Author: Kruisselbrink, Johannes Willem
Title: Evolution strategies for robust optimization
Date: 2012-05-10

Chapter 8

Empirical Study on Finding Robust Optima

In Chapter 7, the problem of finding robust optima in the context of Evolution Strategies has been discussed and an overview has been given of techniques for solving such problems. In this review, the myopic approach, the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach, the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ approach, the adaptive averaging approach, the archive based approach, and the metamodeling approach are shown to be feasible approaches for practical scenarios. An interesting and yet unanswered question is: how do these different techniques compare against each other with respect to performance? This chapter presents the results of an empirical study that is designed to shed some light on this question.

The structure of this chapter is as follows: Section 8.1 describes the general experimental setup adopted in this empirical study. Section 8.2 shows the results of an empirical study for finding the optimal sample sizes for the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ and the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ approach. Section 8.3 shows the results of a full empirical comparison of the different methods for finding robust optima. Section 8.4 closes with a summary and discussion.

8.1 Experimental Setup

The purpose of the experimental study is to find out how the different evaluation techniques for finding robust optima compare when used within the same algorithmic basis, namely the $(5/2_{DI}, 35)$-σSA-ES and the CMA-ES. The general experimental settings, shown in Table 8.1, restrict to one particular search space dimension size, $n = 10$, and an evaluation budget of $10,000$ function evaluations, which is taken as a standard setup throughout this chapter. For the assessment of the quality of each scheme, we record the final solution quality over multiple runs. Here, the final solution quality refers to a highly accurate Monte-Carlo approximation (using $m = 1000$ samples) of the expected objective function value of the solution returned after each optimization run).

General experimental settings	
Search space dimension size	$n = 10$
Evaluation budget per run	10,000
Runs per algorithmic scheme	50
Performance indicators	Final solution quality, approximated with Monte-Carlo integration using 1000 samples, (mean, std, median) over all runs, and rank sum for ranking of the algorithmic schemes

Table 8.1: The general experimental setup.

The test problems are enlisted in Table 8.2 and full descriptions can be found in Appendix B. The set of test problems is constructed based on test problems from literature. It incorporates different difficulties that can be encountered within these types of optimization problems.

The *RO Sphere Problem* is a modified version of the original *Sphere Problem*. Obviously, a myopic approach will perform much better on this test problem than any scheme that aims to approximate the expected objective function. However, it is still useful as a test problem for comparing the convergence limitations of the schemes designed for finding robust optima. The *RO Heaviside Sphere Problem* and the *RO Sawtooth Problem* are problems with a shifted robust optimizer that emerges due to a sharp ridge at the original optimum. The *RO Volcano Problem* is a problem with a plateau where the robust is located in the center. The other problems have multimodal objective functions in which the robust optimizer is classified as emergent. The *RO Pickelhaube Problem* is a problem with two peaks, and the algorithmic challenge is to target the most robust peak. The *RO Branke Multipeak Problem* has 2^n peaks, varying in robustness. The other two multipeak problems are more complex and provide cases in which the emergent optimizer is also shifted with respect to the original local optimizers.

Test problem	**Properties of the underlying signal function**	
RO Sphere Problem	unimodal	robust optimizer equals original optimizer
RO Heaviside Sphere Problem	unimodal	shifted robust optimizer
RO Sawtooth Problem	unimodal	shifted robust optimizer
RO Volcano Problem	unimodal	shifted robust optimizer
RO Pickelhaube Problem	multimodal	emergent robust optimizer
RO Branke's Multipeak Problem	multimodal	emergent robust optimizer
RO Multipeak F1 Problem	multimodal	emergent robust optimizer
RO Multipeak F2 Problem	multimodal	emergent robust optimizer

Table 8.2: The test problems used for empirical comparison.

The different evaluation schemes for finding robust optima that are compared in this empirical study are enlisted in Table 8.3. For the multi-evaluation methods, two variants are considered, namely $\mathrm{MEM}^{-}_{\mathrm{MS}}$ and $\mathrm{MEM}^{+}_{\mathrm{LHS}}$. Theses are considered to be tuned optimally for each test problem, therefore, Section 8.2 presents the results of the tuning of these methods. Both resampling methods are also considered in an adaptive averaging form: UH-$\mathrm{MEM}^{-}_{\mathrm{MS}}$ and UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$. These two adaptive averaging methods use the rank-based adaptive averaging approach for updating the sample size, as described in Section 7.2.5. The ABRSS and the Kriging metamodeling approach are used as presented in Section 7.2.7 and Section 7.2.8 respectively.

Evaluation schemes for finding robust optima	
Myopic	A canonical $(5/2_{DI}, 35)$-σSA-ES and CMA-ES.
$\mathbf{MEM^{-}_{MS}}$	The multi-evaluation method (MEM) using Monte-Carlo integration and resampling the disturbances for all individuals in a generation.
$\mathbf{MEM^{+}_{LHS}}$	The multi-evaluation method (MEM) using Latin Hypercube sampling and the same disturbances for all individuals in a generation.
UH-$\mathbf{MEM^{-}_{MS}}$	The rank-based adaptive averaging method using the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ evaluation scheme.
UH-$\mathbf{MEM^{+}_{LHS}}$	The rank-based adaptive averaging method using the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ evaluation scheme.
ABRSS	The archive based evaluation approach.
Kriging	The Kriging (metamodeling) based evaluation approach.

Table 8.3: The methods considered in the empirical study for finding robust optima.

8.2 Tuning the Static Resampling Schemes

For the empirical comparison of the schemes enlisted in Table 8.3, we consider optimally tuned versions of the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ and the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ schemes for each test problem. Hence, before presenting the results on the full comparison, Section 8.2.1 shows the results of the tuning experiments of the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ evaluation scheme and Section 8.2.2 shows the results of the tuning experiments of the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ evaluation scheme.

8.2.1 The Optimal Sample Size for $\mathrm{MEM}^{-}_{\mathrm{MS}}$

This experiment is done in order to determine, for each test problem, the optimal sample size for the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ evaluation scheme. Different instances of the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-CMA-ES are considered with varying sample sizes: $m = 1, 2, \ldots, 10$. These sample sizes are compared on the test problems listed in Table 8.2 using the experimental setup shown in Table 8.1. The results of these experiments are shown in the tables and figures of

Section 8.2.1.1 and Section 8.2.1.2 for the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-CMA-ES respectively.

