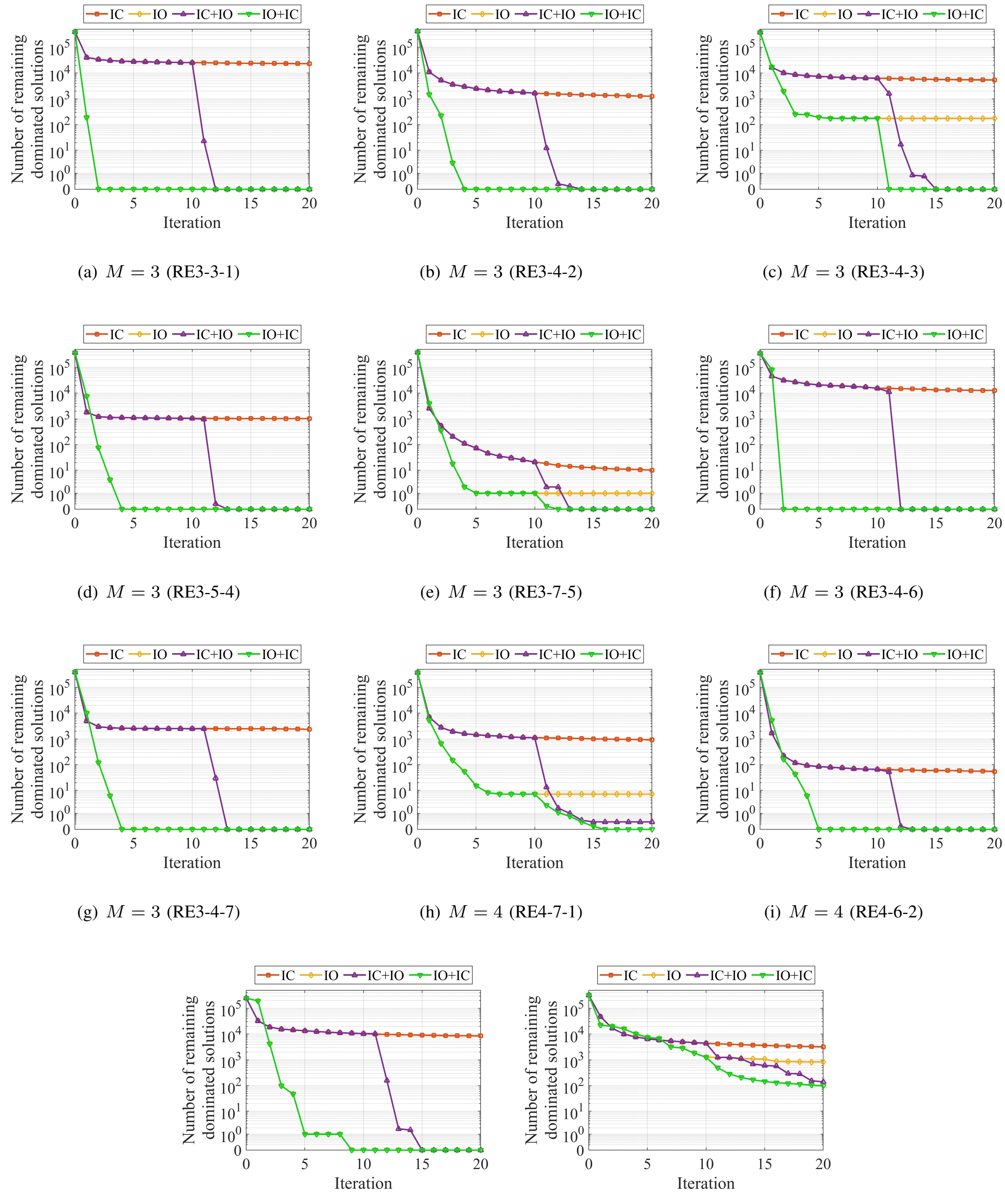
(a) $M = 3$ (DTLZ1) (b) $M = 5$ (DTLZ1) (c) $M = 10$ (DTLZ1)

(d) $M = 3$ (DTLZ2) (e) $M = 5$ (DTLZ2) (f) $M = 10$ (DTLZ2)

(g) $M = 3$ (DTLZ3) (h) $M = 5$ (DTLZ3) (i) $M = 10$ (DTLZ3)

(j) $M = 3$ (DTLZ4) (k) $M = 5$ (DTLZ4) (l) $M = 10$ (DTLZ4)

Fig. 9. Comparison between the hybrid partition methods and the original partition methods for the DTLZ1-4 problems.



(a) $M = 3$ (RE3-3-1)

(b) $M = 3$ (RE3-4-2)

(c) $M = 3$ (RE3-4-3)

(d) $M = 3$ (RE3-5-4)

(e) $M = 3$ (RE3-7-5)

(f) $M = 3$ (RE3-4-6)

(g) $M = 3$ (RE3-4-7)

(h) $M = 4$ (RE4-7-1)

(i) $M = 4$ (RE4-6-2)

(j) $M = 6$ (RE6-3-1)

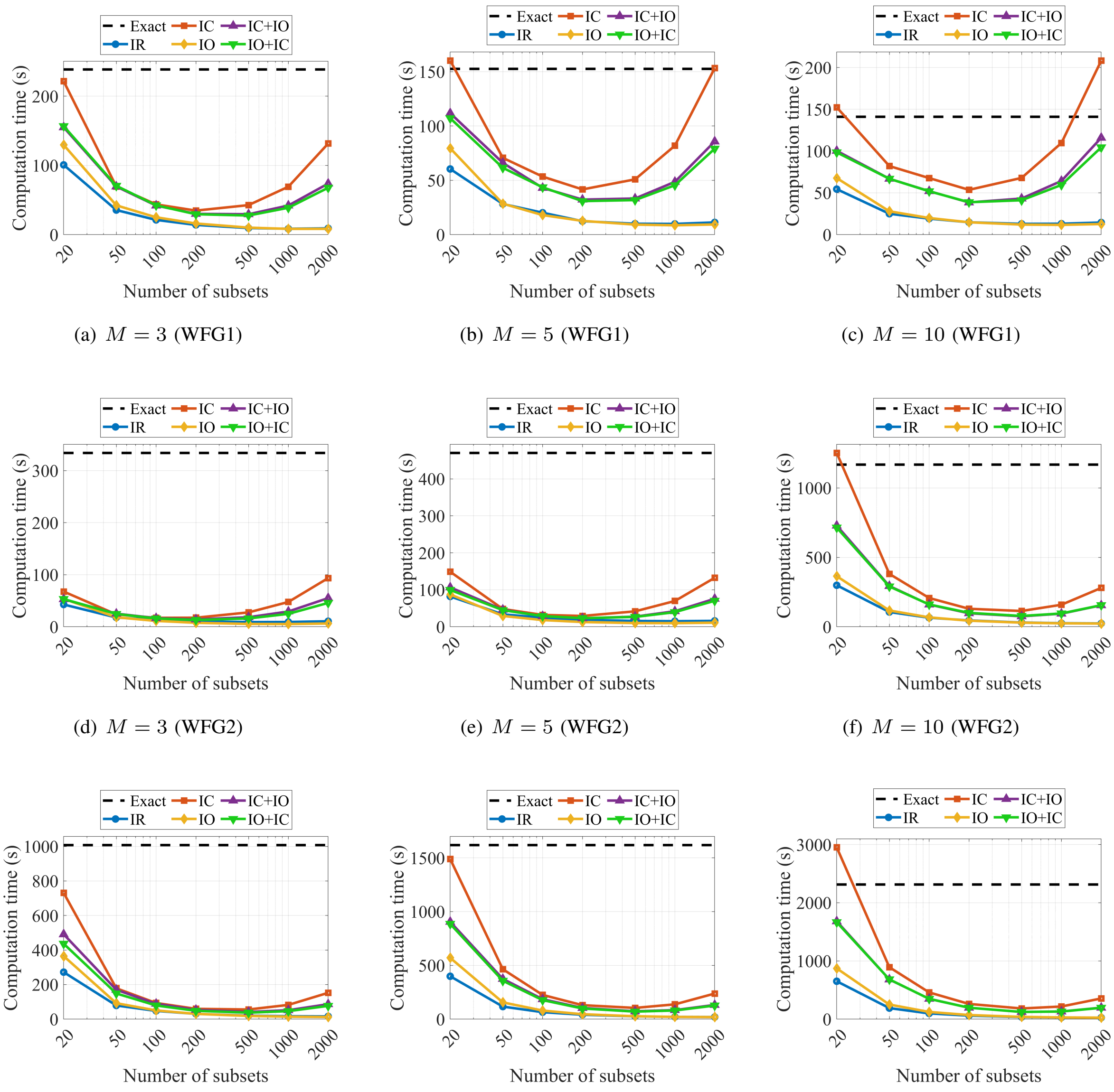
(k) $M = 9$ (RE9-7-1)

