7 APPENDIX

Table 2 lists the SMAPE values (mean and standard deviations over ten repetitions) for three models: the original GPR, the transferred model, and the one trained from scratch measured on each target function in 2D and 10D. We also include the results for two sizes of transfer data: $|\mathcal{T}| \in \{50, 50d\}$. Using the Kruskal-Wallis test and Dunn's posthoc procedure with a 5% significance level, we compare the SMAPE values of the three models and indicate the statistical significance as follows: the transferred model is underlined if it outperforms the original model; Among the transferred model and the one trained from scratch, the better one is indicated with boldface.

Figure 6 and 7 illustrate SMAPE scores (y-axis) of the original, transferred GPRs, and the one trained directly on the transfer data set ("Trained from scratch") as a function of the size of the transfer data set on 2D and 10D BBOB functions. Sample sizes of transfer data for 2D (x-axis) are 10, 20, 30, 50, and 100. For 10D (x-axis), they are 10, 20, 30, 50, 100, 250, 500.

Figure 8 presents experimental outcomes consistent with our earlier findings: as the number of transfer data samples increases, the performance of the transferred GPR gradually becomes inferior to that of the GPR trained from scratch. Furthermore, in most BBOB functions, the performance gap between the transferred GPR and scratch-trained GPR tends to widen once the latter starts outperforming the former.

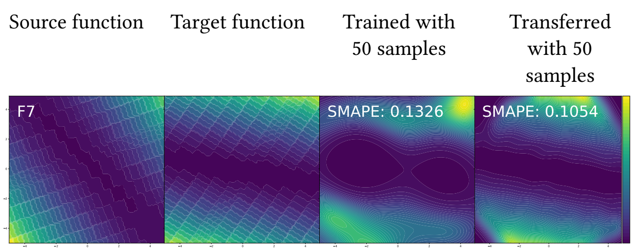

Figure 5: A case study on the transfer learning result, showing the target, trained from scratch and transferred versions of F7 (StepEllipsoid) function in dimensionality 2.