Based on the results, we conclude that for the explicit averaging schemes (the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{-}_{\mathrm{MS}}$-CMA-ES) for each of the test problems with respect to the general experimental setup the optimal sample sizes lie at the values shown in Table 8.4. From these results we observe the trade-off in convergence speed versus convergence accuracy. It depends on the test problem which sample size is most suitable, i.e., there is no clear winner. It seems that for the CMA-ES a slightly higher sample size is required than for the $(5/2_{DI}, 35)$-σSA-ES. Also, we see that for these test problems, the SEM evaluation approach is not optimal in any case.

	$\mathrm{MEM}^{-}_{\mathrm{MS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES	**$\mathrm{MEM}^{-}_{\mathrm{MS}}$-CMA-ES**
RO Sphere Problem	m = 6	m = 10
RO Heaviside Sphere Problem	m = 9	m = 7
RO Sawtooth Problem	m = 4	m = 8
RO Volcano Problem	m = 5	m = 10
RO Pickelhaube Problem	m = 3	m = 4
RO Branke's Multipeak Problem	m = 3	m = 4
RO Multipeak F1 Problem	m = 2	m = 6
RO Multipeak F2 Problem	m = 6	m = 8

Table 8.4: The optimal sample size for the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ approach to achieve best convergence accuracy on a budget of $10,000$ function evaluations.

8.2.1.1 Results $\mathrm{MEM}^{-}_{\mathrm{MS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	8.61	11.88	4.85	20487	10
MEM2 MS-	4.13	0.33	4.06	13628	8
MEM3 MS-	7.64	12.98	3.97	12697	6
MEM4 MS-	4.34	3.00	3.81	9550	2
MEM5 MS-	5.33	7.12	3.85	9790	3
MEM6 MS-	**3.89**	**0.30**	**3.80**	**8695**	**1**
MEM7 MS-	6.27	8.41	3.95	11791	5
MEM8 MS-	6.46	9.32	3.89	10880	4
MEM9 MS-	4.96	5.40	4.03	13067	7
MEM10 MS-	7.42	10.27	4.18	14665	9

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.49	0.39	0.42	20794	10
MEM2 MS-	0.28	0.16	0.26	16241	9
MEM3 MS-	0.25	0.24	0.22	14843	8
MEM4 MS-	0.17	0.08	0.16	11802	7
MEM5 MS-	0.15	0.07	0.13	9861	3
MEM6 MS-	0.21	0.25	0.15	11672	5
MEM7 MS-	0.25	0.40	0.15	11718	6
MEM8 MS-	0.18	0.25	0.12	8944	2
MEM9 MS-	**0.18**	**0.37**	**0.10**	**8183**	**1**
MEM10 MS-	0.22	0.31	0.14	11192	4

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.33	0.12	0.29	19929	10
MEM2 MS-	0.26	0.04	0.25	12821	7
MEM3 MS-	0.30	0.13	0.25	12676	6
MEM4 MS-	**0.25**	**0.04**	**0.24**	**8721**	**1**
MEM5 MS-	0.26	0.06	0.25	10063	3
MEM6 MS-	0.25	0.03	0.24	9248	2
MEM7 MS-	0.27	0.09	0.24	10284	4
MEM8 MS-	0.26	0.06	0.25	10917	5
MEM9 MS-	0.29	0.10	0.26	14090	8
MEM10 MS-	0.29	0.08	0.27	16501	9

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.88	0.40	0.77	17980	10
MEM2 MS-	0.82	0.42	0.73	13189	8
MEM3 MS-	0.77	0.26	0.72	11845	5
MEM4 MS-	0.79	0.37	0.71	9808	2
MEM5 MS-	**0.75**	**0.26**	**0.71**	**8412**	**1**
MEM6 MS-	0.76	0.26	0.72	10592	4
MEM7 MS-	0.77	0.30	0.71	10340	3
MEM8 MS-	0.85	0.43	0.73	13006	7
MEM9 MS-	0.86	0.42	0.73	12872	6
MEM10 MS-	0.99	0.56	0.77	17206	9

RO PICKELHAUBE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.31	0.20	0.29	13072	6
MEM2 MS-	0.26	0.03	0.24	8561	3
MEM3 MS-	**0.26**	**0.03**	**0.24**	**7914**	**1**
MEM4 MS-	0.26	0.03	0.24	7980	2
MEM5 MS-	0.30	0.22	0.28	10740	4
MEM6 MS-	0.27	0.04	0.28	10900	5
MEM7 MS-	0.34	0.27	0.30	14810	7
MEM8 MS-	0.32	0.10	0.29	15474	8
MEM9 MS-	0.34	0.09	0.32	17538	9
MEM10 MS-	0.37	0.11	0.33	18261	10

RO BRANKE MULTIPEAK PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.43	0.09	0.41	17123	10
MEM2 MS-	0.41	0.03	0.40	13275	7
MEM3 MS-	**0.40**	**0.01**	**0.40**	**9237**	**1**
MEM4 MS-	0.42	0.06	0.40	11051	4
MEM5 MS-	0.40	0.01	0.40	11268	5
MEM6 MS-	0.40	0.02	0.40	9877	2
MEM7 MS-	0.40	0.01	0.40	10710	3
MEM8 MS-	0.42	0.06	0.40	12789	6
MEM9 MS-	0.41	0.02	0.40	14230	8
MEM10 MS-	0.43	0.05	0.41	15690	9

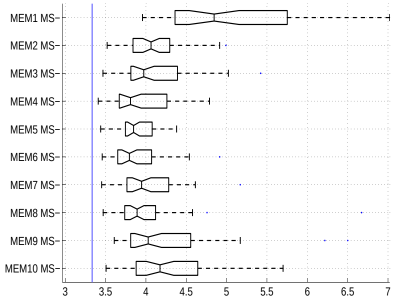
RO MULTIPEAK F1 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	-0.50	0.09	-0.52	13893	7
MEM2 MS-	**-0.56**	**0.08**	**-0.58**	**8450**	**1**
MEM3 MS-	-0.54	0.08	-0.57	9747	2
MEM4 MS-	-0.53	0.09	-0.55	11273	5
MEM5 MS-	-0.53	0.09	-0.56	10464	3
MEM6 MS-	-0.54	0.08	-0.57	10640	4
MEM7 MS-	-0.50	0.10	-0.53	13312	6
MEM8 MS-	-0.48	0.07	-0.47	15946	9
MEM9 MS-	-0.48	0.07	-0.48	16074	10
MEM10 MS-	-0.48	0.08	-0.46	15451	8