Fig. 10. Comparison between the hybrid partition methods and the original partition methods for the RE problems.

S3. Effects of the Number of Subsets for the Performance of Partition Methods

A. Computation Time

Figure 11-13, Figure 14 and Figure 15 show the effects of the number of subsets on the computation time of each partition method for the WFG, DTLZ and RE problems, respectively.

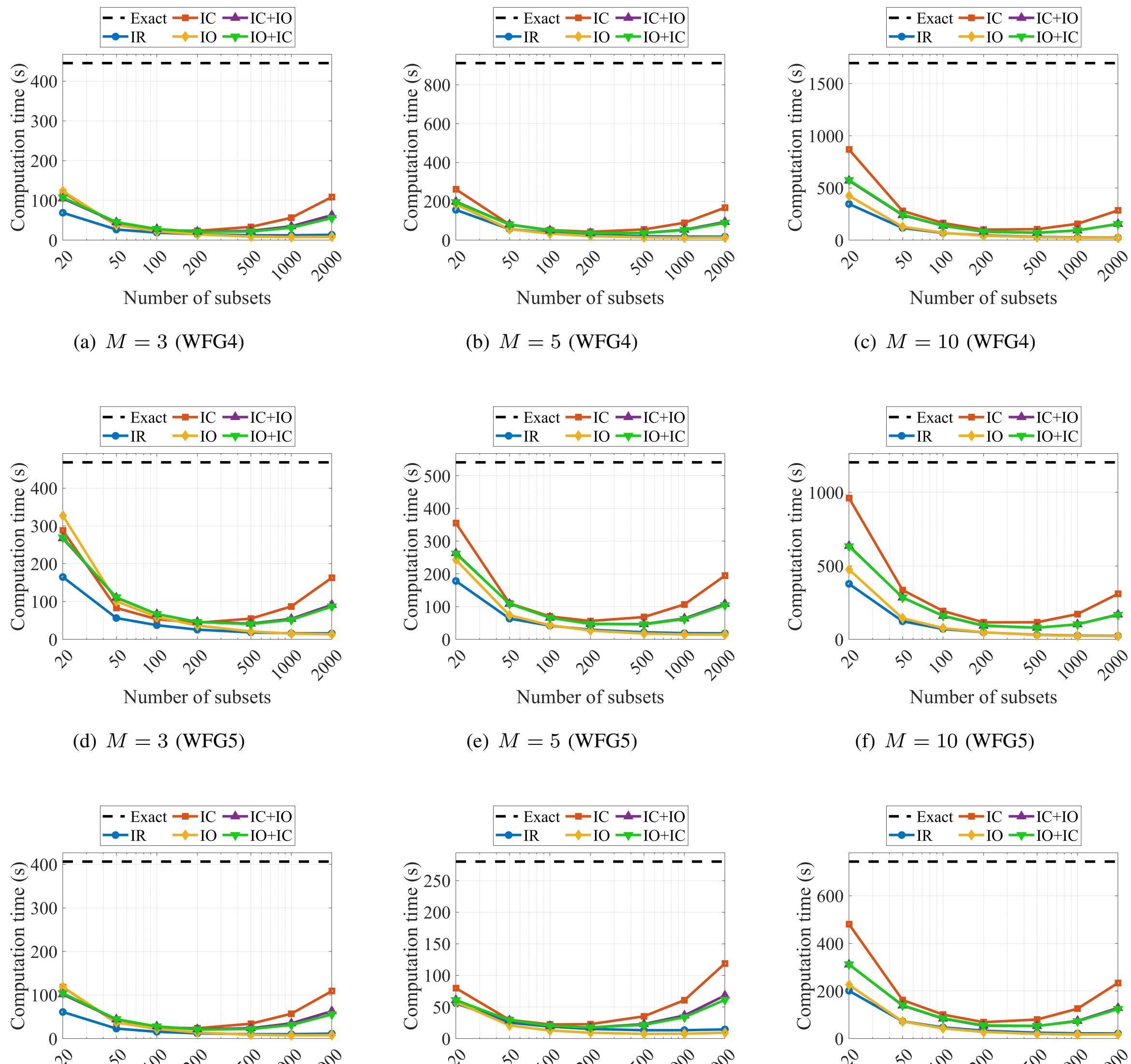


(a) $M = 3$ (WFG1) (b) $M = 5$ (WFG1) (c) $M = 10$ (WFG1)

(d) $M = 3$ (WFG2) (e) $M = 5$ (WFG2) (f) $M = 10$ (WFG2)

(g) $M = 3$ (WFG3) (h) $M = 5$ (WFG3) (i) $M = 10$ (WFG3)

Fig. 11. Computation time of each partition method with different specifications of the number of subsets for the WFG1-3 problems. The number of iterations is 20 in all partition methods. The dashed horizontal line shows the computation time of the exact method (where the number of subsets is one).



(a) $M = 3$ (WFG4) (b) $M = 5$ (WFG4) (c) $M = 10$ (WFG4)

(d) $M = 3$ (WFG5) (e) $M = 5$ (WFG5) (f) $M = 10$ (WFG5)

(g) $M = 3$ (WFG6) (h) $M = 5$ (WFG6) (i) $M = 10$ (WFG6)

Fig. 12. Computation time of each partition method with different specifications of the number of subsets for the WFG4-6 problems. The number of iterations is 20 in all partition methods. The dashed horizontal line shows the computation time of the exact method (where the number of subsets is one).