2D	Original GPR	Train from scratch	Transferred	Train from scratch	Transferred
		50 samples		100 samples	
F1	0.2973 ± 0.0770	0.0431 ± 0.0090	**0.0099 ± 0.0105**	0.0180 ± 0.0041	0.0055 ± 0.0115
F2	0.1172 ± 0.0310	0.0633 ± 0.0095	0.0353 ± 0.0186	0.0542 ± 0.0073	0.0360 ± 0.0187
F3	0.2494 ± 0.0798	0.0499 ± 0.0132	**0.0368 ± 0.0041**	0.0410 ± 0.0067	0.0356 ± 0.0030
F4	0.2286 ± 0.0777	0.0448 ± 0.0097	0.0454 ± 0.0331	0.0350 ± 0.0030	0.0343 ± 0.0109
F5	0.2455 ± 0.0788	0.0126 ± 0.0072	0.0348 ± 0.0258	0.0101 ± 0.0051	0.0350 ± 0.0255
F6	0.3778 ± 0.0950	0.1046 ± 0.0212	0.0623 ± 0.0319	0.0755 ± 0.0159	0.0583 ± 0.0303
F7	0.3623 ± 0.1175	0.1521 ± 0.0239	**0.0739 ± 0.0198**	0.1080 ± 0.0115	0.0739 ± 0.0195
F8	0.3166 ± 0.0703	0.1400 ± 0.0136	0.0850 ± 0.0460	0.1057 ± 0.0136	0.0882 ± 0.0545
F9	0.2830 ± 0.0728	0.1215 ± 0.0096	0.0895 ± 0.0504	0.1083 ± 0.0099	0.0863 ± 0.0509
F10	0.1676 ± 0.0428	0.0615 ± 0.0102	0.0362 ± 0.0163	0.0509 ± 0.0064	0.0373 ± 0.0194
F11	0.1824 ± 0.0571	0.0486 ± 0.0056	0.0340 ± 0.0195	0.0414 ± 0.0070	0.0328 ± 0.0211
F12	0.3180 ± 0.1046	0.0494 ± 0.0085	0.0578 ± 0.0321	0.0426 ± 0.0125	0.0565 ± 0.0308
F13	0.1754 ± 0.0411	0.0676 ± 0.0109	0.0377 ± 0.0249	0.0500 ± 0.0060	0.0402 ± 0.0302
F14	0.5325 ± 0.1454	0.0714 ± 0.0149	0.0287 ± 0.0262	0.0498 ± 0.0060	0.0280 ± 0.0247
F15	0.3241 ± 0.0990	0.0508 ± 0.0126	0.0481 ± 0.0537	0.0388 ± 0.0039	0.0303 ± 0.0039
F16	0.2171 ± 0.0255	0.1648 ± 0.0090	**0.1440 ± 0.0082**	0.1560 ± 0.0033	0.1433 ± 0.0080
F17	0.5335 ± 0.1625	0.1402 ± 0.0123	**0.1187 ± 0.0049**	0.1333 ± 0.0084	**0.1184 ± 0.0046**
F18	0.4117 ± 0.1253	0.1045 ± 0.0083	0.0883 ± 0.0054	0.1013 ± 0.0093	0.0881 ± 0.0050
F19	0.3465 ± 0.0795	0.1926 ± 0.0290	0.1779 ± 0.0114	0.1812 ± 0.0090	0.1735 ± 0.0044
F20	0.4315 ± 0.1291	0.1299 ± 0.0174	**0.0623 ± 0.0304**	0.0856 ± 0.0070	0.0681 ± 0.0413
F21	0.3317 ± 0.0294	0.2489 ± 0.0128	0.1885 ± 0.1028	0.2364 ± 0.0131	0.1805 ± 0.0941
F22	0.3432 ± 0.0386	0.2309 ± 0.0245	0.1621 ± 0.0788	0.2080 ± 0.0223	0.1573 ± 0.0810
F23	0.1527 ± 0.0043	0.1577 ± 0.0099	0.1527 ± 0.0043	0.1548 ± 0.0062	0.1526 ± 0.0042
F24	0.1876 ± 0.0411	0.1399 ± 0.0228	0.1458 ± 0.0354	0.1183 ± 0.0144	0.1405 ± 0.0374
10D	Original GPR	Train from scratch	Transferred	Train from scratch	Transferred