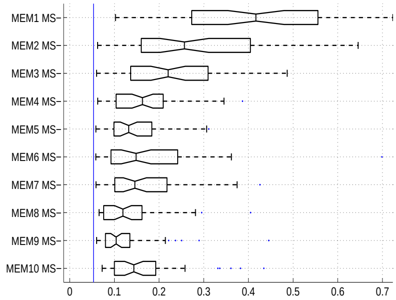
RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	-0.34	0.20	-0.25	18940	10
MEM2 MS-	-0.50	0.22	-0.62	12415	6
MEM3 MS-	-0.52	0.22	-0.64	11268	4
MEM4 MS-	-0.56	0.20	-0.66	10659	3
MEM5 MS-	-0.58	0.16	-0.65	10153	2
MEM6 MS-	**-0.57**	**0.18**	**-0.68**	**9857**	**1**
MEM7 MS-	-0.53	0.18	-0.63	12242	5
MEM8 MS-	-0.53	0.17	-0.59	12522	7
MEM9 MS-	-0.51	0.17	-0.55	13298	8
MEM10 MS-	-0.51	0.14	-0.52	13896	9

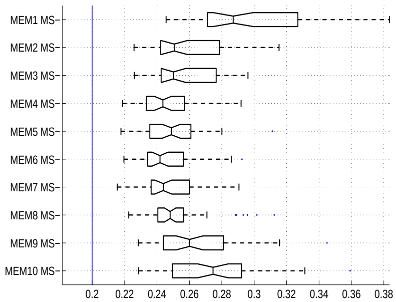
RO SPHERE PROBLEM



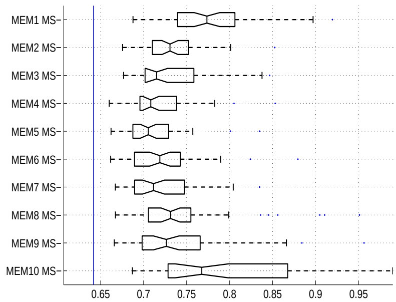
RO HEAVISIDE SPHERE PROBLEM



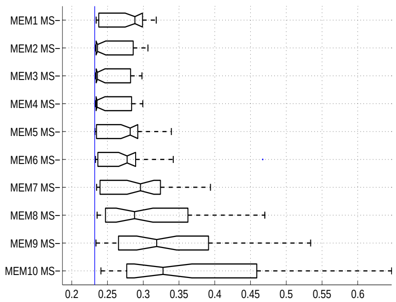
RO SAWTOOTH PROBLEM



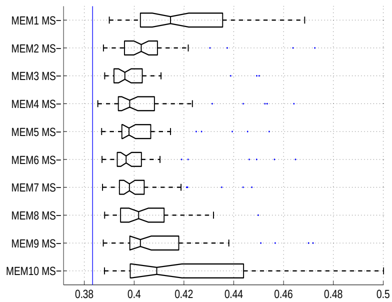
RO VOLCANO PROBLEM



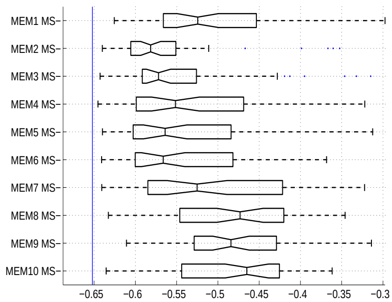
RO PICKELHAUBE PROBLEM



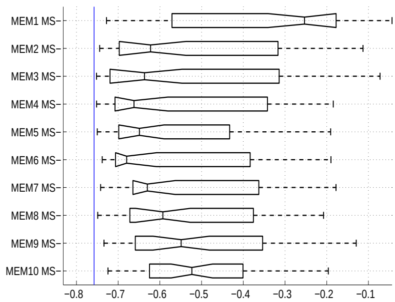
RO BRANKE MULTIPEAK PROBLEM



RO MULTIPEAK F1 PROBLEM



RO MULTIPEAK F2 PROBLEM



8.2.1.2 Results $\mathrm{MEM}^{-}_{\mathrm{MS}}$-CMA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	4.76	0.92	4.60	21779	10
MEM2 MS-	4.99	6.47	3.97	17971	9
MEM3 MS-	3.93	0.31	3.81	15537	8
MEM4 MS-	3.81	0.23	3.75	13268	7
MEM5 MS-	4.02	2.15	3.70	11048	6
MEM6 MS-	5.17	10.44	3.65	10029	4
MEM7 MS-	3.65	0.12	3.63	8952	3
MEM8 MS-	6.52	10.36	3.66	10610	5
MEM9 MS-	**4.21**	**4.10**	**3.57**	**7873**	**1**
MEM10 MS-	5.55	9.55	3.62	8183	2

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.27	0.15	0.26	20171	10
MEM2 MS-	0.23	0.48	0.16	16851	9
MEM3 MS-	0.16	0.09	0.14	15401	8
MEM4 MS-	0.23	0.56	0.10	12720	6
MEM5 MS-	0.11	0.04	0.10	11926	5
MEM6 MS-	0.26	0.63	0.11	12924	7
MEM7 MS-	0.09	0.07	0.08	8429	2
MEM8 MS-	0.17	0.40	0.09	10358	4
MEM9 MS-	0.14	0.36	0.08	9127	3
MEM10 MS-	**0.08**	**0.03**	**0.07**	**7343**	**1**

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.29	0.03	0.29	21324	10
MEM2 MS-	0.27	0.03	0.26	18009	9
MEM3 MS-	0.25	0.03	0.25	15046	8
MEM4 MS-	0.25	0.03	0.24	12988	7
MEM5 MS-	0.25	0.05	0.24	11193	5
MEM6 MS-	0.25	0.08	0.23	10102	4
MEM7 MS-	0.25	0.06	0.24	11798	6
MEM8 MS-	**0.23**	**0.01**	**0.23**	**7403**	**1**
MEM9 MS-	0.24	0.05	0.23	9228	3
MEM10 MS-	0.26	0.10	0.23	8159	2

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.76	0.06	0.76	20957	10
MEM2 MS-	0.71	0.04	0.71	16305	9
MEM3 MS-	0.75	0.30	0.70	15816	8
MEM4 MS-	0.70	0.03	0.69	12757	7
MEM5 MS-	0.69	0.02	0.69	11052	5
MEM6 MS-	0.73	0.33	0.69	11309	6
MEM7 MS-	0.72	0.29	0.68	9737	4
MEM8 MS-	0.76	0.38	0.68	9622	3
MEM9 MS-	0.68	0.02	0.68	8966	2
MEM10 MS-	**0.68**	**0.02**	**0.68**	**8729**	**1**