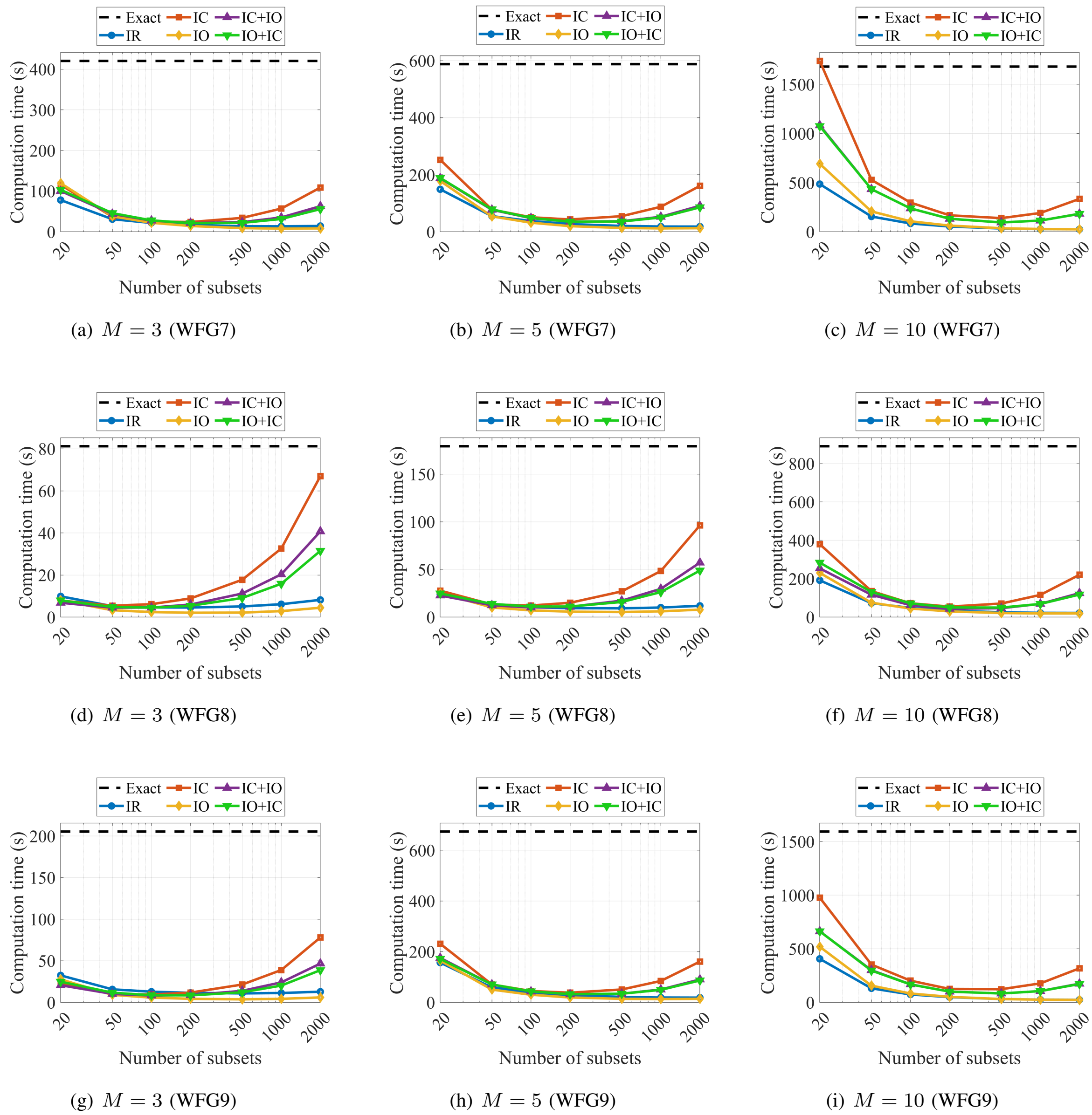
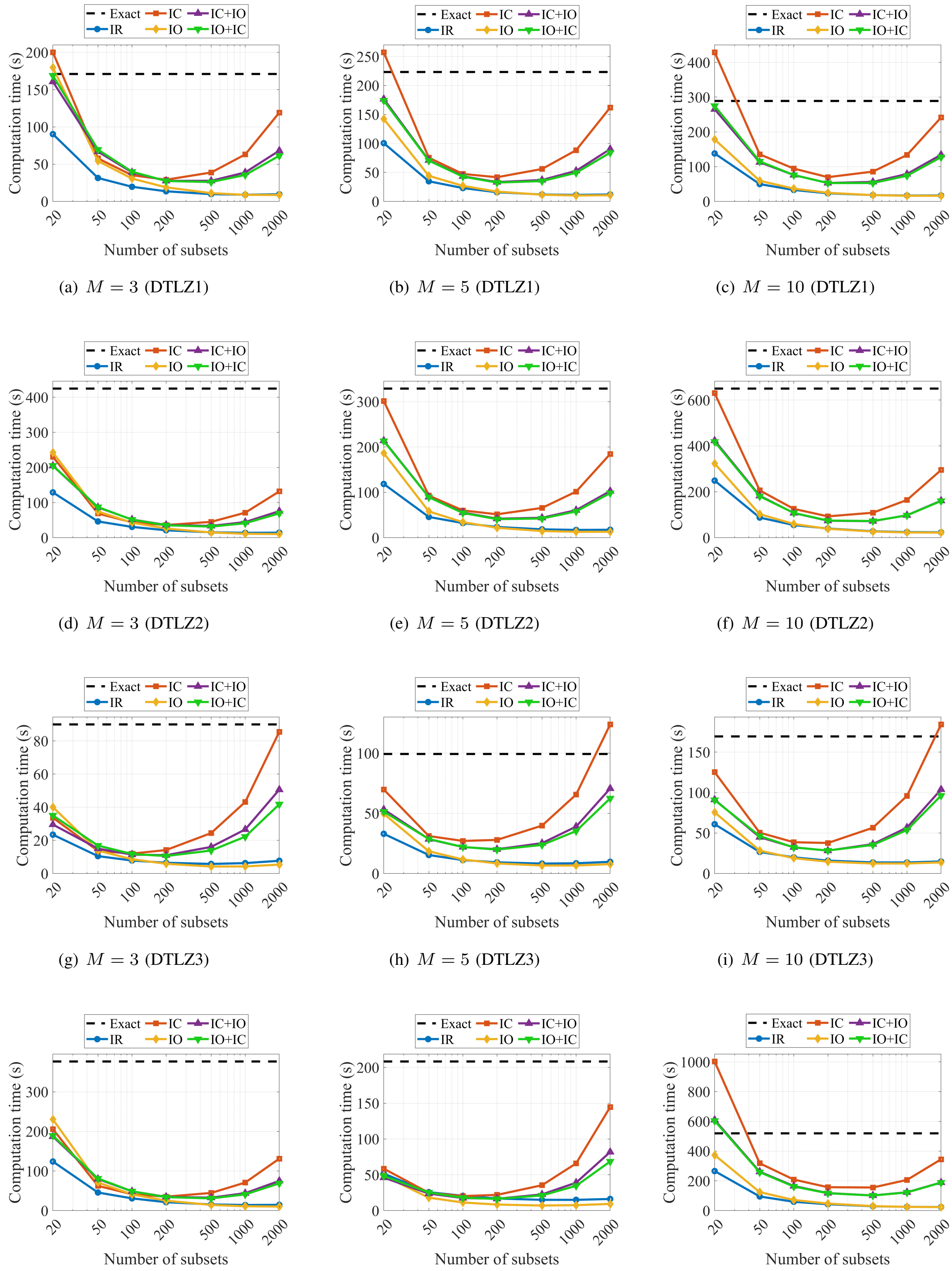


Fig. 13. Computation time of each partition method with different specifications of the number of subsets for the WFG7-9 problems. The number of iterations is 20 in all partition methods. The dashed horizontal line shows the computation time of the exact method (where the number of subsets is one).