		50 samples		500 samples	
F1	0.0700 ± 0.0119	1.4140 ± 0.8951	**0.0043 ± 0.0013**	0.0079 ± 0.0006	**0.0033 ± 0.0013**
F2	0.0809 ± 0.0132	1.2238 ± 0.9507	**0.0276 ± 0.0059**	0.0252 ± 0.0048	0.0221 ± 0.0071
F3	0.0648 ± 0.0026	1.4149 ± 0.8936	**0.0370 ± 0.0066**	0.0154 ± 0.0022	0.0137 ± 0.0015
F4	0.1188 ± 0.0048	1.4282 ± 0.8734	**0.0915 ± 0.0041**	0.0466 ± 0.0038	0.0554 ± 0.0037
F5	0.0481 ± 0.0083	1.4013 ± 0.9144	**0.0034 ± 0.0008**	0.0025 ± 0.0003	0.0030 ± 0.0003
F6	0.0638 ± 0.0064	1.2098 ± 0.9678	**0.0202 ± 0.0028**	0.0140 ± 0.0022	0.0137 ± 0.0015
F7	0.1072 ± 0.0104	1.4140 ± 0.8950	**0.0294 ± 0.0048**	0.0195 ± 0.0029	**0.0113 ± 0.0012**
F8	0.0653 ± 0.0074	1.2178 ± 0.9579	**0.0360 ± 0.0027**	0.0200 ± 0.0020	0.0182 ± 0.0012
F9	0.0435 ± 0.0013	1.2220 ± 0.9529	**0.0359 ± 0.0034**	0.0285 ± 0.0050	**0.0166 ± 0.0008**
F10	0.0759 ± 0.0138	1.0156 ± 0.9844	**0.0153 ± 0.0029**	0.0187 ± 0.0049	0.0123 ± 0.0037
F11	0.1429 ± 0.0118	1.2462 ± 0.9232	**0.0766 ± 0.0033**	0.0913 ± 0.0116	**0.0734 ± 0.0025**
F12	0.0659 ± 0.0057	1.0222 ± 0.9778	**0.0211 ± 0.0038**	0.0117 ± 0.0010	0.0106 ± 0.0021
F13	0.0357 ± 0.0052	1.4045 ± 0.9095	**0.0119 ± 0.0015**	0.0201 ± 0.0488	0.0036 ± 0.0006
F14	0.1921 ± 0.0246	1.4246 ± 0.8789	**0.0574 ± 0.0087**	0.0342 ± 0.0031	0.0240 ± 0.0054
F15	0.0831 ± 0.0082	1.4205 ± 0.8853	**0.0418 ± 0.0038**	0.0156 ± 0.0017	0.0175 ± 0.0054
F16	0.0745 ± 0.0135	1.4214 ± 0.8838	**0.0748 ± 0.0141**	0.0700 ± 0.0007	0.0748 ± 0.0142
F17	0.2301 ± 0.0267	1.4332 ± 0.8657	**0.0977 ± 0.0163**	0.0585 ± 0.0021	0.0664 ± 0.0138
F18	0.1841 ± 0.0231	1.4261 ± 0.8766	**0.0654 ± 0.0115**	0.0403 ± 0.0010	**0.0370 ± 0.0014**
F19	0.0986 ± 0.0031	1.6226 ± 0.7547	**0.0850 ± 0.0070**	0.0527 ± 0.0010	**0.0337 ± 0.0013**
F20	0.0730 ± 0.0088	1.4112 ± 0.8994	**0.0290 ± 0.0041**	0.0160 ± 0.0025	0.0135 ± 0.0012
F21	0.0156 ± 0.0004	1.4041 ± 0.9103	0.0155 ± 0.0005	**0.0122 ± 0.0007**	0.0148 ± 0.0004
F22	0.0118 ± 0.0005	1.4036 ± 0.9110	**0.0116 ± 0.0005**	0.0090 ± 0.0024	0.0105 ± 0.0005
F23	0.1223 ± 0.0009	1.6241 ± 0.7517	0.1223 ± 0.0009	**0.1183 ± 0.0011**	0.1222 ± 0.0009
F24	0.2718 ± 0.0059	1.4718 ± 0.8067	**0.2393 ± 0.0038**	0.1787 ± 0.0194	0.2333 ± 0.0031

