MBORE: Multi-objective Bayesian Optimisation by Density-Ratio Estimation (Supplementary Material)

GEORGE DE ATH, University of Exeter, United Kingdom
TINKLE CHUGH, University of Exeter, United Kingdom
ALMA A. M. RAHAT, Swansea University, United Kingdom

CCS Concepts: • **Computing methodologies** → **Modeling methodologies**; • **Theory of computation** → **Gaussian processes**; • **Applied computing** → **Multi-criterion optimization and decision-making**.

Additional Key Words and Phrases: Bayesian optimisation, Surrogate modelling, Scalarisation methods, Efficient multi-objective optimisation, Acquisition function

ACM Reference Format:
George De Ath, Tinkle Chugh, and Alma A. M. Rahat. 2022. MBORE: Multi-objective Bayesian Optimisation by Density-Ratio Estimation (Supplementary Material) . In *Genetic and Evolutionary Computation Conference (GECCO '22), July 9–13, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 25 pages. https://doi.org/10.1145/3512290.3528769

In this supplementary materials document we provide the results for all the experiments carried out in this work. In Section A we provide additional details on the benchmarks and methods used. Section B summarises the performance of each method based on model (XGB, MLP or GP) with respect to a given benchmark and scalariser, as well as for problem dimensionality and number of objectives. Section C shows the computational timing per problem dimensionality for the WFG and DTLZ benchmarks. Finally, in Section D we provide convergence plots in terms of hypervolume and IGD+ for all test problems.

A EXPERIMENTAL SET UP

In the following sections we give the reference and ideal points for all three benchmarks used, as well as the number of weight vectors used for the augmented Tchebycheff scalarisation method.

A.1 Benchmark details

Reference points and ideal points were used normalise objective values before the hypervolume and IGD+ indicators were calculated. This ensures that objectives are weighted equally in both measures. The i-th element of a given vector of solutions $\mathbf{f} = \{f^1, \dots, f^M\}$ is normalised by calculating $\hat{f}^i = (f^i - q^i)/(r^i - q^i)$, where $\mathbf{q} = \{q^1, \dots, q^M\}$ and $\mathbf{r} = \{r^1, \dots, r^M\}$ are the ideal and reference points respectively for a given problem. The i-th element of $\mathbf{q}$ and $\mathbf{r}$ was found by optimising the i-th objective of the corresponding test problem with CMA-ES [3], either minimising it for $\mathbf{q}$ or maximising for $\mathbf{r}$. For completeness, we include the ideal and reference points used for all test problems evaluated in this work. Note that, for conciseness, in the following tables we use a python-like vector-building notation to denote vectors, e.g. $[10] * M$ is a vector of length

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
GECCO '22, July 9–13, 2022, Boston, MA, USA
© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9237-2/22/07…$15.00
https://doi.org/10.1145/3512290.3528769

M containing with each element containing the value 10, and $[10] * (M-1) + [20]$ is another M-length vector with its last value being 20 and its $M-1$ other values being 10.

Problem ID(s)	d	Reference Point	Ideal Point
1	2	$[120] * M$	$[0] * M$
	5	$[450] * M$	$[0] * M$
	10	$[1000] * M$	$[0] * M$
2, 4, 5	2	$[2] * M$	$[0] * M$
	5	$[2] * M$	$[0] * M$
	10	$[4] * M$	$[0] * M$
3	2	$[250] * M$	$[0] * M$
	5	$[1000] * M$	$[0] * M$
	10	$[2000] * M$	$[0] * M$
6	2	$[2.5] * M$	$[0] * M$
	5	$[5] * M$	$[0] * M$
	10	$[10] * M$	$[0] * M$
7	2	$[1.5] * (M-1) + [23]$	See below
	5	$[1.5] * (M-1) + [60]$	See below
	10	$[1.5] * (M-1) + [110]$	See below

Table 1. DTLZ [2] reference and ideal points for a given dimensionality d and number of objectives M.

M	Ideal point
2	$[0, 2.307]$
3	$[0, 0, 2.614]$
5	$[0, 0, 0, 0, 3.228]$
10	$[0, 0, 0, 0, 0, 0, 0, 0, 0, 4.763]$

Table 2. Ideal points for the DTLZ [2] test problem number 7 for a given number of objectives and $d \leq 10$.

M	Reference Point	Ideal Point
2	$[3, 5]$	$[0, 0]$
3	$[3, 5, 7]$	$[0, 0, 0]$
5	$[3, 5, 7, 9, 11]$	$[0, 0, 0, 0, 0]$
10	$[3, 5, 7, 9, 11, 13, 15, 17, 19, 21]$	$[0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$

Table 3. WFG [4] reference and ideal points for a given number of objectives M and $d \in \{6, 8, 10, 20, 50, 100\}$.

Problem (d, M)	Reference Point	Ideal Point
RE2-4-1 (4, 2)	[2995, 0.051]	[1237, 0.002]
RE2-2-4 (2, 2)	[6005, 45]	[60.5, 0]
RE3-3-1 (3, 3)	[817, 8250000, 19360000]	[0, 0.3, 0]
RE3-4-2 (4, 3)	[334, 17600, 425100000]	[0.01, 0.0004, 0]
RE3-5-4 (5, 3)	[1705, 11.8, 0.27]	[−0.73, 1.13, 0]
RE3-4-7 (4, 3)	[1.01, 1.25, 1.1]	[0, 0, −0.44]
RE4-7-1 (7, 4)	[43, 4.5, 13.1, 14.2]	[15.5, 3.5, 10.6, 0]
RE4-6-2 (6, 4)	[0, 20100, 31100, 15.4]	[−2757, 3962, 1947, 0]
RE6-3-1 (3, 6)	[83100, 1351, 2854000, 16028000, 358000, 99800]	[63840, 30, 285346, 183749, 7.2, 0]
RE9-7-1 (7, 9)	[42.7, 1.6, 350, 1.1, 1.57, 1.75, 1.25, 1.3, 1.06]	[15.5, 0, 0, 0.09, 0.36, 0.52, 0.73, 0.61, 0.66]

Table 4. Real-world benchmark [5] reference and ideal points.

A.2 Augmented Tchebycheff Scalarisation

We used $\rho = 0.05$ in the scalariser's calculations. The weight vectors for the scalarisation were calculated via the Riesz' s-Energy method [1]. The following table details the number of weight vectors used for a given number of objectives M.

Number of objectives	2	3	4	5	6	7	8	9	10
Number of weight vectors	100	105	120	126	132	112	156	90	275

Table 5. Number of weight vectors used for a given number of objectives.

A.3 Gaussian Process Priors

A zero-mean GP was used for BO, alongside an ARD Matérn kernel with $\nu = 5/2$:

$$\kappa_{Matern}(\mathbf{x}, \mathbf{x}' \mid \boldsymbol{\theta}) = \sigma_o^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} r\right)^{\nu} K_{\nu}\left(\sqrt{2\nu} r\right), \tag{1}$$

where K_ν is a modified Bessel function and $r^2 = \sum_{i=1}^{d} (x_i - x_i')^2 / \omega_i^2$ is the squared distance between $\mathbf{x}$ and $\mathbf{x}'$ scaled by the length-scales ω_i. We place uniform priors on the kernel's hyperparameters $\boldsymbol{\theta} = \{\omega_1, \ldots, \omega_d, \sigma_o\}$. Given that the locations with which the GPs are trained on are rescaled to the unit hypercube and that its target function values are standardised, we define them to be $\omega \sim \mathcal{U}(10^{-4}, \sqrt{d})$ and $\sigma_o \sim \mathcal{U}(10^{-4}, 10)$ respectively. This allows for interactions to occur on a sensible length-scale, i.e. the largest distance in $[0, 1]^d$ is $\sqrt{d}$, and a plausible output-scale, i.e. not too large given the standardised target function.

B OVERALL SUMMARY OF PERFORMANCE FOR ALL PROBLEM SETS

		Hypervolume			IGD+		
	Benchmark	XGB	MLP	GP	XGB	MLP	GP
HYPI	DTLZ	36	4	29	40	1	28
	WFG	37	0	36	36	1	34
	RW	4	0	6	5	0	6
	WFG (HD)	10	7	18	10	9	18
DomRank	DTLZ	21	11	37	30	5	36
	WFG	36	0	45	40	1	35
	RW	1	1	8	2	1	8
	WFG (HD)	5	15	17	7	16	10
PHC	DTLZ	36	12	22	41	8	22
	WFG	39	0	36	38	0	35
	RW	7	0	5	9	0	3
	WFG (HD)	18	6	7	17	5	6
AT	DTLZ	13	8	43	27	2	42
	WFG	1	0	63	4	2	62
	RW	2	0	8	2	0	8
	WFG (HD)	0	4	23	0	8	23

Table 6. Performance summary of MBORE (MLP and XGB) and BO (GP) for a given scalarisation method (HYPI, DomRank, PHC and AT), on each benchmark. Table values correspond to the number of times each model was the best or statistically equivalent to the best method for a given benchmark and scalarisation method. Note how the mono-surrogate-based BO (GP) performs the best across all problems for AT scalariser, whereas MBORE with XGB does for PHC.