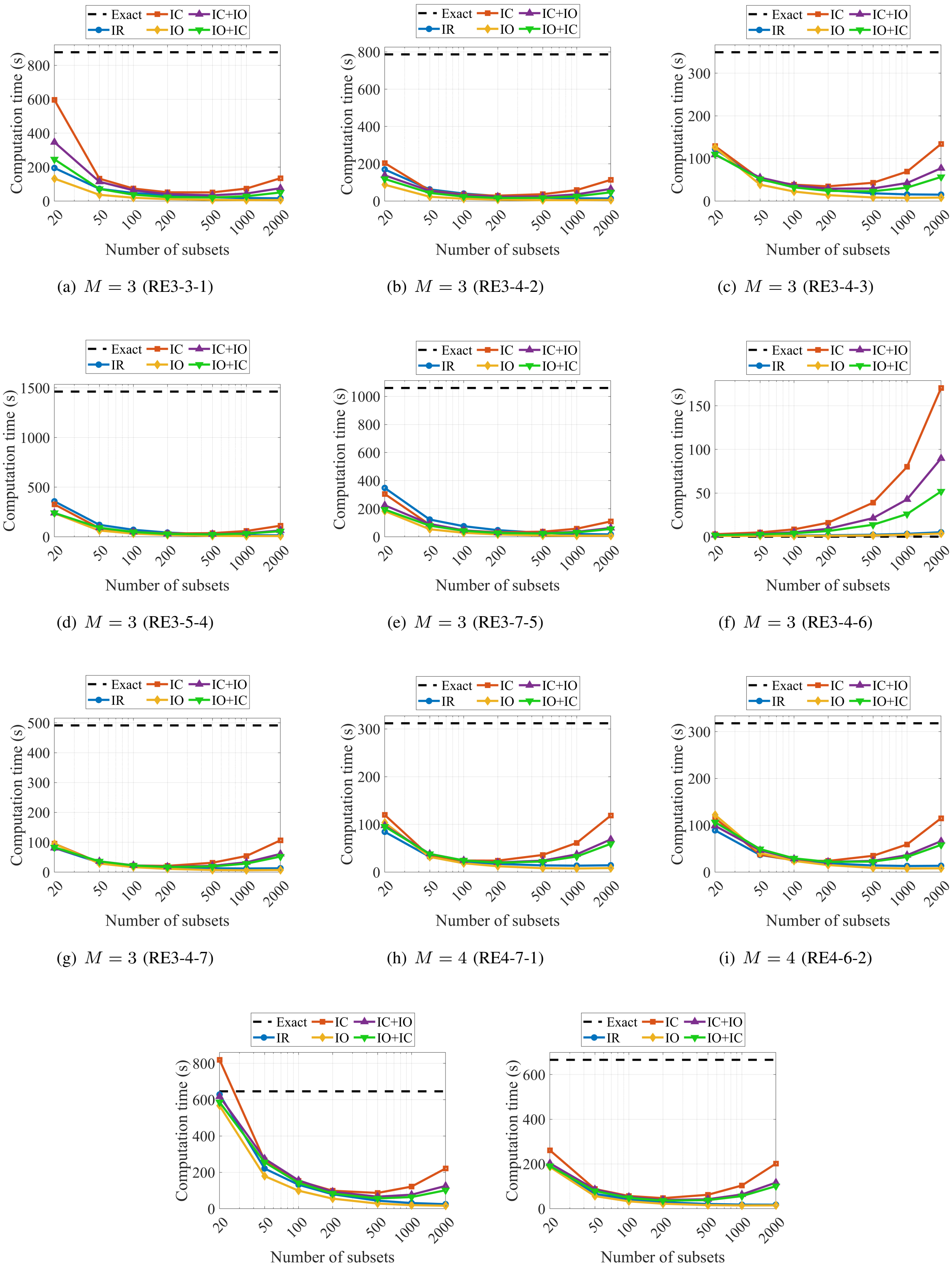
(a) $M = 3$ (DTLZ1) (b) $M = 5$ (DTLZ1) (c) $M = 10$ (DTLZ1)

(d) $M = 3$ (DTLZ2) (e) $M = 5$ (DTLZ2) (f) $M = 10$ (DTLZ2)

(g) $M = 3$ (DTLZ3) (h) $M = 5$ (DTLZ3) (i) $M = 10$ (DTLZ3)

(j) $M = 3$ (DTLZ4) (k) $M = 5$ (DTLZ4) (l) $M = 10$ (DTLZ4)

Fig. 14. Computation time of each partition method with different specifications of the number of subsets for the DTLZ1-4 problems. The number of iterations is 20 in all partition methods. The dashed horizontal line shows the computation time of the exact method (where the number of subsets is one).



(a) $M = 3$ (RE3-3-1)

(b) $M = 3$ (RE3-4-2)

(c) $M = 3$ (RE3-4-3)

(d) $M = 3$ (RE3-5-4)

(e) $M = 3$ (RE3-7-5)

(f) $M = 3$ (RE3-4-6)

(g) $M = 3$ (RE3-4-7)

(h) $M = 4$ (RE4-7-1)

(i) $M = 4$ (RE4-6-2)

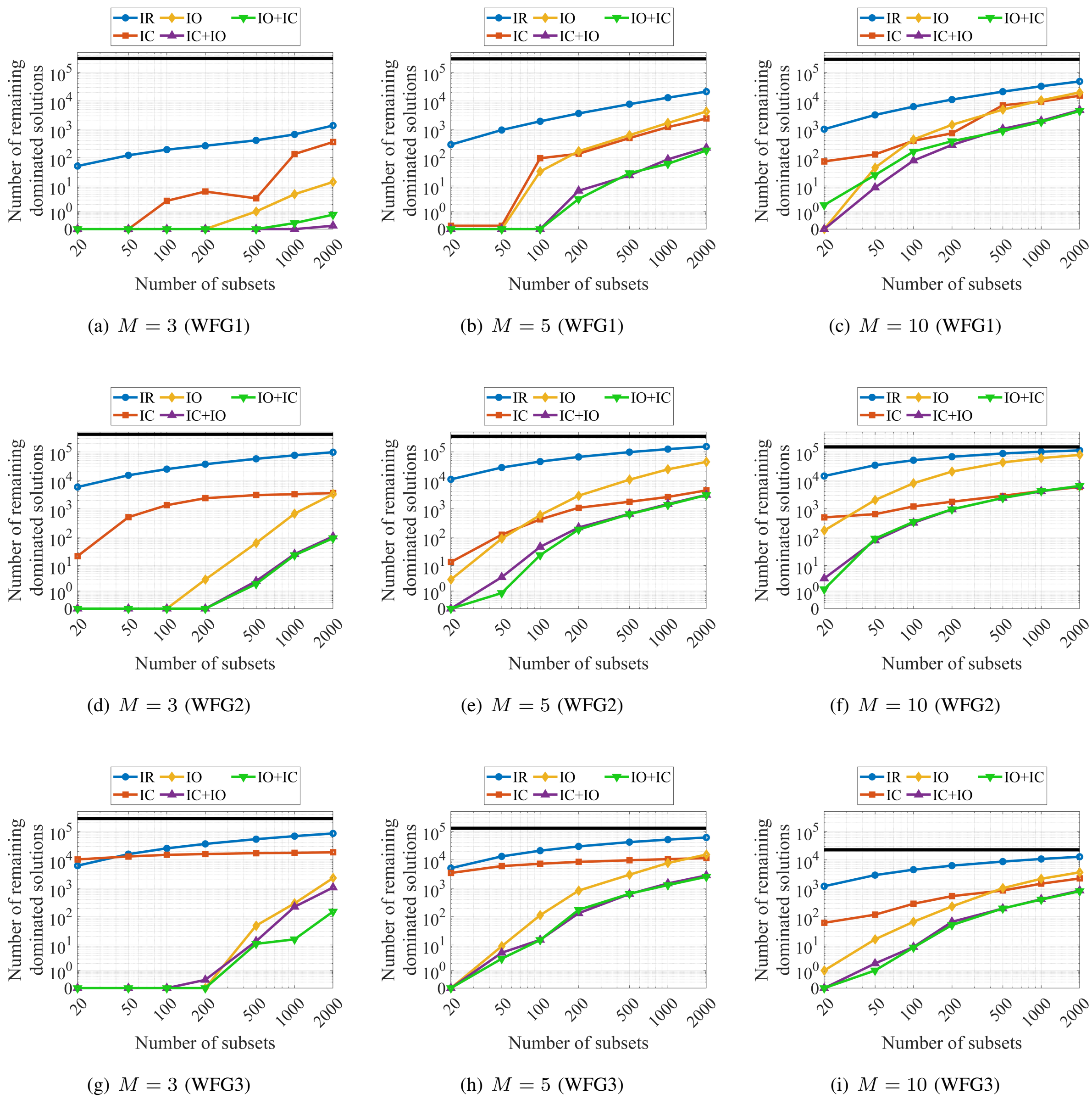
(j) $M = 6$ (RE6-3-1)