RO BRANKE MULTIPEAK PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	0.26	0.03	0.24	12457	7
MEM2 MS-	0.26	0.03	0.28	11990	4
MEM3 MS-	0.26	0.02	0.28	11023	3
MEM4 MS-	**0.29**	**0.22**	**0.28**	**10173**	**1**
MEM5 MS-	0.32	0.30	0.28	10280	2
MEM6 MS-	0.36	0.26	0.28	12318	6
MEM7 MS-	0.36	0.23	0.28	12174	5
MEM8 MS-	0.57	0.40	0.30	16495	10
MEM9 MS-	0.37	0.21	0.28	13432	8
MEM10 MS-	0.43	0.30	0.28	14908	9

RO MULTIPEAK F1 PROBLEM

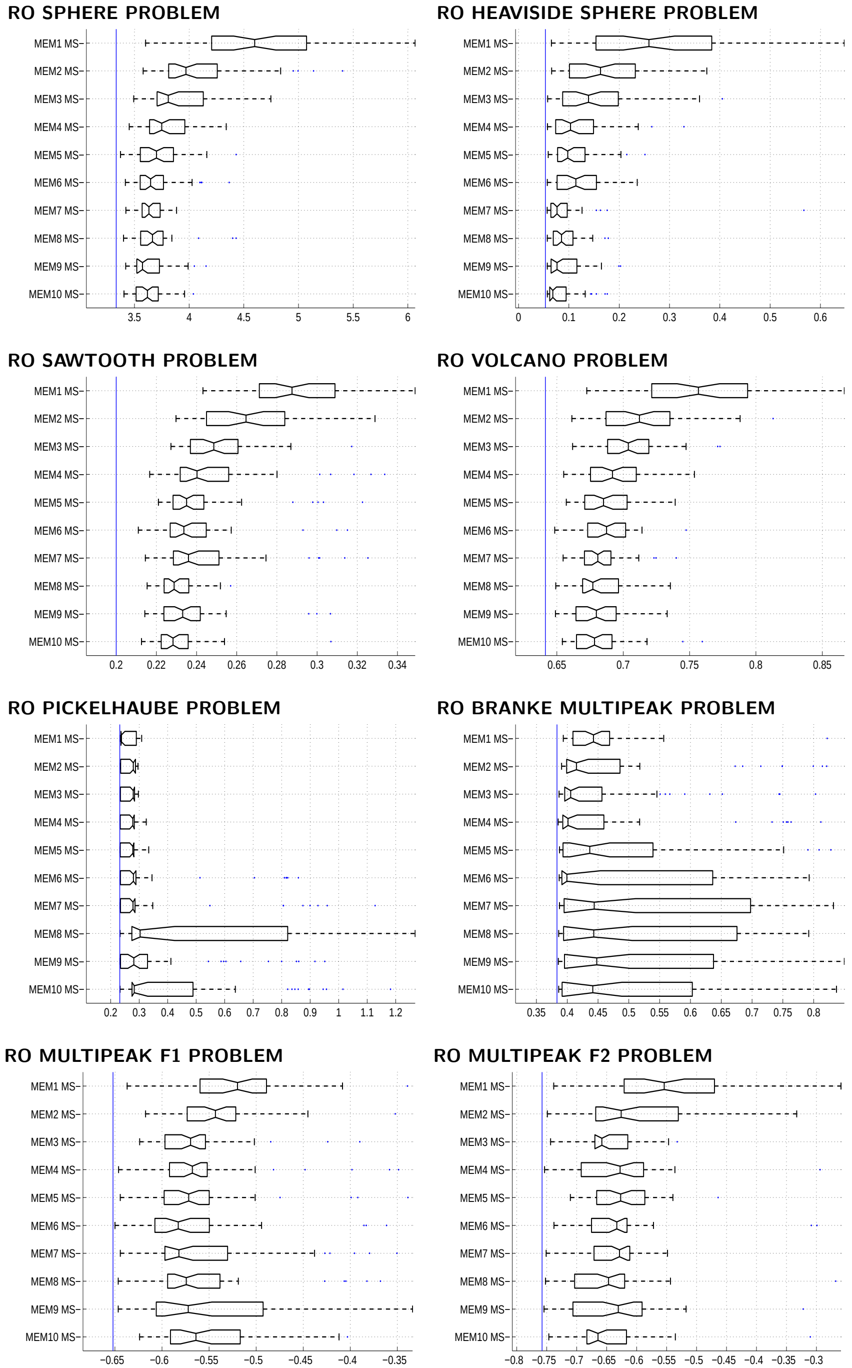
	Mean	Std	Med	∑#	#
MEM1 MS-	0.45	0.07	0.44	13547	7
MEM2 MS-	0.48	0.13	0.41	13000	6
MEM3 MS-	0.46	0.10	0.41	11511	3
MEM4 MS-	**0.46**	**0.12**	**0.40**	**10393**	**1**
MEM5 MS-	0.49	0.14	0.44	12248	5
MEM6 MS-	0.49	0.14	0.40	11151	2
MEM7 MS-	0.53	0.15	0.44	13875	10
MEM8 MS-	0.53	0.14	0.44	13685	8
MEM9 MS-	0.52	0.14	0.45	13790	9
MEM10 MS-	0.50	0.13	0.44	12050	4

RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	-0.51	0.07	-0.52	17313	10
MEM2 MS-	-0.53	0.07	-0.54	15347	9
MEM3 MS-	-0.56	0.05	-0.57	11377	3
MEM4 MS-	-0.55	0.08	-0.57	11941	6
MEM5 MS-	-0.56	0.06	-0.57	11243	2
MEM6 MS-	**-0.56**	**0.07**	**-0.58**	**10063**	**1**
MEM7 MS-	-0.54	0.09	-0.58	11649	5
MEM8 MS-	-0.55	0.07	-0.57	11595	4
MEM9 MS-	-0.54	0.09	-0.57	11942	7
MEM10 MS-	-0.54	0.07	-0.56	12780	8

RO FNIM F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 MS-	-0.53	0.12	-0.55	19015	10
MEM2 MS-	-0.60	0.11	-0.63	13772	9
MEM3 MS-	-0.64	0.05	-0.66	11116	3
MEM4 MS-	-0.63	0.09	-0.63	12606	7
MEM5 MS-	-0.60	0.12	-0.63	13417	8
MEM6 MS-	-0.63	0.08	-0.63	11119	4
MEM7 MS-	-0.64	0.05	-0.63	11705	6
MEM8 MS-	**-0.64**	**0.07**	**-0.65**	**10395**	**1**
MEM9 MS-	-0.61	0.14	-0.63	11465	5
MEM10 MS-	-0.64	0.07	-0.66	10640	2