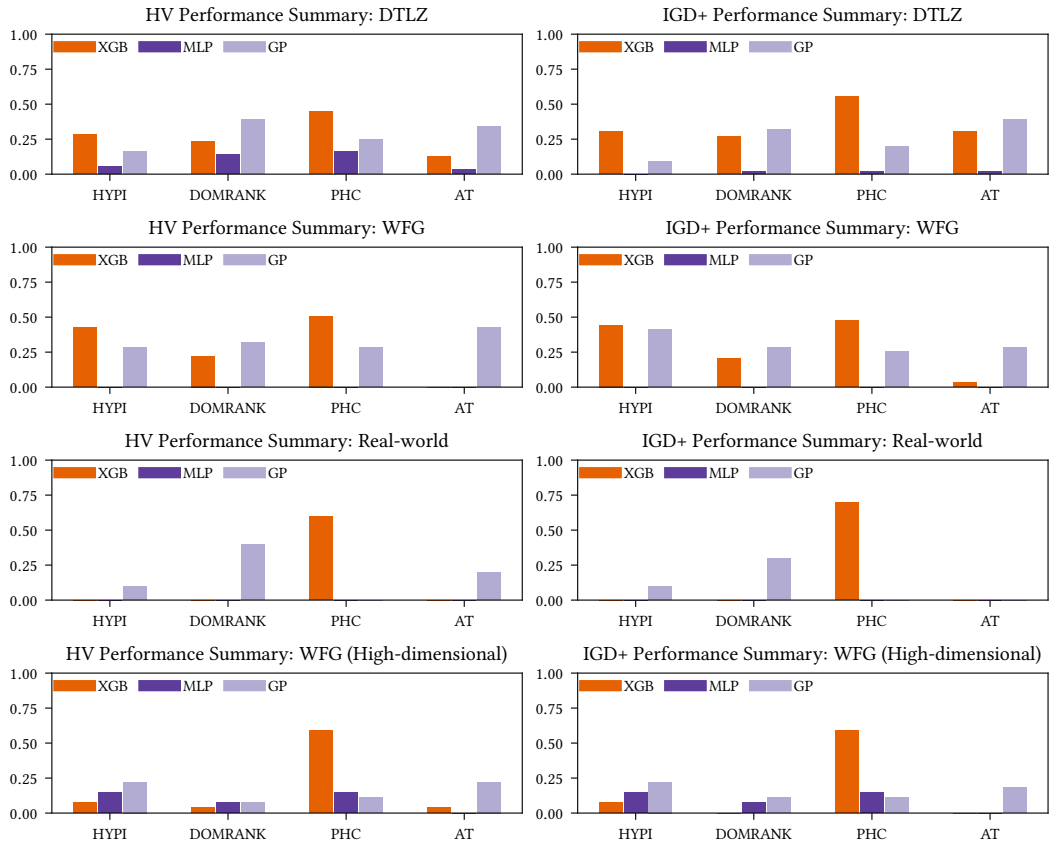

Fig. 1. Performance summary for the sets of benchmark problems evaluated: (*left column*) hypervolume (*right column*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

B.1 Performance Summary based on dimensionality/objectives: DTLZ

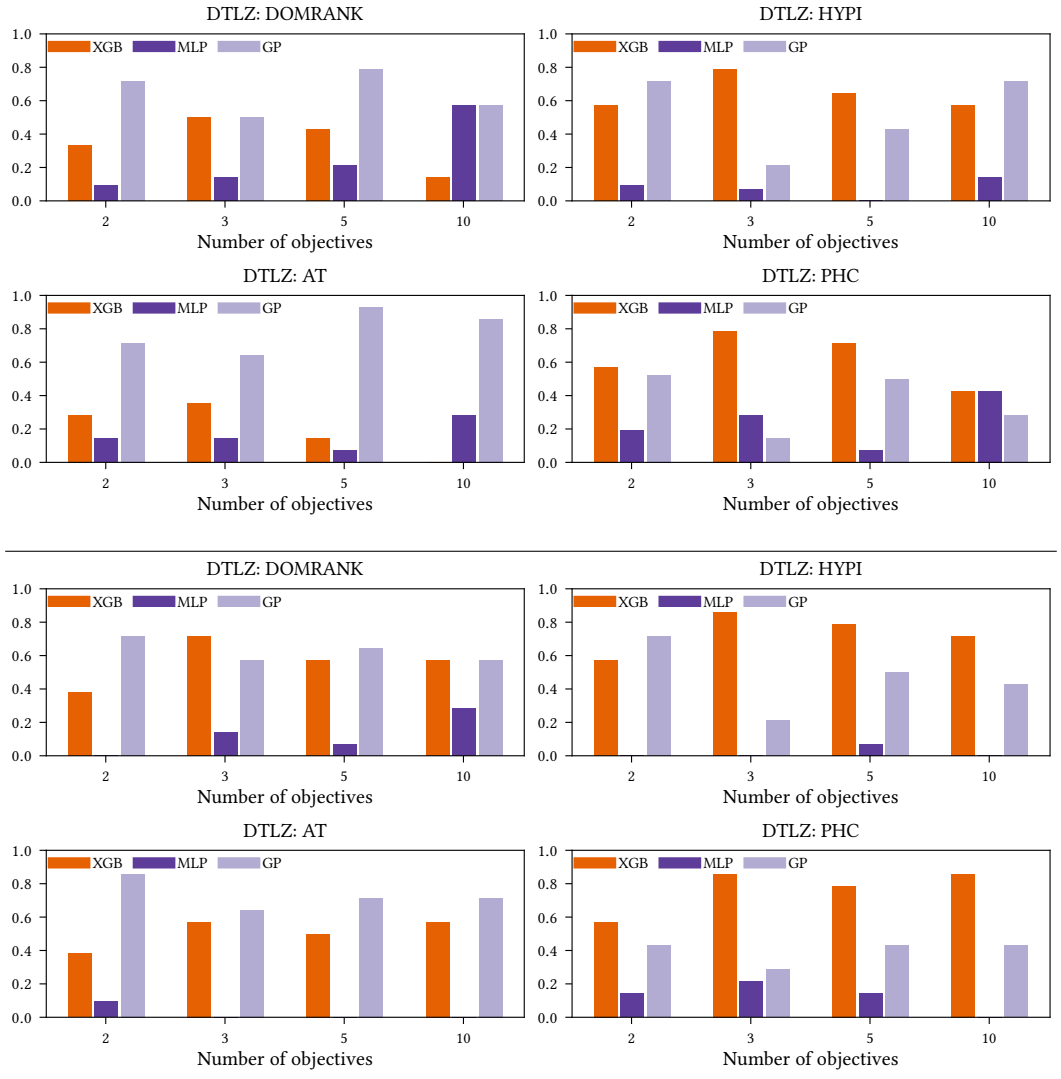

Fig. 2. Performance summary for the DTLZ test problems based on the number of objectives a problem has: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

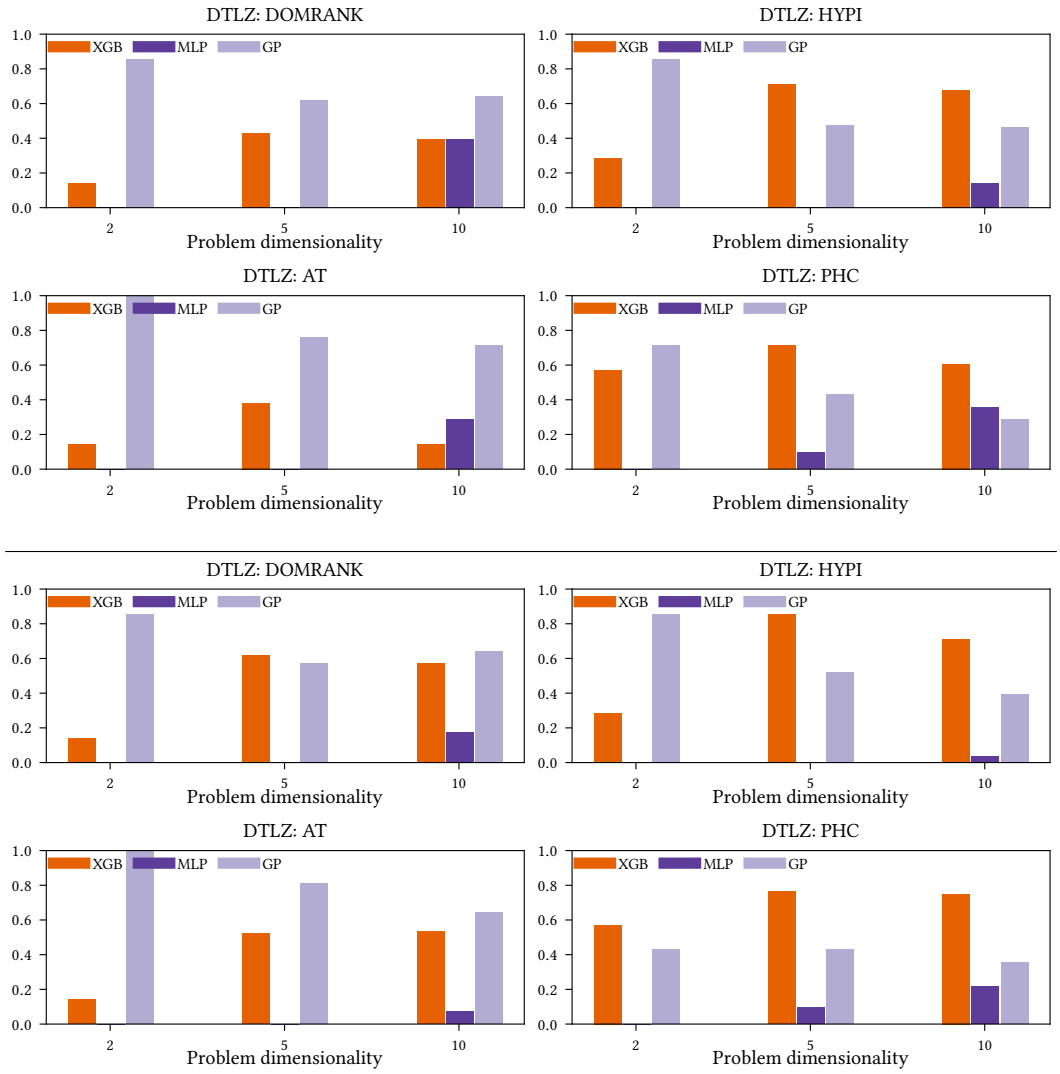

Fig. 3. Performance summary for the DTLZ test problems based on the problem's dimensionality: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