(k) $M = 9$ (RE9-7-1)

Fig. 15. Computation time of each partition method with different specifications of the number of subsets for the RE problems. The number of iterations is 20 in all partition methods. The dashed horizontal line shows the computation time of the exact method (where the number of subsets is one).

B. Number of Remaining Dominated Solutions

Figure 16-18, Figure 19 and Figure 20 show the effects of the number of subsets on the removal performance (i.e., the number of remaining dominated solutions) of each partition method for the WFG, DTLZ and RE problems, respectively.

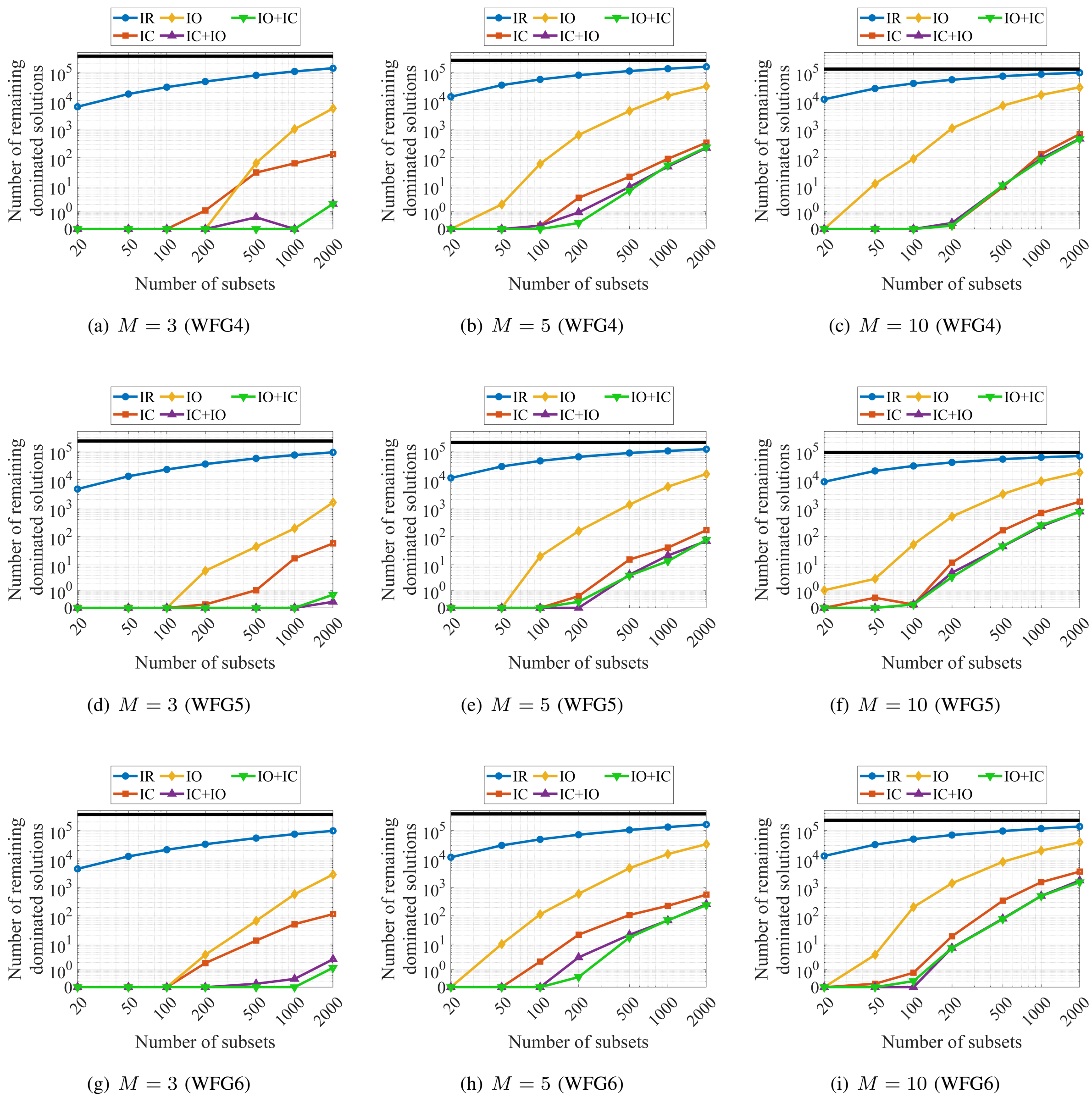


(a) $M = 3$ (WFG1) (b) $M = 5$ (WFG1) (c) $M = 10$ (WFG1)

(d) $M = 3$ (WFG2) (e) $M = 5$ (WFG2) (f) $M = 10$ (WFG2)

(g) $M = 3$ (WFG3) (h) $M = 5$ (WFG3) (i) $M = 10$ (WFG3)

Fig. 16. Number of remaining dominated solutions by each partition method with different specifications of the number of subsets for the WFG1-3 problems. The bold horizontal line shows the number of dominated solutions included in each solution set with 500,000 solutions.