Table 2: On 2D and 10D BBOB functions, the mean and standard deviation (over ten repetitions) of the SMAPE value measured for the original and the transferred models, and the one trained with the transfer data set ("trained from scratch"). We include results obtained with two sample sizes of the transfer data: $|\mathcal{T}| \in \{50, 50d\}$. Using the Kruskal-Wallis test and Dunn's posthoc procedure at a 5% significance level, we compare the three models and indicate the statistical significance as follows: the transferred model is underlined if it significantly outperforms the original model, while among the transferred model and the one trained from scratch, the significantly better one is indicated with boldface.

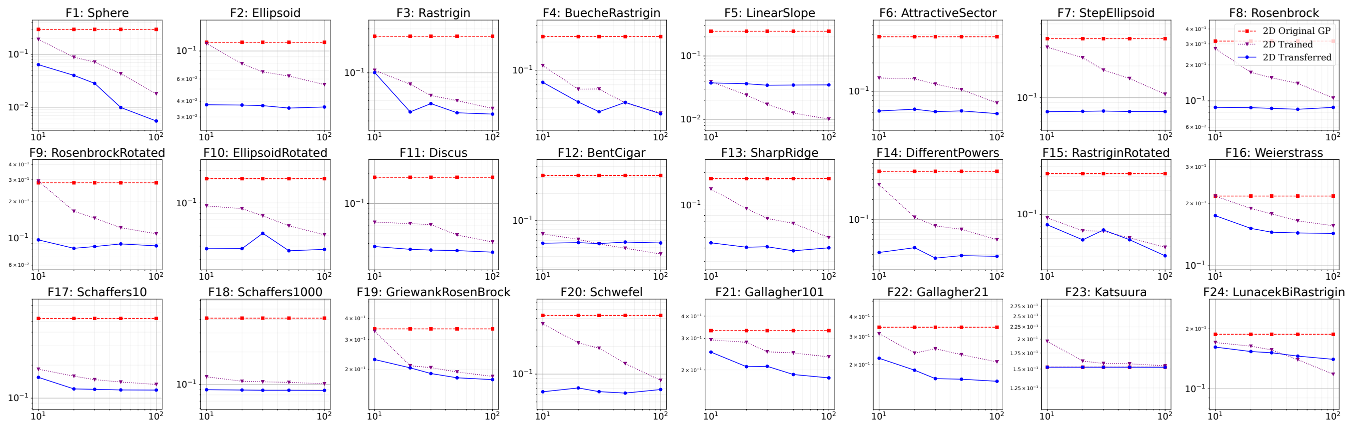

Figure 6: SMAPE scores (y-axis) of the original, transferred GPRs, and the one trained directly on the transfer data set ("Trained from scratch") as a function of the size of the transfer data set on 2D BBOB functions. Sample sizes of transfer data for 2D (x-axis) are 10, 20, 30, 50, and 100.

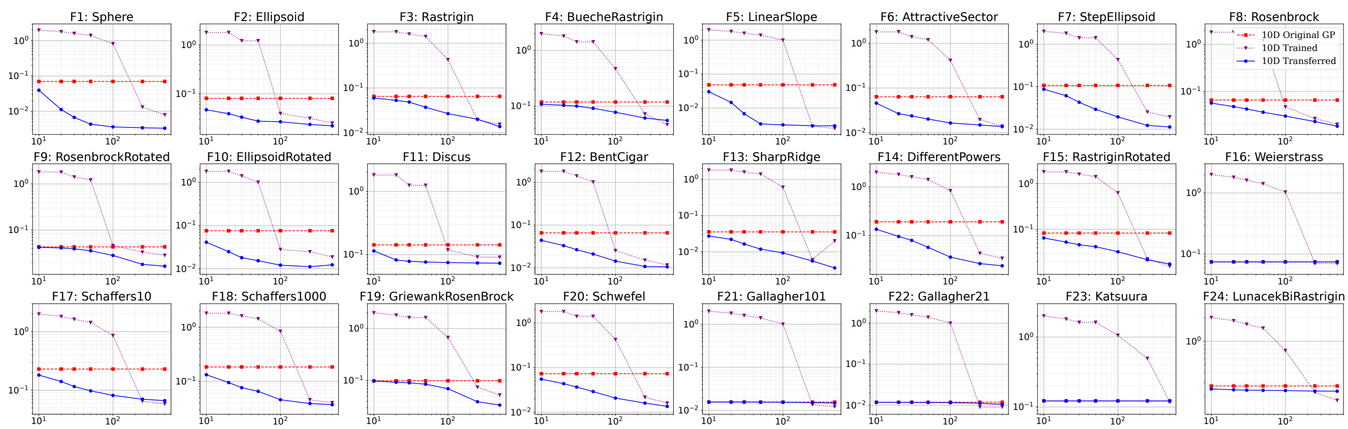

Figure 7: SMAPE scores (y-axis) of the original, transferred GPRs, and the one trained directly on the transfer data set ("Trained from scratch") as a function of the size of the transfer data set on 10D BBOB functions. Sample sizes of transfer data for 10D (x-axis) are 10, 20, 30, 50, 100, 250, and 500.