B.2 Performance Summary based on dimensionality/objectives: WFG

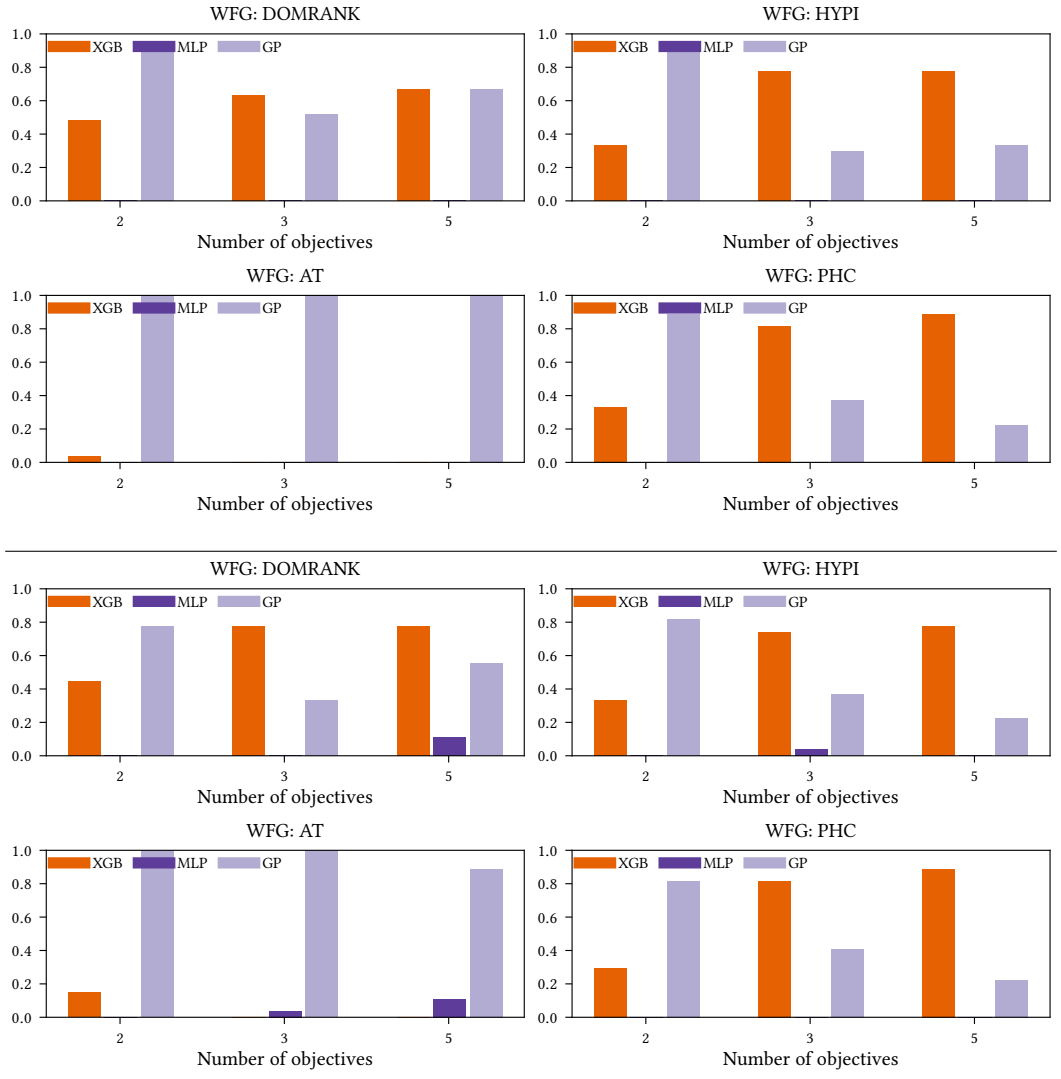

Fig. 4. Performance summary for the WFG test problems based on the number of objectives a problem has: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

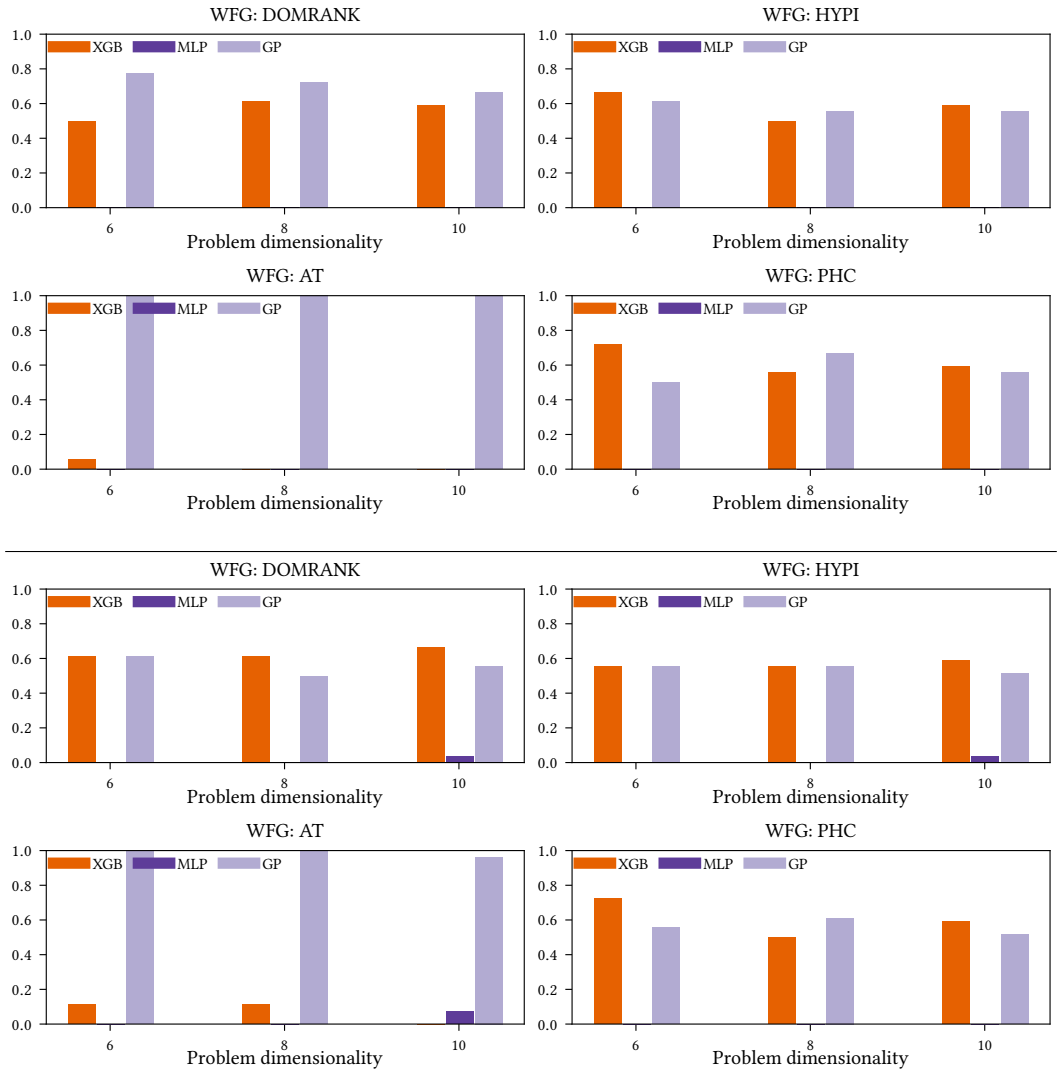

Fig. 5. Performance summary for the WFG test problems based on the problem's dimensionality: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

B.3 Performance Summary based on dimensionality/objectives: WFG (high-dimensional)

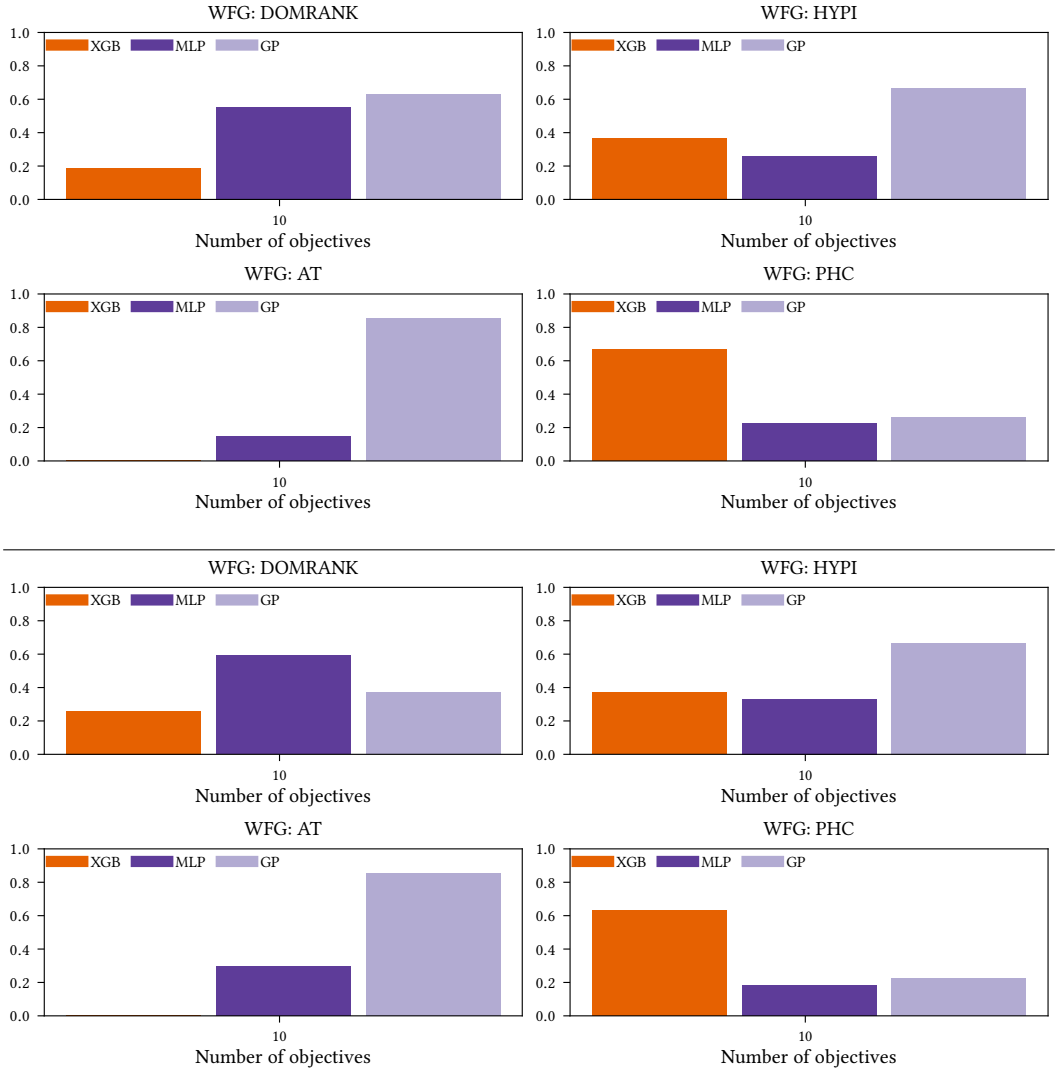

Fig. 6. Performance summary for the high-dimensional WFG test problems based on the number of objectives a problem has: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