(a) $M = 3$ (WFG4)

(b) $M = 5$ (WFG4)

(c) $M = 10$ (WFG4)

(d) $M = 3$ (WFG5)

(e) $M = 5$ (WFG5)

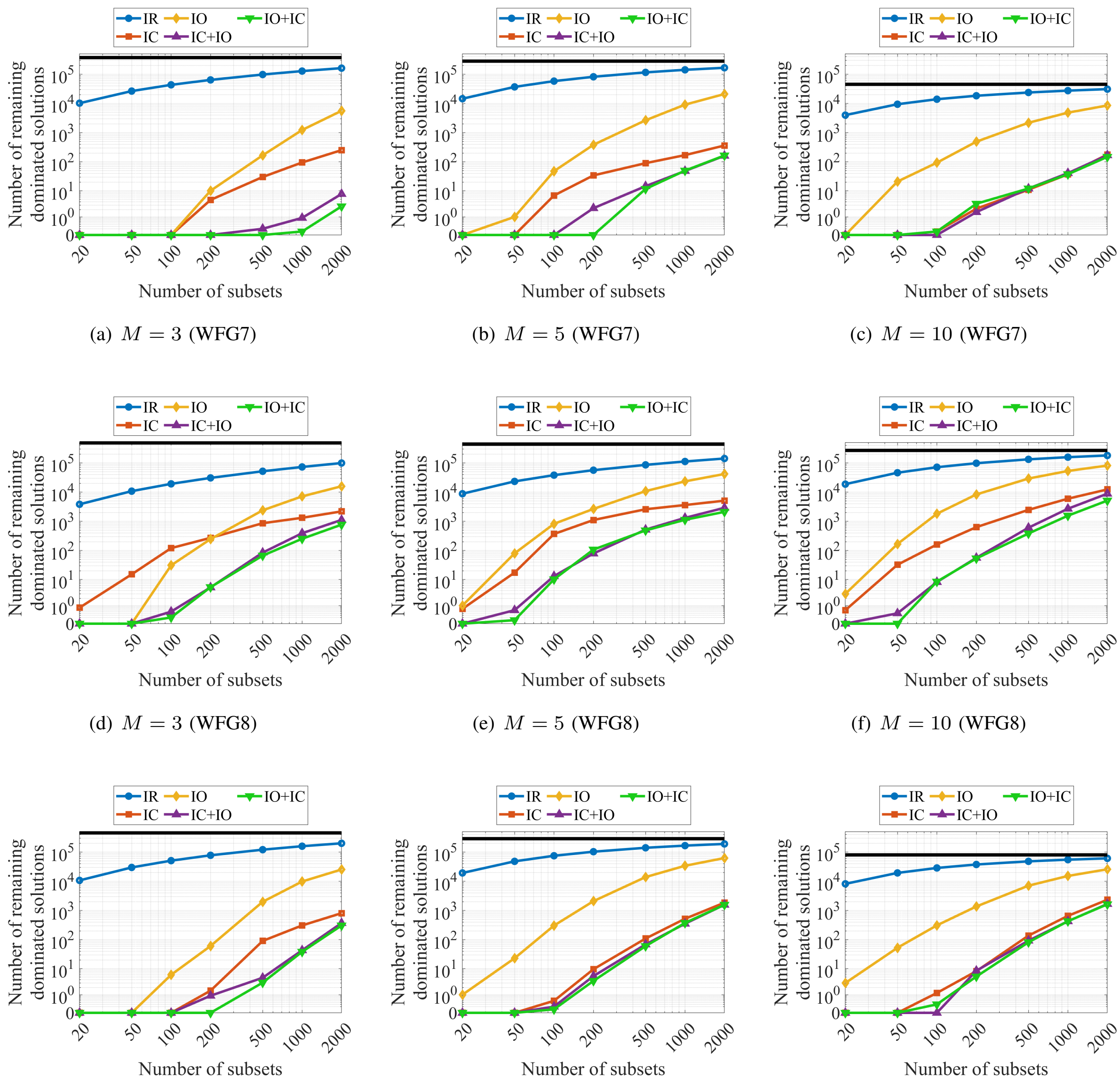
(f) $M = 10$ (WFG5)

(g) $M = 3$ (WFG6)

(h) $M = 5$ (WFG6)

(i) $M = 10$ (WFG6)

Fig. 17. Number of remaining dominated solutions by each partition method with different specifications of the number of subsets for the WFG4-6 problems. The bold horizontal line shows the number of dominated solutions included in each solution set with 500,000 solutions.

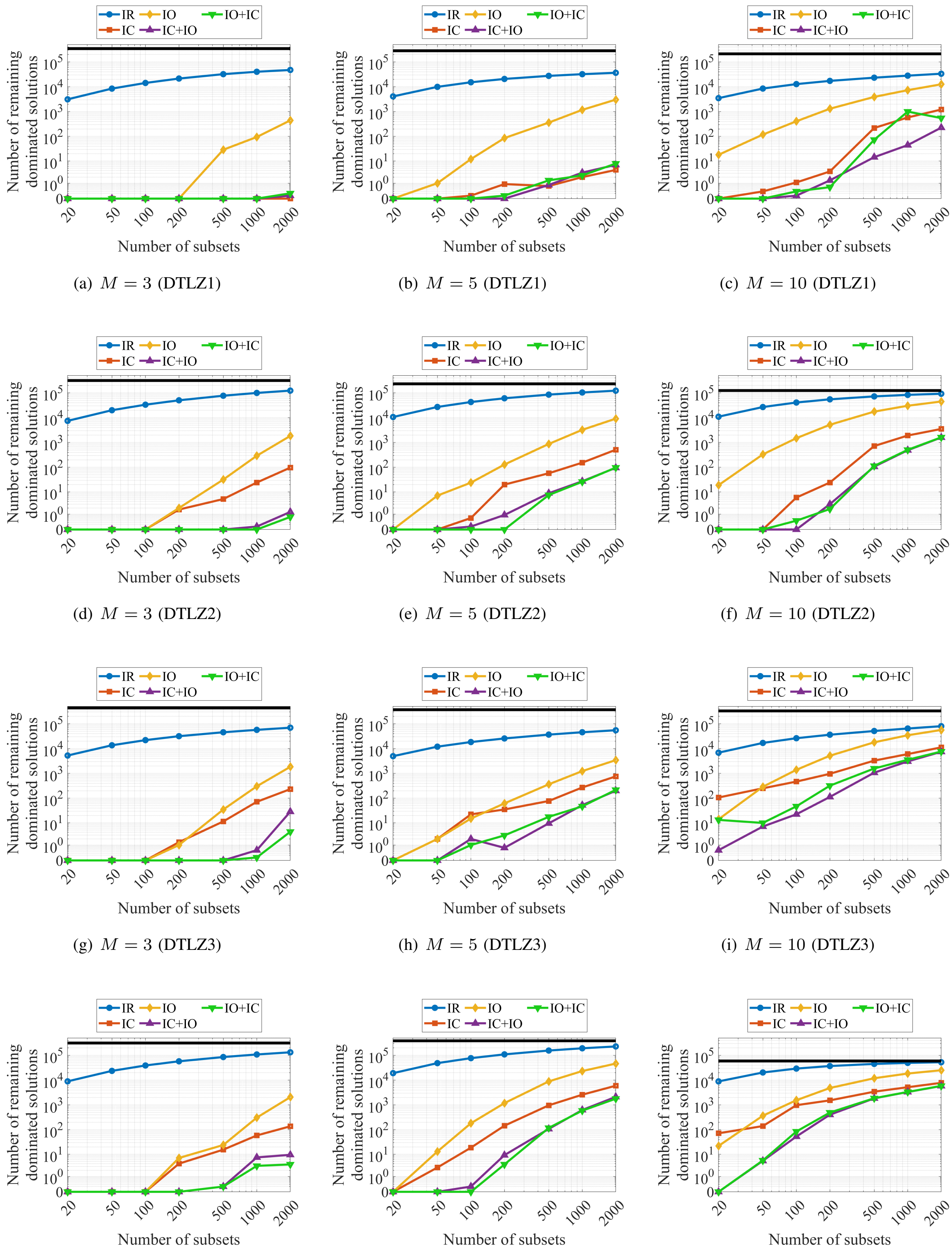


(a) $M = 3$ (WFG7) (b) $M = 5$ (WFG7) (c) $M = 10$ (WFG7)

(d) $M = 3$ (WFG8) (e) $M = 5$ (WFG8) (f) $M = 10$ (WFG8)

(g) $M = 3$ (WFG9) (h) $M = 5$ (WFG9) (i) $M = 10$ (WFG9)

Fig. 18. Number of remaining dominated solutions by each partition method with different specifications of the number of subsets for the WFG7-9 problems. The bold horizontal line shows the number of dominated solutions included in each solution set with 500,000 solutions.