RO SPHERE PROBLEM
MEM1 MS
MEM2 MS
MEM3 MS
MEM4 MS
MEM5 MS
MEM6 MS
MEM7 MS
MEM8 MS
MEM9 MS
MEM10 MS
3.5 4 4.5 5 5.5 6
RO HEAVISIDE SPHERE PROBLEM
0 0.1 0.2 0.3 0.4 0.5 0.6
RO SAWTOOTH PROBLEM
0.2 0.22 0.24 0.26 0.28 0.3 0.32 0.34
RO VOLCANO PROBLEM
0.65 0.7 0.75 0.8 0.85
RO PICKELHAUBE PROBLEM
0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2
RO BRANKE MULTIPEAK PROBLEM
0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75 0.8
RO MULTIPEAK F1 PROBLEM
−0.65 −0.6 −0.55 −0.5 −0.45 −0.4 −0.35
RO MULTIPEAK F2 PROBLEM
−0.8 −0.75 −0.7 −0.65 −0.6 −0.55 −0.5 −0.45 −0.4 −0.35 −0.3

8.2.2 The Optimal Sample Size for $\mathrm{MEM}^{+}_{\mathrm{LHS}}$

A first experiment is done in order to determine, for each test problem, the optimal sample size for the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ evaluation scheme. Different instances of the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-CMA-ES are considered with varying sample sizes: $m = 1, 2, \ldots, 10$. These sample sizes are compared on the test problems listed in Table 8.2 using the experimental setup shown in Table 8.1. The results of these experiments are shown in the tables and figures of Section 8.2.2.1 and Section 8.2.2.2 for the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-CMA-ES respectively.

Based on the results, we conclude that for the explicit averaging schemes, the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES and the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$-CMA-ES, for each of the test problems with respect to the general experimental setup the optimal sample sizes lie at the values shown in Table 8.5. The results show a similar picture as observed in the results of Section 8.2.1. Hence, also here the trade-off between convergence accuracy and convergence speed is well visible. Moreover, it seems that the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ method has a slightly higher optimal sample size, as compared to the $\mathrm{MEM}^{-}_{\mathrm{MS}}$. Also here, the CMA-ES seems to accept a higher sample size than the $(5/2_{DI}, 35)$-σSA-ES.

	$\mathrm{MEM}^{+}_{\mathrm{LHS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES	**$\mathrm{MEM}^{+}_{\mathrm{LHS}}$-CMA-ES**
RO Sphere Problem	m = 5	m = 8
RO Heaviside Sphere Problem	m = 8	m = 10
RO Sawtooth Problem	m = 7	m = 10
RO Volcano Problem	m = 6	m = 10
RO Pickelhaube Problem	m = 4	m = 3
RO Branke's Multipeak Problem	m = 5	m = 6
RO Multipeak F1 Problem	m = 5	m = 6
RO Multipeak F2 Problem	m = 3	m = 8

Table 8.5: The optimal sample size for the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach to achieve best convergence accuracy on a budget of $10,000$ function evaluations.

8.2.2.1 Results $\text{MEM}^{+}_{\text{LHS}}$-$(5/2_{DI}, 35)$-$\sigma$SA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	7.72	10.14	4.50	22457	10
MEM2 LHS+	4.36	6.10	3.48	17008	8
MEM3 LHS+	4.91	7.63	3.37	10973	5
MEM4 LHS+	3.36	0.01	3.35	6628	2
MEM5 LHS+	**3.97**	**4.35**	**3.35**	**5229**	**1**
MEM6 LHS+	5.65	9.21	3.35	6997	3
MEM7 LHS+	6.12	9.88	3.35	7696	4
MEM8 LHS+	5.18	8.64	3.41	14000	6
MEM9 LHS+	5.00	7.43	3.42	15258	7
MEM10 LHS+	5.48	8.21	3.54	19004	9

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.53	0.32	0.52	19521	9
MEM2 LHS+	0.64	0.26	0.67	21204	10
MEM3 LHS+	0.43	0.24	0.36	17431	8
MEM4 LHS+	0.31	0.31	0.25	14883	7
MEM5 LHS+	0.22	0.28	0.17	11439	6
MEM6 LHS+	0.25	0.47	0.17	10409	5
MEM7 LHS+	0.16	0.27	0.12	7649	3
MEM8 LHS+	**0.17**	**0.28**	**0.11**	**7185**	**1**
MEM9 LHS+	0.24	0.41	0.12	8203	4
MEM10 LHS+	0.13	0.09	0.12	7326	2

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.38	0.11	0.35	22446	10
MEM2 LHS+	0.28	0.05	0.27	19425	9
MEM3 LHS+	0.26	0.07	0.24	15067	8
MEM4 LHS+	0.25	0.08	0.23	10824	6
MEM5 LHS+	0.23	0.01	0.23	8006	2
MEM6 LHS+	0.25	0.08	0.23	8719	3
MEM7 LHS+	**0.24**	**0.06**	**0.23**	**7889**	**1**
MEM8 LHS+	0.26	0.10	0.23	9876	4
MEM9 LHS+	0.24	0.05	0.23	10530	5
MEM10 LHS+	0.27	0.09	0.24	12468	7

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.76	0.07	0.76	21241	10
MEM2 LHS+	0.66	0.01	0.66	13464	6
MEM3 LHS+	0.76	0.44	0.65	9709	4
MEM4 LHS+	0.72	0.34	0.65	7378	3
MEM5 LHS+	0.72	0.32	0.65	7193	2
MEM6 LHS+	**0.68**	**0.24**	**0.65**	**5547**	**1**
MEM7 LHS+	0.76	0.43	0.65	10555	5
MEM8 LHS+	0.84	0.54	0.66	13604	7
MEM9 LHS+	0.86	0.51	0.68	17686	8
MEM10 LHS+	0.83	0.42	0.69	18873	9

RO PICKELHAUBE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.27	0.03	0.28	13646	7
MEM2 LHS+	0.26	0.02	0.27	10019	4
MEM3 LHS+	0.30	0.25	0.27	8408	3
MEM4 LHS+	**0.25**	**0.03**	**0.23**	**6950**	**1**
MEM5 LHS+	0.28	0.19	0.23	7814	2
MEM6 LHS+	0.28	0.13	0.26	11643	5
MEM7 LHS+	0.30	0.21	0.27	13508	6
MEM8 LHS+	0.30	0.08	0.28	15522	8
MEM9 LHS+	0.38	0.23	0.32	17867	9
MEM10 LHS+	0.44	0.22	0.40	19873	10