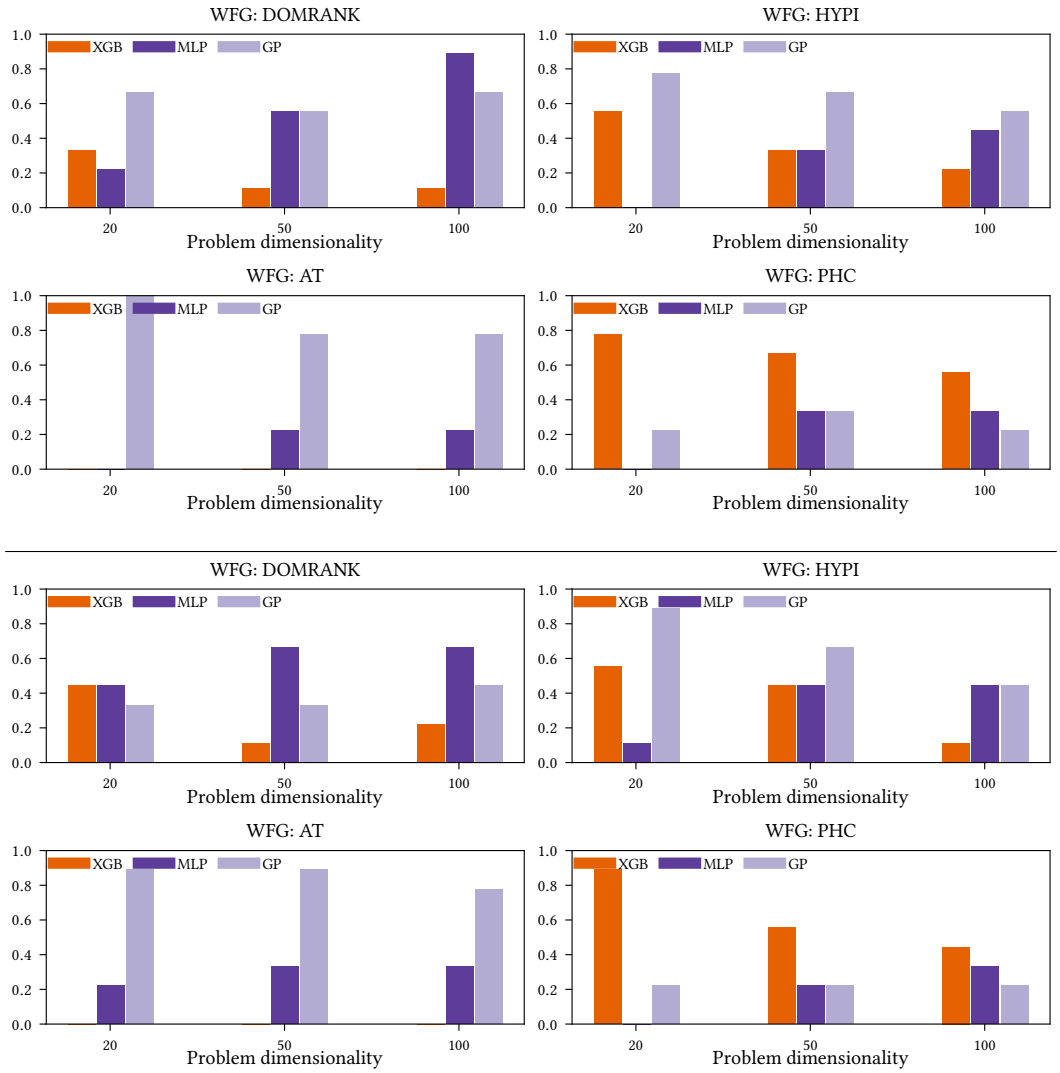

Fig. 7. Performance summary for the high-dimensional WFG test problems based on the problem's dimensionality: (*upper*) hypervolume (*lower*) IGD+. Bar heights correspond to the proportion of times that a method is best or statistically equivalent to the best method the benchmark's problems.

C COMPUTATIONAL TIMING FOR DTLZ AND WFG

The following plots show the computation time for each problem dimensionality across the DTLZ ($d \in \{2, 5, 10\}$) and WFG ($d \in \{6, 8, 10, 20, 50, 100\}$) benchmarks. Note that the timings of the real-world problems are not included because they are not configurable, i.e. only one dimensionality can be used per problem.

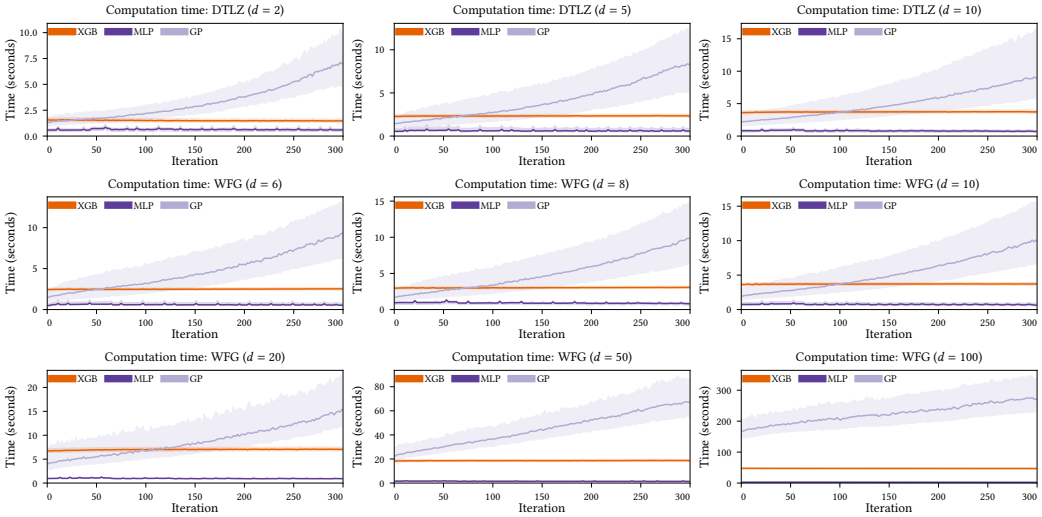

Fig. 8. Computation time taken per iteration on the DTLZ and WFG benchmark problems. The median computation time over all scalarisers and problems is shown as the solid lines, with their corresponding interquartile ranges shaded.

D CONVERGENCE PLOTS

D.1 DTLZ Convergence Plots

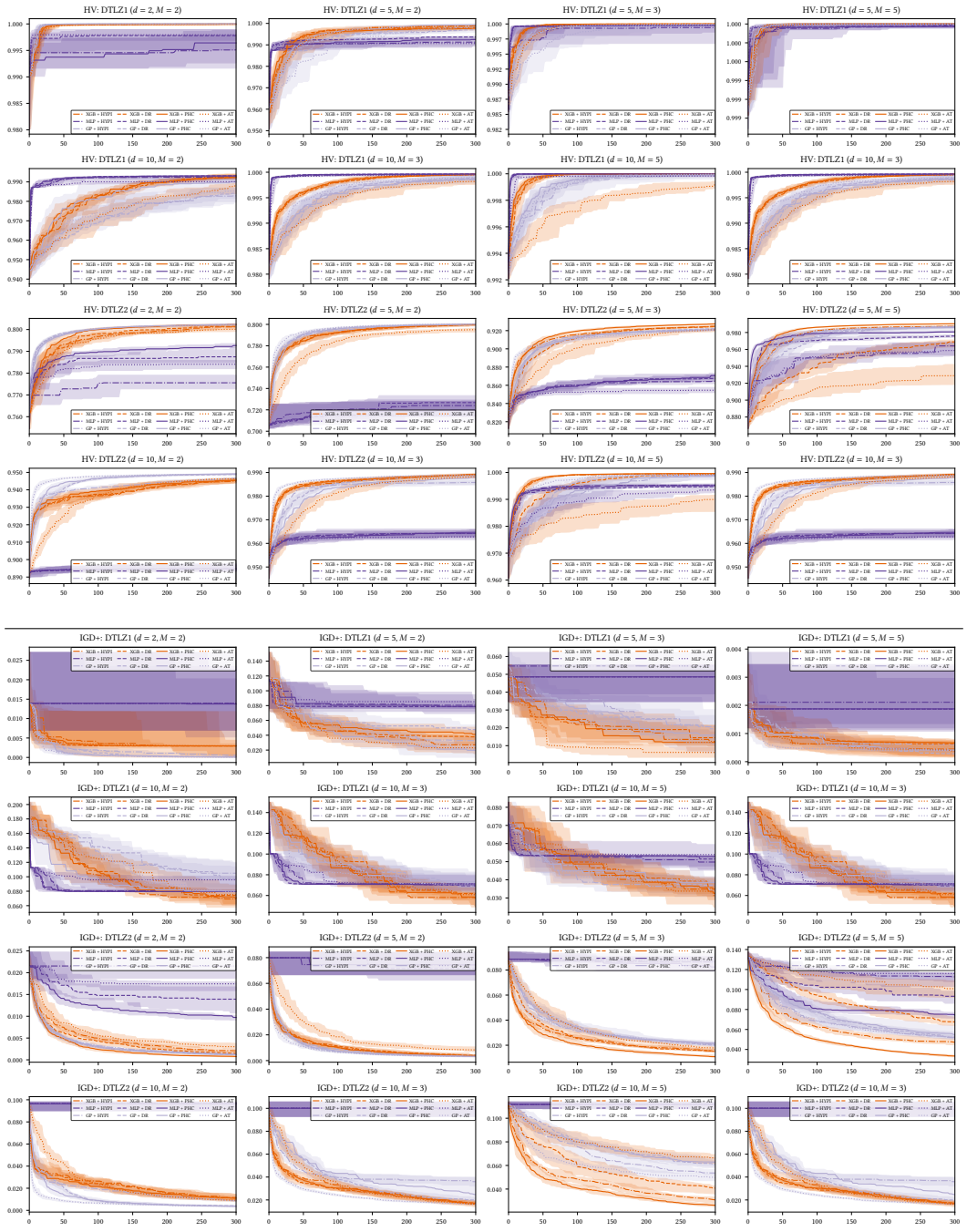

Fig. 9. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for DTLZ1 and DTLZ2.