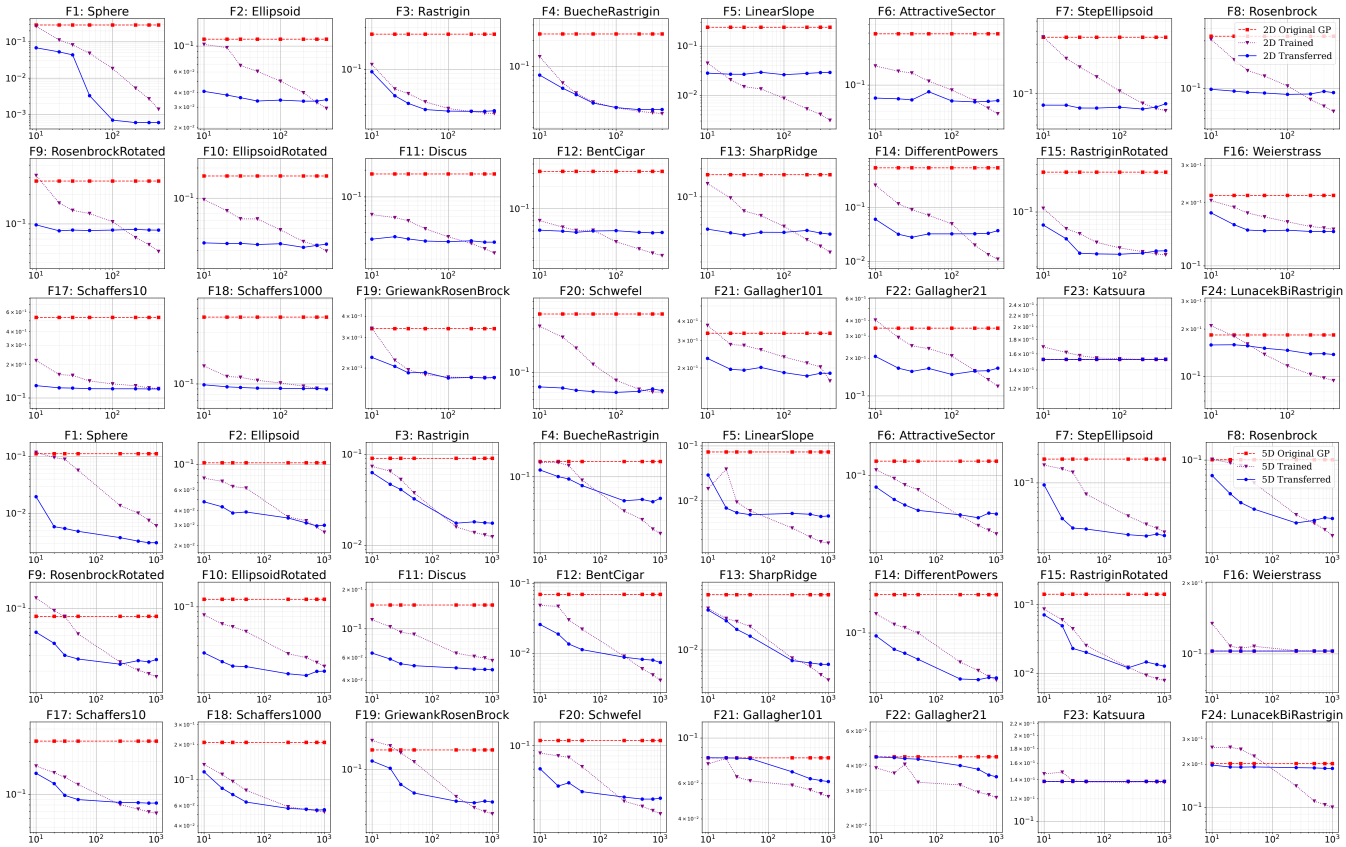

Figure 8: SMAPE scores (y-axis) of the original, transferred GPRs, and the one trained directly on the transfer data set ("Trained from scratch") as a function of the size of the transfer data set on 2D and 5D BBOB functions based on relatively larger sample sizes. Sample sizes of transfer data for both 2D and 5D are 10, 20, 30, 50, $50d$, $100d$, $150d$, $200d$.