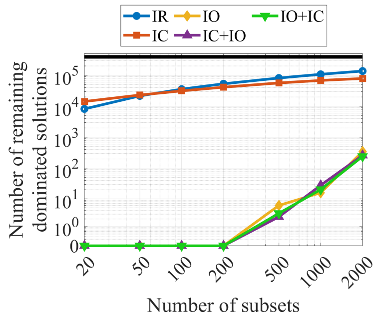
(a) $M = 3$ (DTLZ1) (b) $M = 5$ (DTLZ1) (c) $M = 10$ (DTLZ1)

(d) $M = 3$ (DTLZ2) (e) $M = 5$ (DTLZ2) (f) $M = 10$ (DTLZ2)

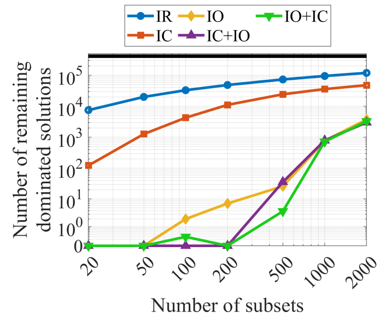
(g) $M = 3$ (DTLZ3) (h) $M = 5$ (DTLZ3) (i) $M = 10$ (DTLZ3)

(j) $M = 3$ (DTLZ4) (k) $M = 5$ (DTLZ4) (l) $M = 10$ (DTLZ4)

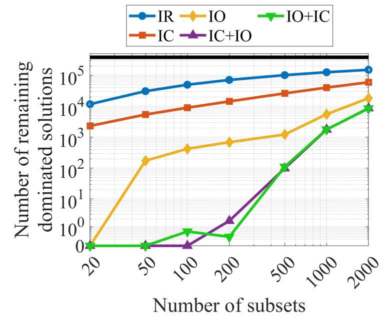
Fig. 19. Number of remaining dominated solutions by each partition method with different specifications of the number of subsets for the DTLZ1-4 problems. The bold horizontal line shows the number of dominated solutions included in each solution set with 500,000 solutions.



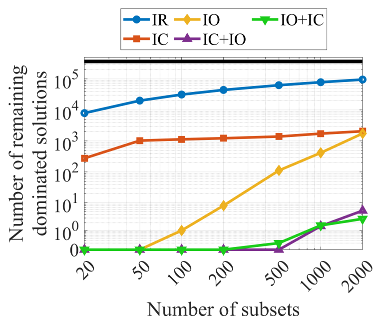
(a) $M = 3$ (RE3-3-1)



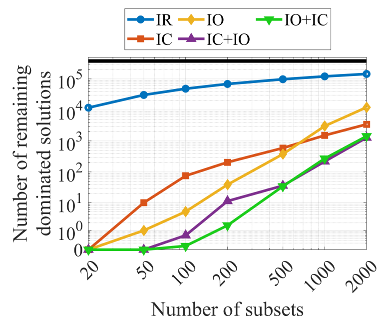
(b) $M = 3$ (RE3-4-2)



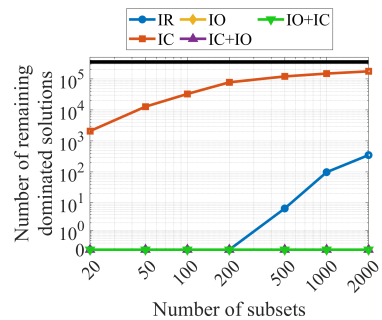
(c) $M = 3$ (RE3-4-3)



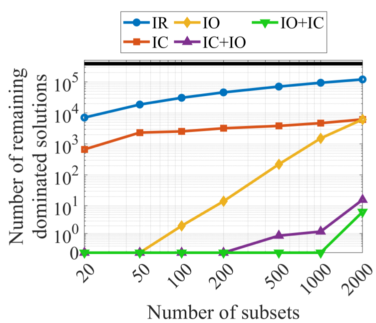
(d) $M = 3$ (RE3-5-4)



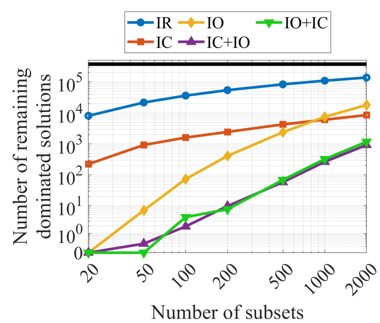
(e) $M = 3$ (RE3-7-5)



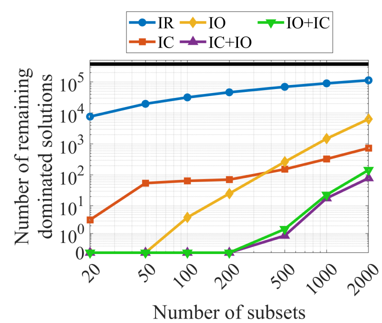
(f) $M = 3$ (RE3-4-6)



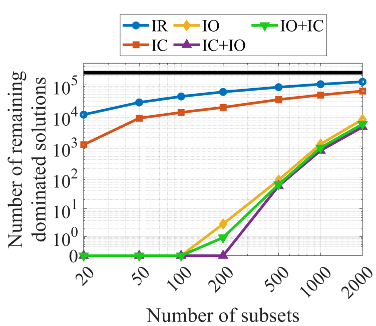
(g) $M = 3$ (RE3-4-7)



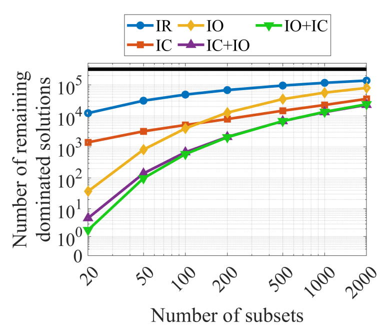
(h) $M = 4$ (RE4-7-1)



(i) $M = 4$ (RE4-6-2)



(j) $M = 6$ (RE6-3-1)



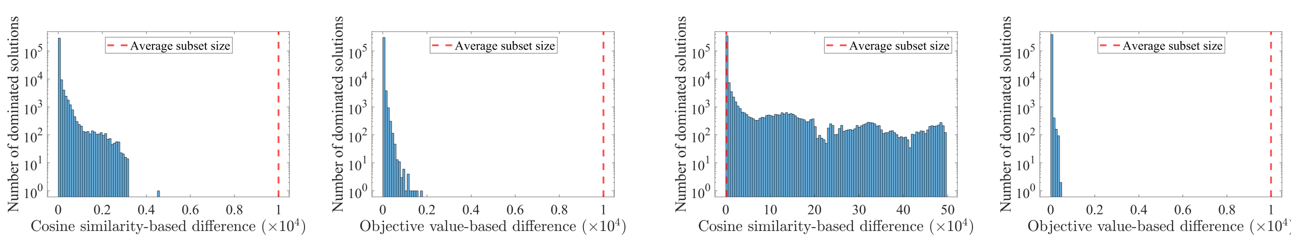
(k) $M = 9$ (RE9-7-1)

Fig. 20. Number of remaining dominated solutions by each partition method with different specifications of the number of subsets for the RE problems. The bold horizontal line shows the number of dominated solutions included in each solution set with 500,000 solutions.