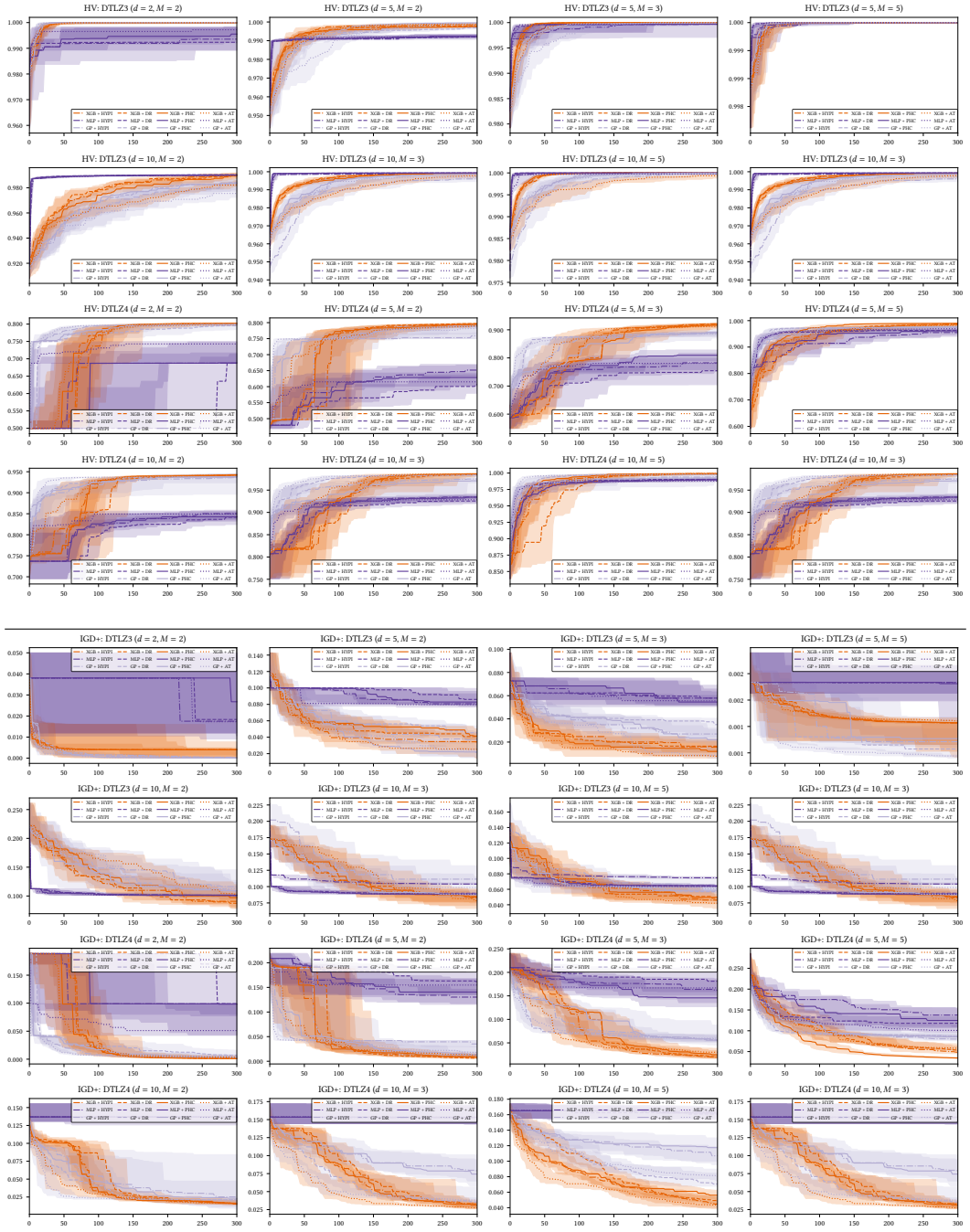

Fig. 10. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for DTLZ3 and DTLZ4.

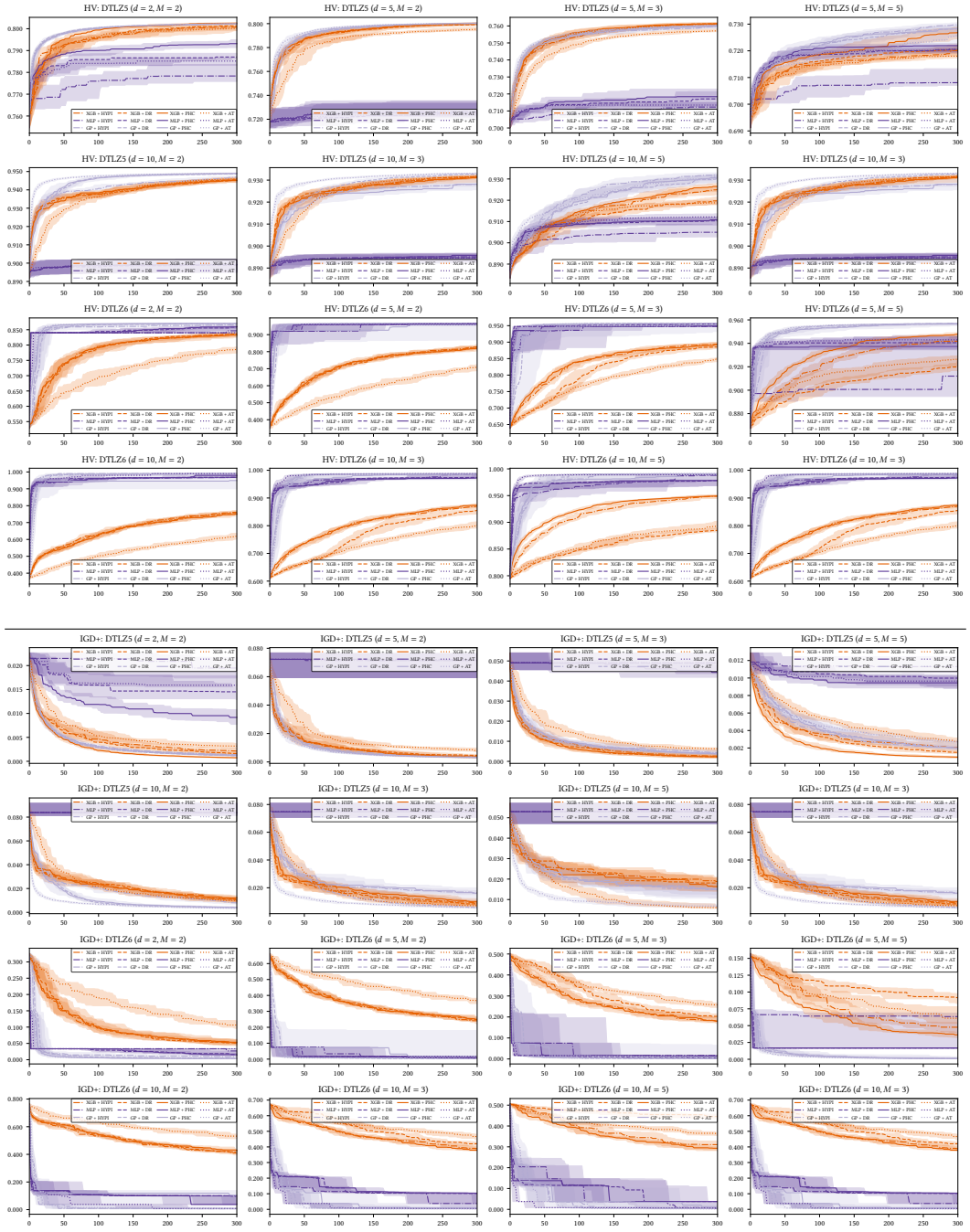

Fig. 11. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for DTLZ5 and DTLZ6.

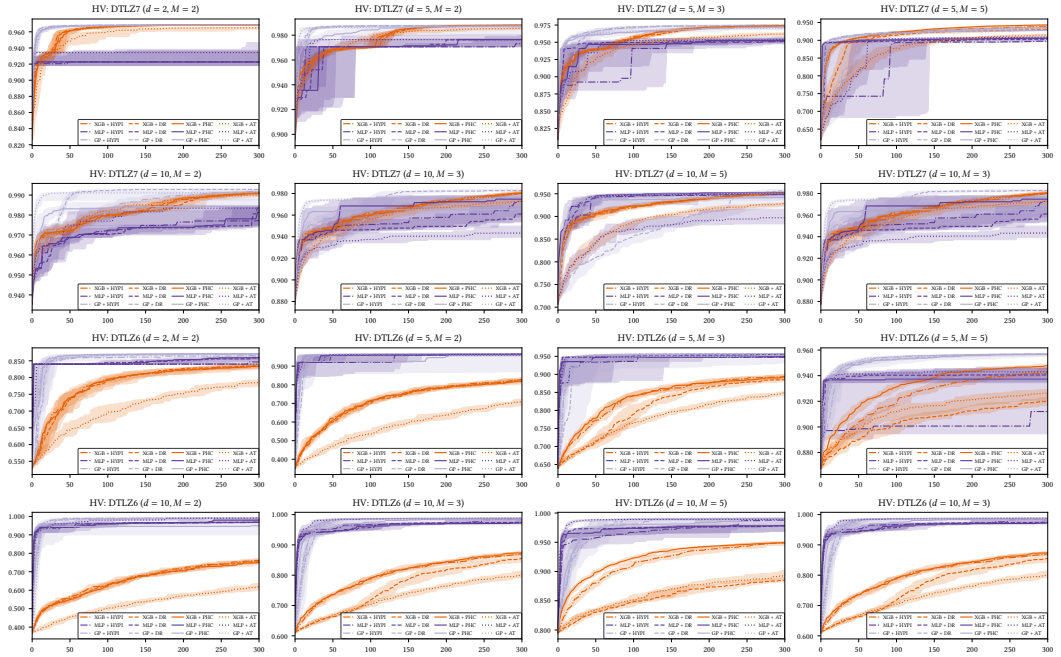

Fig. 12. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for DTLZ7.