RO BRANKE MULTIPEAK PROBLEM

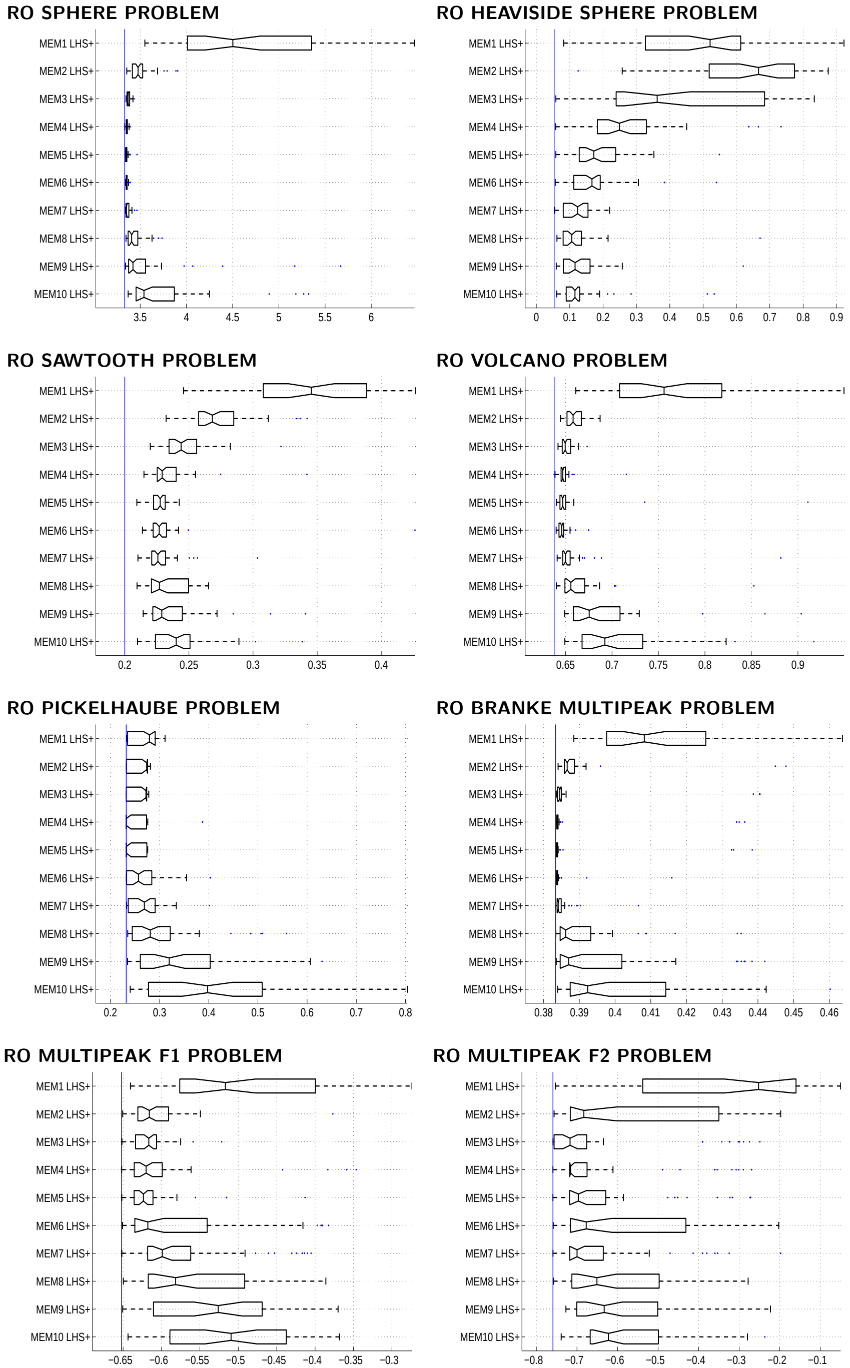
	Mean	Std	Med	∑#	#
MEM1 LHS+	0.42	0.06	0.41	21356	10
MEM2 LHS+	0.40	0.04	0.39	15944	7
MEM3 LHS+	0.39	0.04	0.38	11200	5
MEM4 LHS+	0.39	0.01	0.38	6913	3
MEM5 LHS+	**0.39**	**0.01**	**0.38**	**5516**	**1**
MEM6 LHS+	0.38	0.00	0.38	5742	2
MEM7 LHS+	0.39	0.00	0.38	9205	4
MEM8 LHS+	0.39	0.02	0.39	14390	6
MEM9 LHS+	0.41	0.04	0.39	16018	8
MEM10 LHS+	0.42	0.05	0.39	18966	9

RO MULTIPEAK F1 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	-0.49	0.10	-0.52	18600	10
MEM2 LHS+	-0.61	0.04	-0.62	9928	4
MEM3 LHS+	-0.62	0.02	-0.62	8693	3
MEM4 LHS+	-0.60	0.07	-0.62	8388	2
MEM5 LHS+	**-0.62**	**0.04**	**-0.62**	**7648**	**1**
MEM6 LHS+	-0.58	0.09	-0.62	10752	5
MEM7 LHS+	-0.57	0.07	-0.60	12625	6
MEM8 LHS+	-0.55	0.08	-0.58	14524	7
MEM9 LHS+	-0.53	0.08	-0.53	16398	8
MEM10 LHS+	-0.51	0.08	-0.51	17694	9

RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	-0.35	0.22	-0.25	20208	10
MEM2 LHS+	-0.57	0.20	-0.68	12574	6
MEM3 LHS+	**-0.64**	**0.17**	**-0.72**	**8437**	**1**
MEM4 LHS+	-0.64	0.16	-0.72	9752	2
MEM5 LHS+	-0.64	0.14	-0.70	9753	3
MEM6 LHS+	-0.60	0.17	-0.68	12047	5
MEM7 LHS+	-0.65	0.14	-0.70	10254	4
MEM8 LHS+	-0.61	0.14	-0.65	12637	7
MEM9 LHS+	-0.59	0.14	-0.63	14434	8
MEM10 LHS+	-0.58	0.12	-0.62	15154	9


RO SPHERE PROBLEM
MEM1 LHS+
MEM2 LHS+
MEM3 LHS+
MEM4 LHS+
MEM5 LHS+
MEM6 LHS+
MEM7 LHS+
MEM8 LHS+
MEM9 LHS+
MEM10 LHS+
3.5 4 4.5 5 5.5 6
RO HEAVISIDE SPHERE PROBLEM
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
RO SAWTOOTH PROBLEM
0.2 0.25 0.3 0.35 0.4
RO VOLCANO PROBLEM
0.65 0.7 0.75 0.8 0.85 0.9
RO PICKELHAUBE PROBLEM
0.2 0.3 0.4 0.5 0.6 0.7 0.8
RO BRANKE MULTIPEAK PROBLEM
0.38 0.39 0.4 0.41 0.42 0.43 0.44 0.45 0.46
RO MULTIPEAK F1 PROBLEM
−0.65 −0.6 −0.55 −0.5 −0.45 −0.4 −0.35 −0.3
RO MULTIPEAK F2 PROBLEM
−0.8 −0.7 −0.6 −0.5 −0.4 −0.3 −0.2 −0.1