S4. Effects of the Solution Distribution

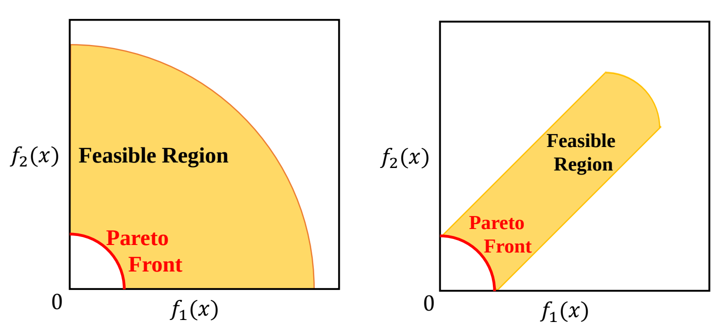
To further analyze these results, we quantify the cosine similarity-based difference and the objective value-based difference between the dominated solutions and their corresponding dominating solutions in the following manner. For a dominated solution $\boldsymbol{s}$ in the solution set S, we rank the other solutions in S based on their cosine similarity with $\boldsymbol{s}$. The most similar solution to $\boldsymbol{s}$ has Rank 1 and the least similar solution has Rank $(N-1)$ where N is the solution set size ($N = 500,000$ in our computational experiments). The cosine similarity-based difference $Diff_c(\boldsymbol{s}, S)$ is defined as the minimum rank over all dominating solutions of $\boldsymbol{s}$. As we can see, larger $Diff_c(\boldsymbol{s}, S)$ means that it is more difficult for the cosine similarity-based partition method to include the dominated solution $\boldsymbol{s}$ and its corresponding dominating solutions(s) in the same subset. In a similar manner, we define the objective value-based difference $Diff_o(\boldsymbol{s}, S)$ using the objective value-based ranking of the solutions in S for each dominated solution $\boldsymbol{s}$. First, all solutions in S are sorted using the i-th objective value. The solution with the smallest objective value has Rank 1, and the solution with the largest objective value has Rank N. Let us denote the rank of solution $\boldsymbol{s}$ based on the i-th objective by $\sigma_i(\boldsymbol{s})$ for $i = 1, 2, ..., M$. The objective value-based difference between a dominated solution $\boldsymbol{s}$ and its dominating solution $\boldsymbol{s'}$ is defined by the difference between their ranks as $|\sigma_i(s) - \sigma_i(s')|$. When this objective value-based difference is small, it is likely that $\boldsymbol{s}$ and $\boldsymbol{s'}$ are included in the same subset by the objective value-based partition method for the i-th objective. Since all objectives are used in the iterated version, we define the objective value-based difference $Diff_o(s, S)$ between s and its dominating solution(s) in S as the minimum objective value-based difference over all dominating solutions for all objectives. That is, $Diff_o(\boldsymbol{s}, S) = min\{|\sigma_i(\boldsymbol{s}) - \sigma_i(\boldsymbol{s'})||1 \leq i \leq M, \boldsymbol{s'} \in S'\}$ where M is the number of objectives and S' is the set of solutions in S which dominate $\boldsymbol{s}$.

Figure 21 shows the cosine similarity-based difference and objective value-based difference distributions for all dominated solutions for the 3-objective DTLZ2 and RE3-3-1 problems. As a reference, the average size of subsets (i.e., 10,000) is shown by a red dotted line. Roughly speaking, a dominated solution $\boldsymbol{s}$ is likely to be group together with at least one of its corresponding dominating solution(s) by the objective value-based (cosine similarity-based) partition if its objective value-based (cosine similarity-based) difference is smaller than the average size of subsets. The dominated solutions and their corresponding dominating solutions have both small cosine similarity-based difference and small objective value-based difference in the 3-objective DTLZ2 problem as shown in Figure 21 (a). This is because the DTLZ2 problem has a regular shape Pareto front and a regular shape feasible region as shown in Figure 22 (a). The WFG4 problem also has a regular Pareto front and a regular feasible region as shown in Figure 22 (b). This explains why very similar results were obtained in Figure 4 for the DTLZ2 and WFG4 problems.



(a) $M = 3$ (DTLZ2) (b) $M = 3$ (RE3-3-1)

Fig. 21. Distributions of the cosine similarity-based difference and the objective value-based difference for all dominated solutions for the 3-objective DTLZ2 problem in (a) and the 3-objective RE3-3-1 problem in (b). A dominated solution is likely to be removed by the cosine similarity-based (objective value-based) partition method if its cosine similarity-based (objective value-based) difference is smaller than the average size of subsets.



(a) DTLZ2-4 (b) WFG4-9

Fig. 22. An illustration of the Pareto front and the feasible region of two-objective DTLZ2-4 and WFG 4-9 problems.

For the RE3-3-1 problem, each dominated solution has large cosine similarity-based difference while it has small objective similarity-based difference as shown in Figure 21 (b). This observation means that cosine similarity is less useful to group

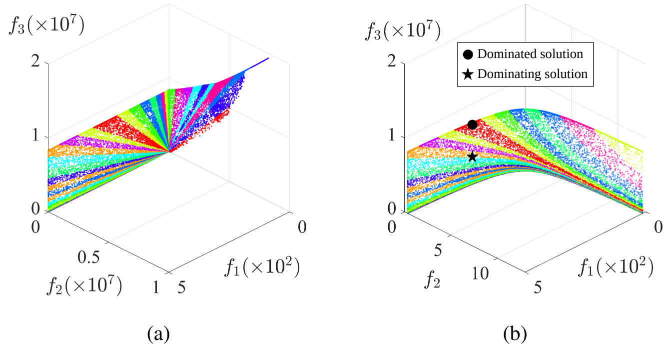


Fig. 23. (a) Plot of 50 subsets obtained by the cosine similarity-based partition method at the 20-th iteration for the RE3-3-1 problem. (b) A focused plot into the range of $[0, 13]$ of the f_2 values whereas the range of the f_2 values in (a) is $[0, 10^7]$.

the dominated solutions and their corresponding dominating solutions for the RE3-3-1 problem. In Figure 23 (a), we plot 50 subsets at the 20th iteration of the cosine similarity-based partition method for the RE3-3-1 problem. As we can see all, each objective has a totally different range. After carefully examinations of all solutions, we found that all each solution has a small f_1 value (i.e., $f_1 < 13$) or a small f_2 value (i.e., $f_2 < 13$). However, their f_3 values vary from 0 to 2×10^7. That is, all solutions are close to the $f_2 - f_3$ plane with $f_1 = 0$ or the $f_1 - f_3$ plane with $f_2 = 0$. Figure 23 (b) focuses on the subspace of the objective space with $0 \leq f_2 \leq 13$. In Figure 23 (b), we show a dominated solution and its single corresponding dominating solution. These two solutions have similar f_1 and f_2 values, but different f_3 values. Thus, this dominated solution is not grouped together with its dominating solution in the same subset, and not removed by the cosine similarity-based partition method. There are 14,770 such remaining dominated solutions in Figure 23 (b). This explains why the cosine similarity-based partition method shows poor performance whereas the objective value-based partition method works well for the RE3-3-1 problem.