D.2 WFG Convergence Plots

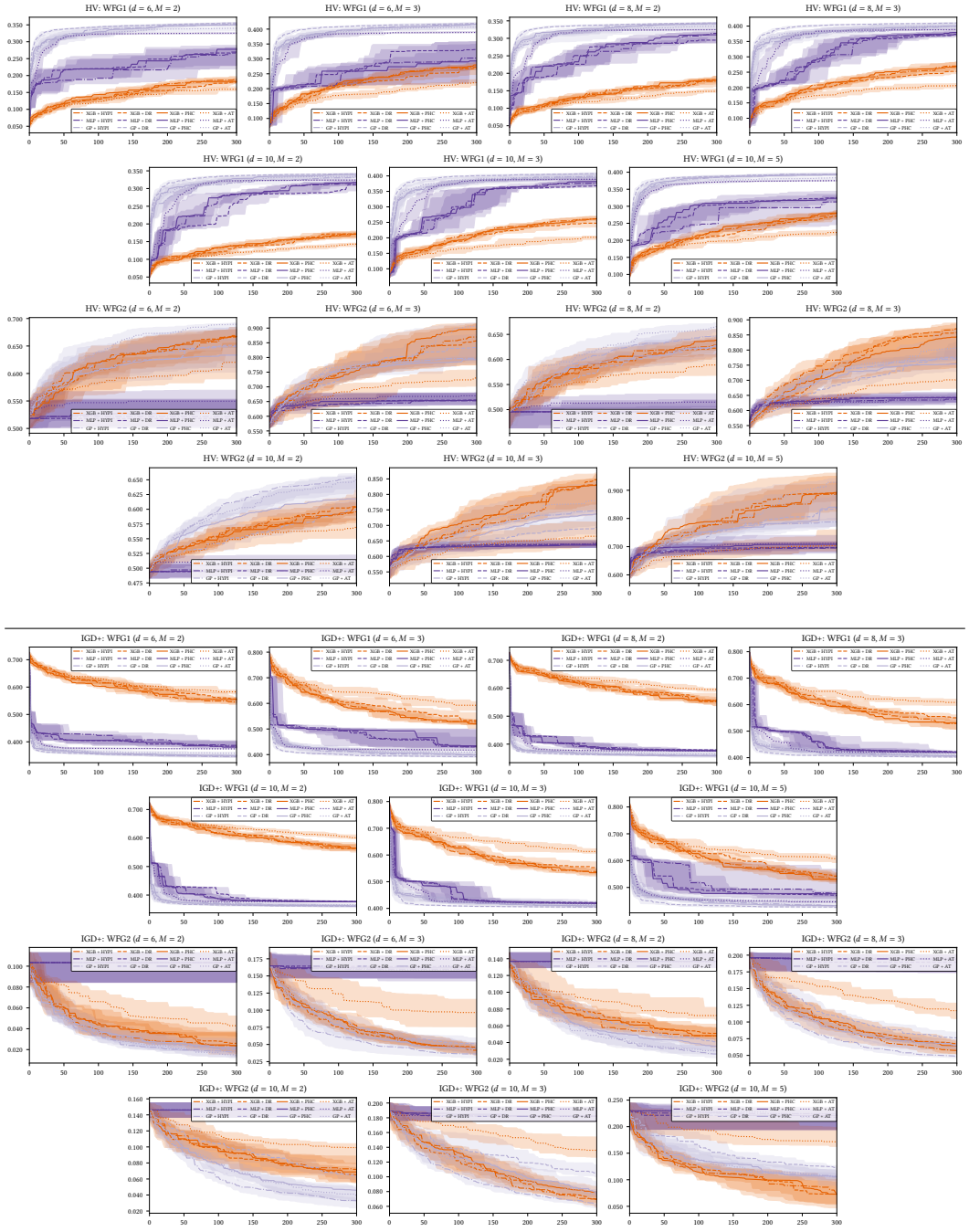

Fig. 13. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for WFG1 and WFG2.

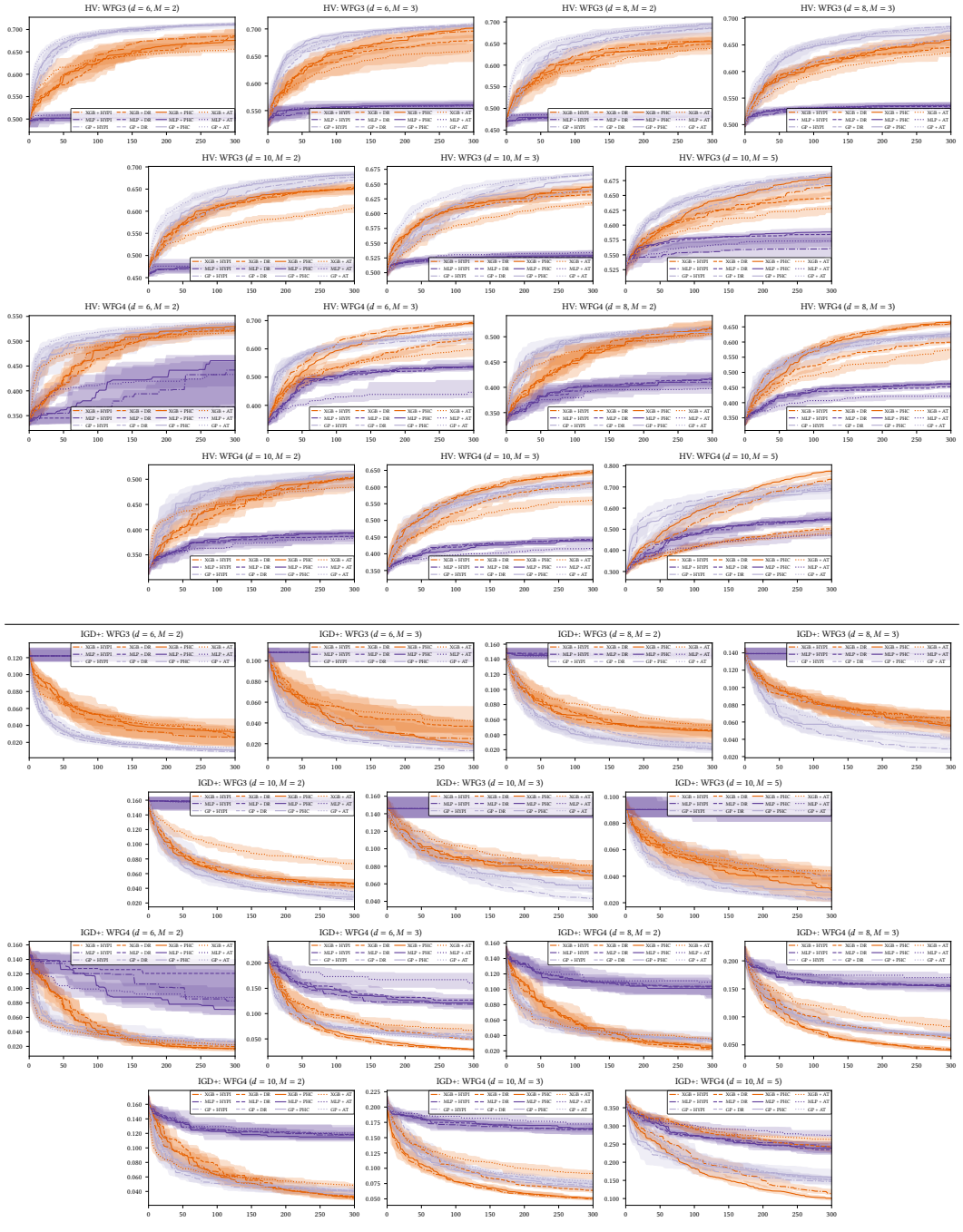

Fig. 14. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for WFG3 and WFG4.

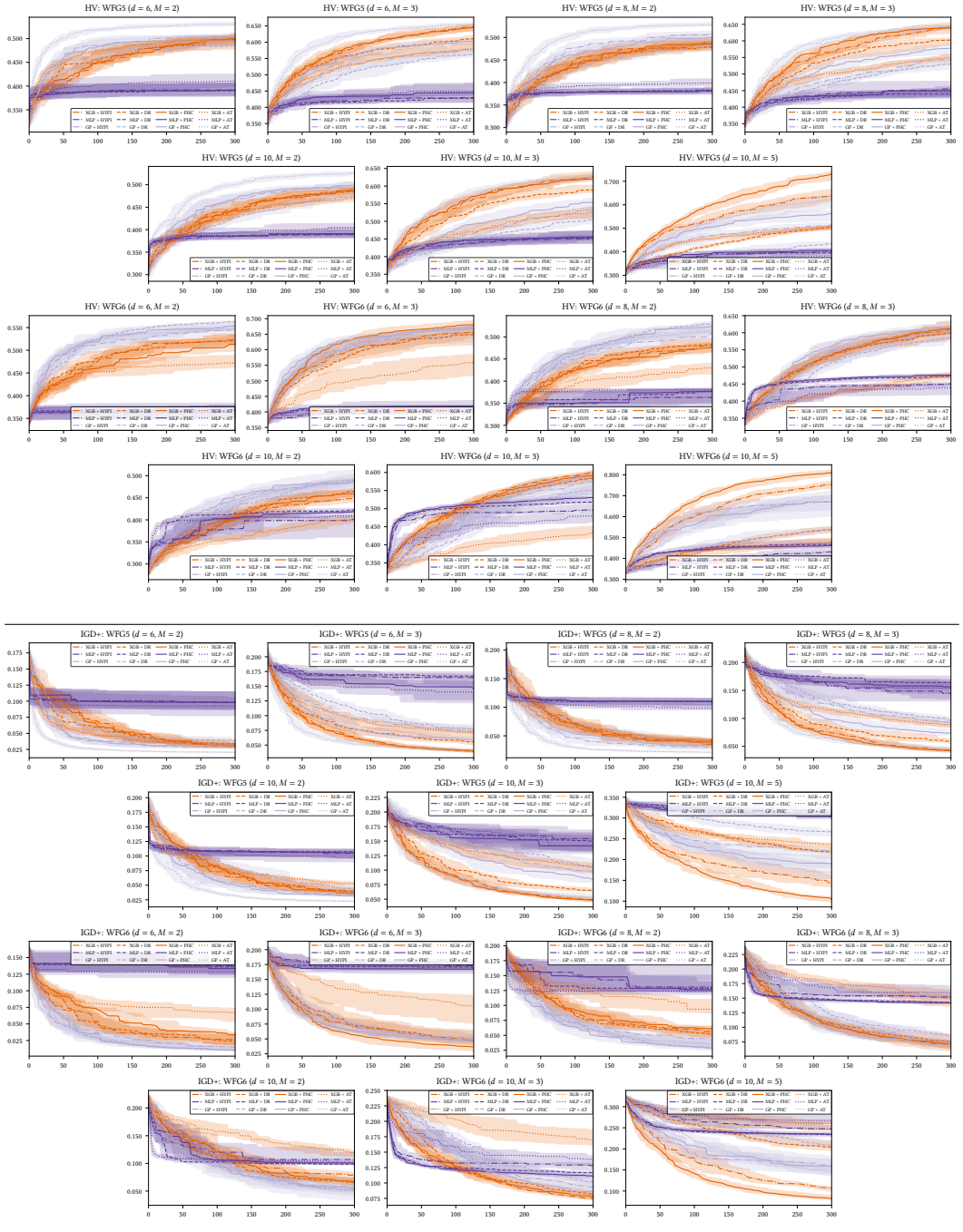

Fig. 15. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for WFG5 and WFG6.

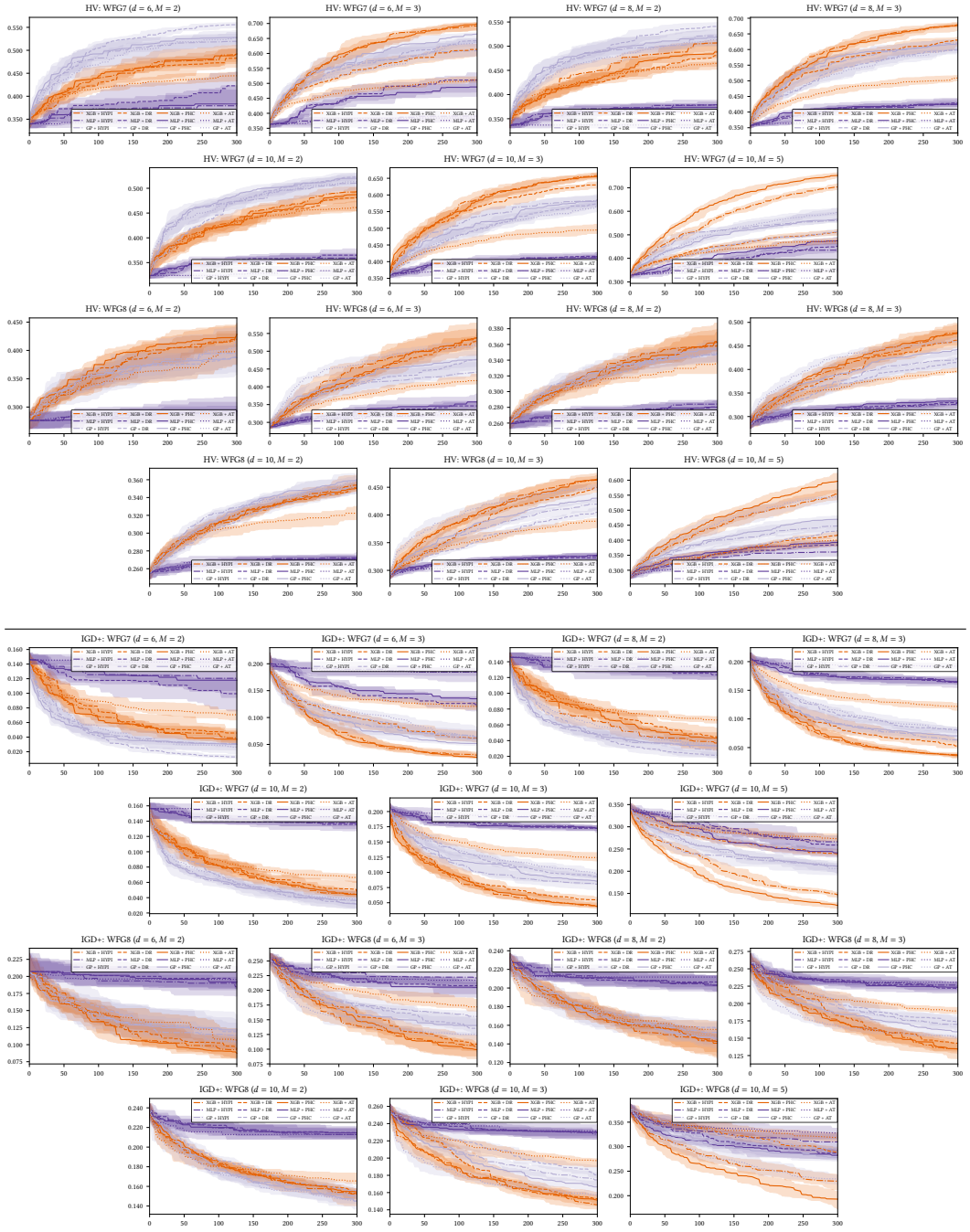

Fig. 16. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for WFG7 and WFG8.

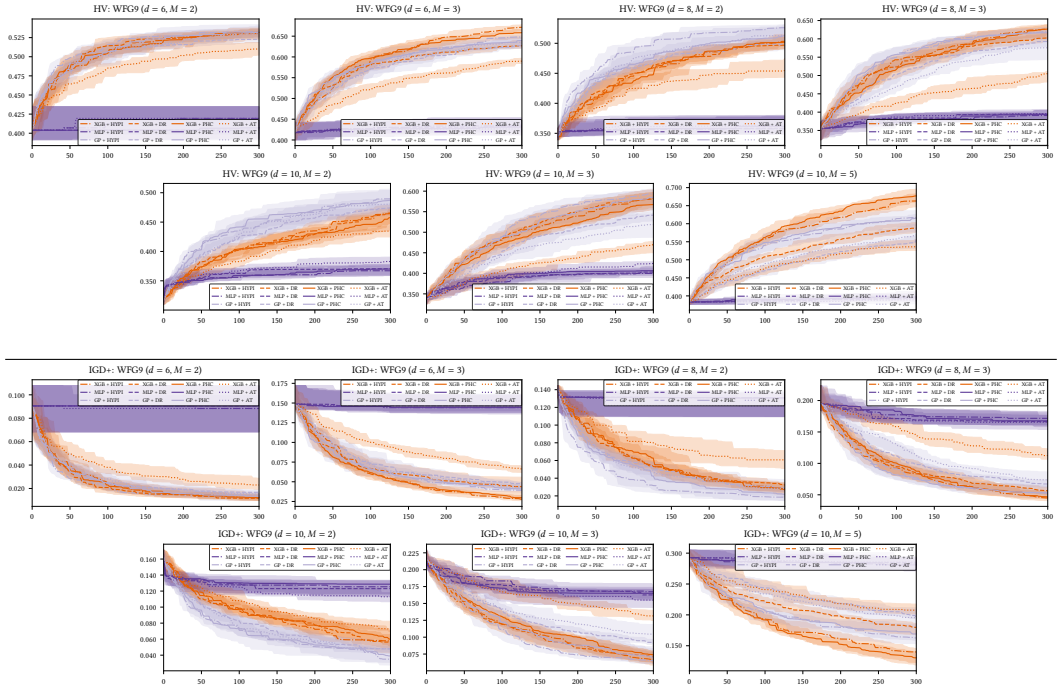

Fig. 17. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for WFG9.

D.3 Real-world Convergence Plots

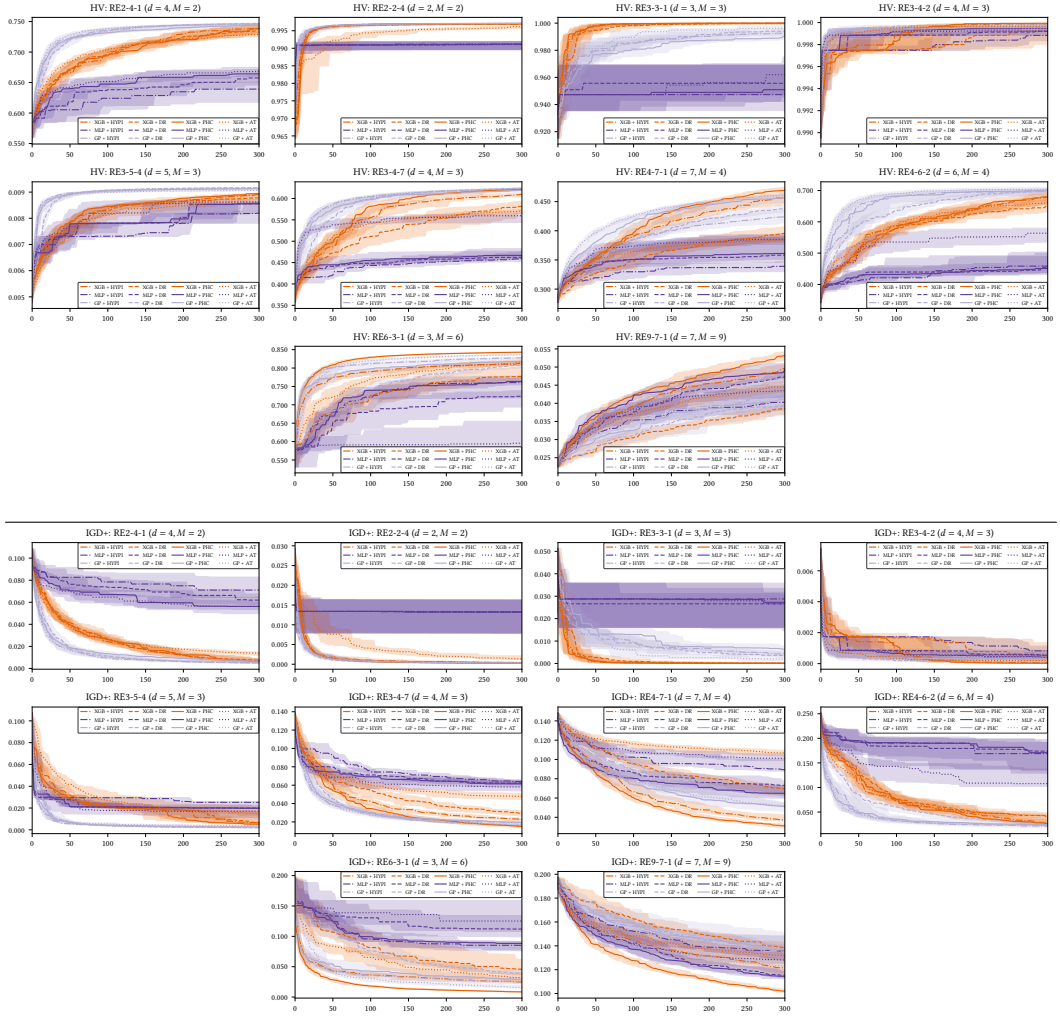

Fig. 18. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots the Real-world benchmark.

D.4 WFG (high-dimensional) Convergence Plots

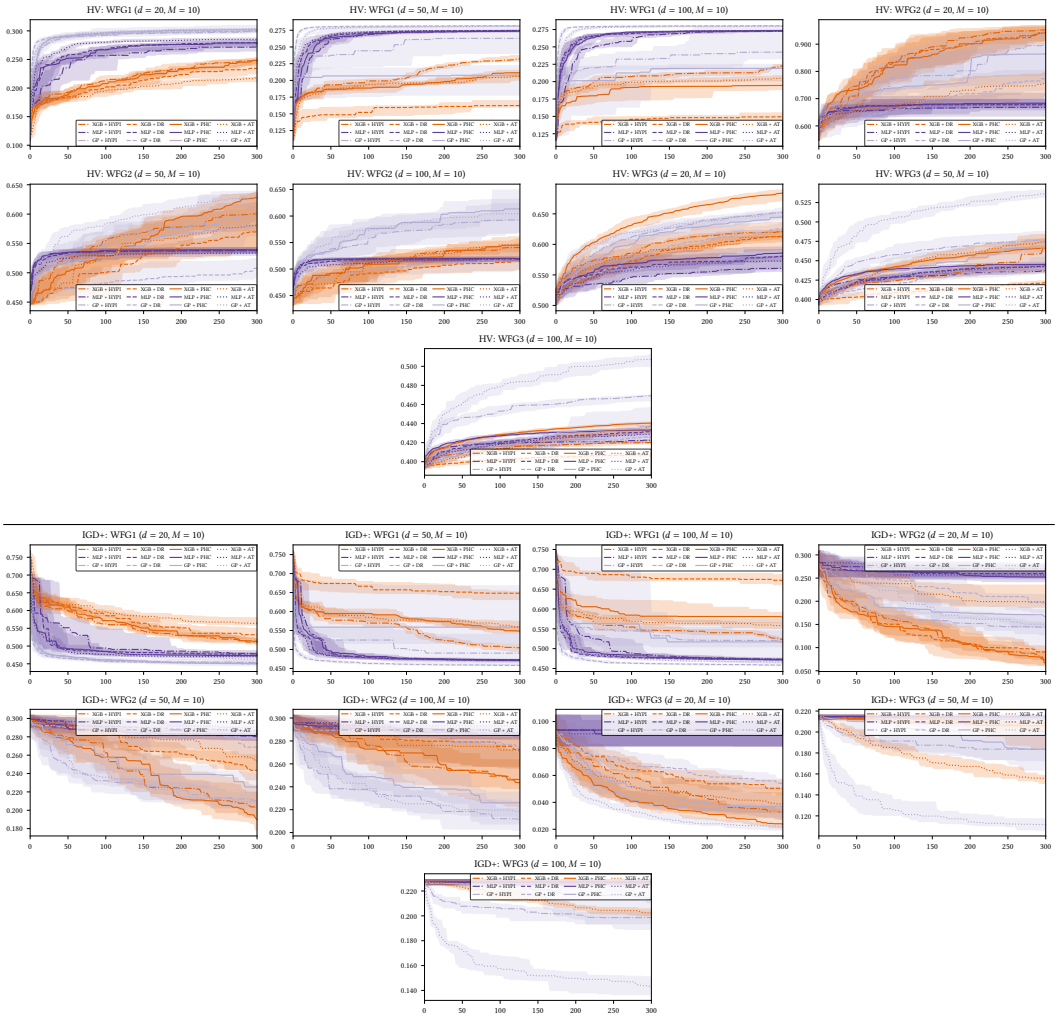

Fig. 19. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for the high-dimensional WFG1, WFG2 and WFG3 problems.

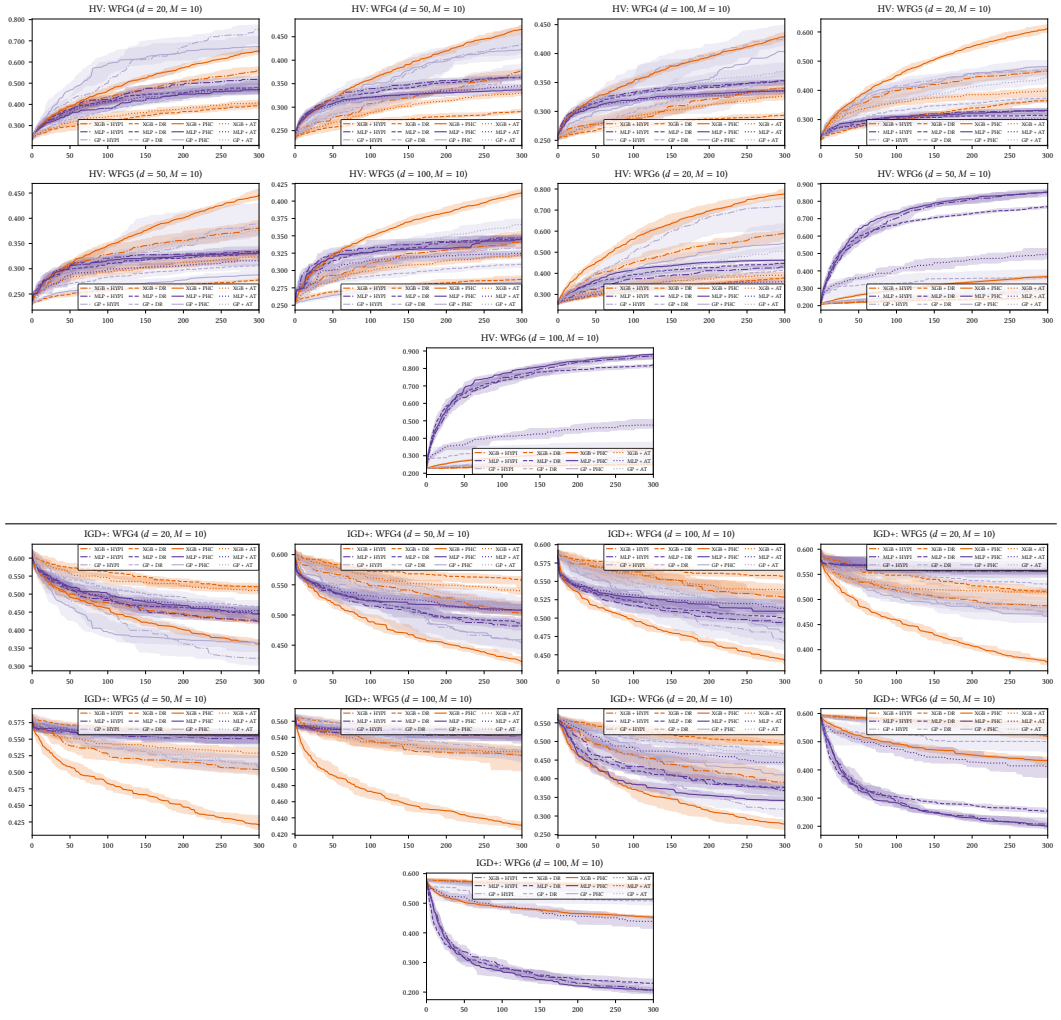

Fig. 20. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for the high-dimensional WFG4, WFG5 and WFG6 problems.

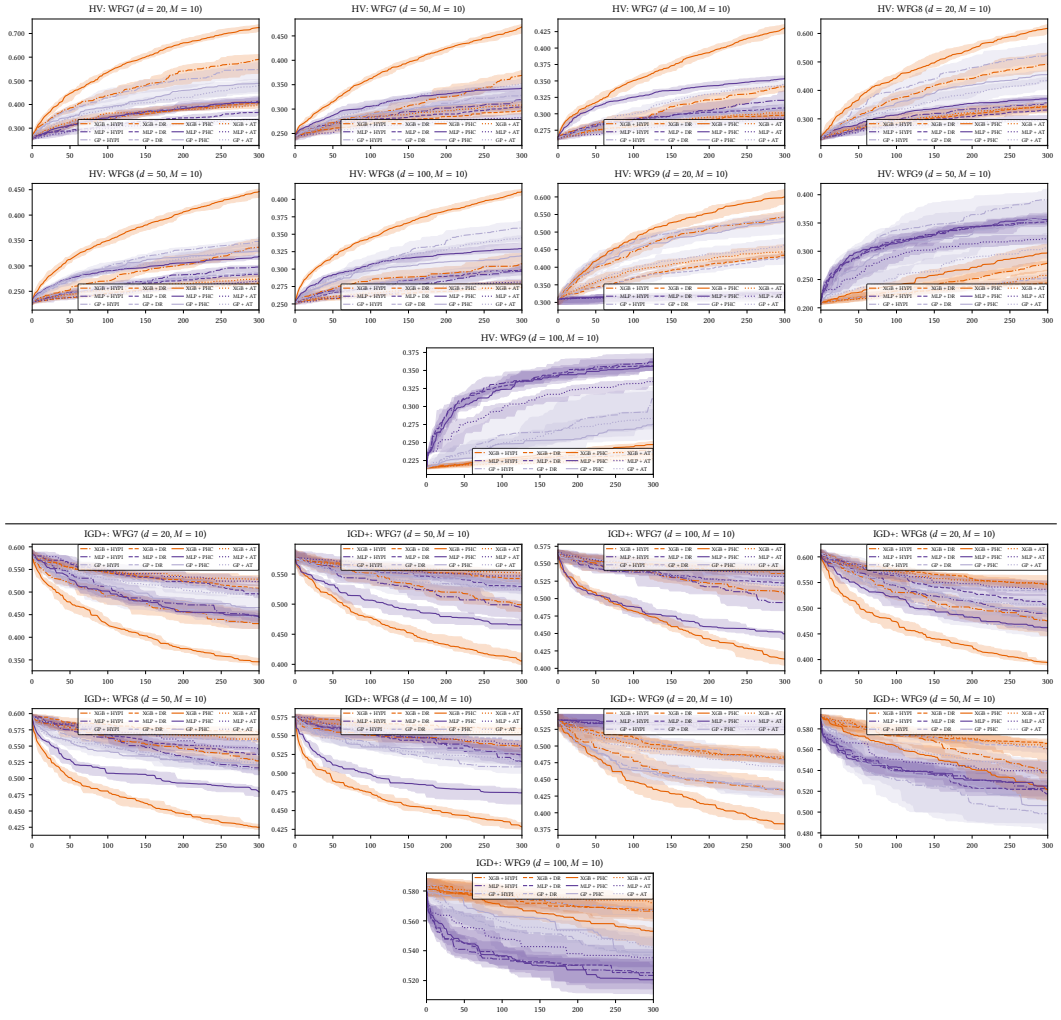

Fig. 21. Hypervolume (*upper*) and IGD+ (*lower*) convergence plots for the high-dimensional WFG7, WFG8 and WFG9 problems.

REFERENCES

[1] Julian Blank, Kalyanmoy Deb, Yashesh Dhebar, Sunith Bandaru, and Haitham Seada. 2021. Generating Well-Spaced Points on a Unit Simplex for Evolutionary Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation* 25, 1 (2021), 48–60.

[2] Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. 2005. Scalable Test Problems for Evolutionary Multiobjective Optimization. In *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications*. Springer, London, 105–145.

[3] Nikolaus Hansen. 2009. Benchmarking a BI-population CMA-ES on the BBOB-2009 function testbed. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers (GECCO '09)*. Association for Computing Machinery, New York, NY, USA, 2389–2396.

[4] S. Huband, P. Hingston, L. Barone, and L. While. 2006. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation* 10, 5 (2006), 477–506.

[5] Ryoji Tanabe and Hisao Ishibuchi. 2020. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing* 89 (2020), 106078.