8.2.2.2 Results $\text{MEM}^{+}_{\text{LHS}}$-CMA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	4.18	0.52	4.05	23549	10
MEM2 LHS+	4.06	4.47	3.42	20890	9
MEM3 LHS+	4.41	7.43	3.36	17957	8
MEM4 LHS+	3.35	0.01	3.34	15081	7
MEM5 LHS+	3.34	0.00	3.34	12612	6
MEM6 LHS+	4.45	7.88	3.34	9429	5
MEM7 LHS+	4.64	9.22	3.34	7519	4
MEM8 LHS+	**3.34**	**0.00**	**3.34**	**5681**	**1**
MEM9 LHS+	3.34	0.00	3.33	6016	2
MEM10 LHS+	3.34	0.01	3.34	6516	3

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.58	0.21	0.65	17032	8
MEM2 LHS+	0.70	0.35	0.72	18879	9
MEM3 LHS+	0.68	0.18	0.74	19527	10
MEM4 LHS+	0.59	0.38	0.74	16824	7
MEM5 LHS+	0.50	0.51	0.38	14130	6
MEM6 LHS+	0.28	0.22	0.19	10187	4
MEM7 LHS+	0.29	0.32	0.20	10288	5
MEM8 LHS+	0.16	0.07	0.15	6904	3
MEM9 LHS+	0.22	0.45	0.12	6435	2
MEM10 LHS+	**0.17**	**0.35**	**0.11**	**5044**	**1**

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.34	0.06	0.33	22652	10
MEM2 LHS+	0.28	0.05	0.27	19178	9
MEM3 LHS+	0.26	0.06	0.25	16762	8
MEM4 LHS+	0.25	0.05	0.24	14660	7
MEM5 LHS+	0.24	0.06	0.23	13380	6
MEM6 LHS+	0.25	0.10	0.23	11685	5
MEM7 LHS+	0.24	0.09	0.22	8434	4
MEM8 LHS+	0.22	0.02	0.22	6097	2
MEM9 LHS+	0.24	0.09	0.22	6698	3
MEM10 LHS+	**0.23**	**0.08**	**0.21**	**5704**	**1**

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.73	0.06	0.71	23447	10
MEM2 LHS+	0.66	0.01	0.65	19907	9
MEM3 LHS+	0.65	0.00	0.65	15461	8
MEM4 LHS+	0.69	0.30	0.65	13183	7
MEM5 LHS+	0.64	0.00	0.64	10220	6
MEM6 LHS+	0.64	0.00	0.64	8402	2
MEM7 LHS+	0.69	0.36	0.64	8420	3
MEM8 LHS+	0.67	0.18	0.64	9092	4
MEM9 LHS+	0.68	0.30	0.64	9126	5
MEM10 LHS+	**0.68**	**0.26**	**0.64**	**7992**	**1**

RO BRANKE MULTIPEAK PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	0.30	0.22	0.28	16067	10
MEM2 LHS+	0.25	0.02	0.23	11381	4
MEM3 LHS+	**0.29**	**0.19**	**0.25**	**10555**	**1**
MEM4 LHS+	0.27	0.07	0.27	11202	3
MEM5 LHS+	0.37	0.38	0.27	11768	5
MEM6 LHS+	0.36	0.37	0.27	12038	6
MEM7 LHS+	0.33	0.19	0.27	11198	2
MEM8 LHS+	0.40	0.31	0.27	12270	7
MEM9 LHS+	0.35	0.22	0.27	12784	8
MEM10 LHS+	0.45	0.29	0.28	15987	9

RO MULTIPEAK F1 PROBLEM

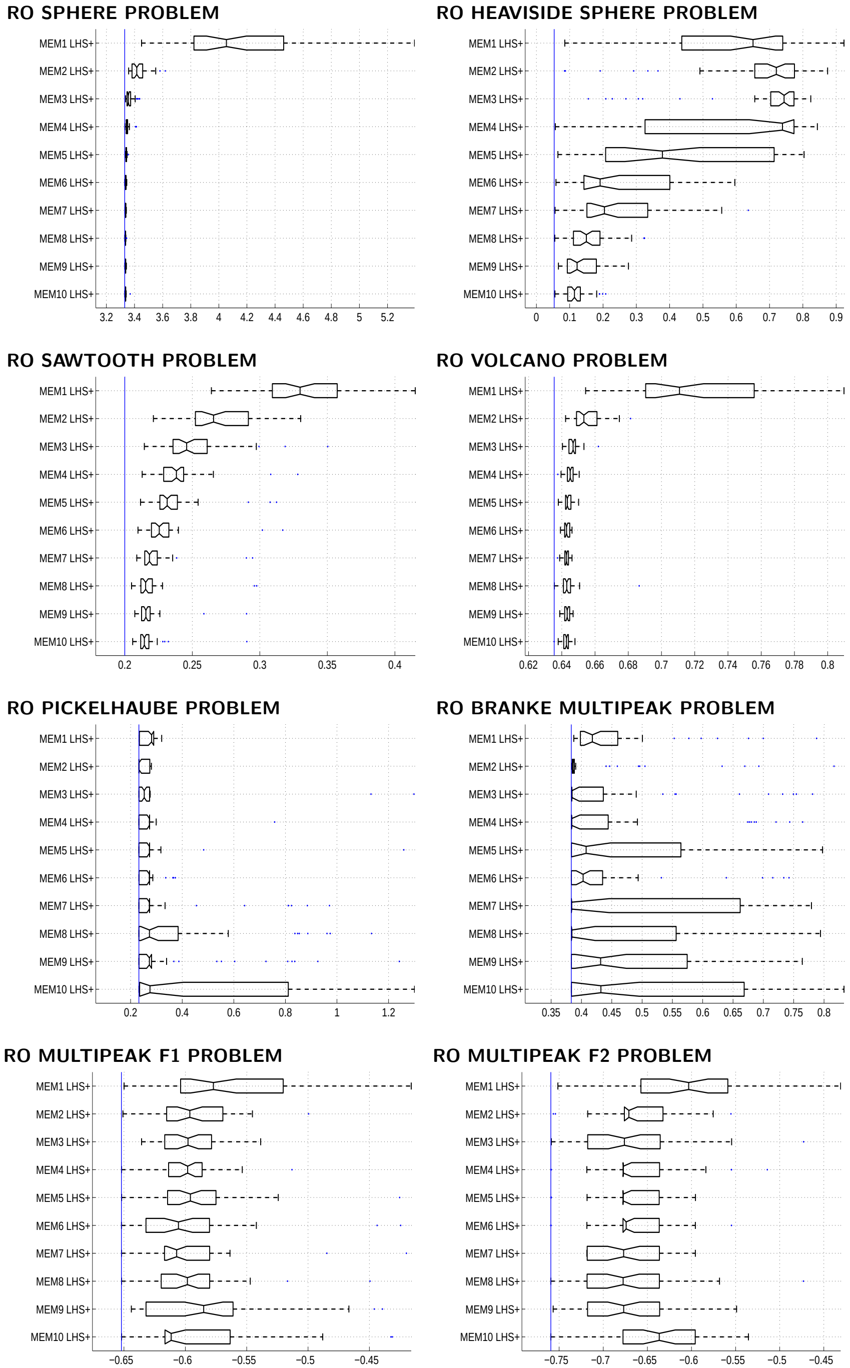
	Mean	Std	Med	∑#	#
MEM1 LHS+	0.45	0.09	0.42	15641	10
MEM2 LHS+	0.42	0.09	0.39	12631	6
MEM3 LHS+	0.45	0.11	0.38	13058	8
MEM4 LHS+	0.45	0.12	0.38	11931	4
MEM5 LHS+	0.48	0.13	0.41	12026	5
MEM6 LHS+	**0.44**	**0.10**	**0.40**	**10595**	**1**
MEM7 LHS+	0.50	0.15	0.38	11928	3
MEM8 LHS+	0.47	0.13	0.38	10956	2
MEM9 LHS+	0.48	0.13	0.43	12888	7
MEM10 LHS+	0.50	0.14	0.43	13596	9

RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	-0.55	0.08	-0.58	15431	10
MEM2 LHS+	-0.59	0.05	-0.60	12617	7
MEM3 LHS+	-0.59	0.04	-0.60	12144	4
MEM4 LHS+	-0.58	0.07	-0.60	12287	6
MEM5 LHS+	-0.58	0.06	-0.60	13128	8
MEM6 LHS+	**-0.59**	**0.07**	**-0.61**	**10713**	**1**
MEM7 LHS+	-0.58	0.08	-0.61	11500	2
MEM8 LHS+	-0.59	0.06	-0.60	11948	3
MEM9 LHS+	-0.56	0.10	-0.59	13280	9
MEM10 LHS+	-0.57	0.08	-0.61	12202	5

RO FNIM F2 PROBLEM

	Mean	Std	Med	∑#	#
MEM1 LHS+	-0.59	0.11	-0.60	18079	10
MEM2 LHS+	-0.66	0.05	-0.67	13923	9
MEM3 LHS+	-0.65	0.08	-0.68	12630	7
MEM4 LHS+	-0.66	0.05	-0.68	11758	6
MEM5 LHS+	-0.67	0.05	-0.68	10415	2
MEM6 LHS+	-0.64	0.09	-0.67	11753	5
MEM7 LHS+	-0.66	0.05	-0.68	11226	3
MEM8 LHS+	**-0.66**	**0.08**	**-0.68**	**10107**	**1**
MEM9 LHS+	-0.64	0.11	-0.68	11673	4
MEM10 LHS+	-0.64	0.09	-0.64	13686	8


RO SPHERE PROBLEM
MEM1 LHS+
MEM2 LHS+
MEM3 LHS+
MEM4 LHS+
MEM5 LHS+
MEM6 LHS+
MEM7 LHS+
MEM8 LHS+
MEM9 LHS+
MEM10 LHS+
3.2 3.4 3.6 3.8 4 4.2 4.4 4.6 4.8 5 5.2
RO HEAVISIDE SPHERE PROBLEM
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
RO SAWTOOTH PROBLEM
0.2 0.25 0.3 0.35 0.4
RO VOLCANO PROBLEM
0.62 0.64 0.66 0.68 0.7 0.72 0.74 0.76 0.78 0.8
RO PICKELHAUBE PROBLEM
0.2 0.4 0.6 0.8 1 1.2
RO BRANKE MULTIPEAK PROBLEM
0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75 0.8
RO MULTIPEAK F1 PROBLEM
−0.65 −0.6 −0.55 −0.5 −0.45
RO MULTIPEAK F2 PROBLEM
−0.75 −0.7 −0.65 −0.6 −0.55 −0.5 −0.45

8.3 Full Comparison of Schemes Finding Robust Optima

Finally, a full empirical comparison between all schemes for finding robust optima (see Table 8.3) is performed. The $\mathrm{MEM}^{-}_{\mathrm{MS}}$ and the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ schemes are assumed to be near optimally tuned, with the settings as found in Table 8.4 and Table 8.5. The question is: do the advanced evaluation schemes provide yield better results than optimally tuned static schemes?

Approach	**Strategy parameters**
Myopic	Default
MEM$^{-}_{\mathbf{MS}}$	See Table 8.4
MEM$^{+}_{\mathbf{LHS}}$	See Table 8.5
UH-MEM$^{-}_{\mathbf{MS}}$	$\theta = 0.6$, $\alpha = 1.2$
UH-MEM$^{+}_{\mathbf{LHS}}$	$\theta = 0.6$, $\alpha = 1.2$
ABRSS	Reference set sample size: $m = 2n = 20$
Kriging	Samples for metamodel construction: $n_{\mathrm{krig}} = 2n = 20$. Sampling on the metamodel for approximation of the expected fitness: $\mathrm{MEM50}^{+}_{\mathrm{LHS}}$

Table 8.6: Algorithm settings used for the empirical comparison between the evaluation techniques for finding robust optima.

Section 8.3.1 and Section 8.3.2 show the results for the $(5/2_{DI}, 35)$-σSA-ES and the CMA-ES respectively. Table 8.7 shows the combined rank scores of the compared schemes on all benchmark problems.

	$(5/2_{DI}, 35)$**-σSA-ES**	**CMA-ES**
Myopic	75,157	80,532
MEM$^{-}_{\mathbf{MS}}$	94,463	85,032
MEM$^{+}_{\mathbf{LHS}}$	48,882	49,925
UH-MEM$^{-}_{\mathbf{MS}}$	88,812	83,494
UH-MEM$^{+}_{\mathbf{LHS}}$	**46,923**	**48,684**
ABRSS	66,191	65,102
Kriging	70,972	78,631

Table 8.7: Combined rank sums on the set of test problems.

From the results, we see that the myopic approach outperforms the other approaches on three of the eight test problems. For the sphere problem this is not surprising, being a unimodal test problem where the original optimizer is also the robust optimizer. On the other two problems, Branke's multipeak problem and the pickelhaube problem, which are multimodal problems with emergent robust optimizers, the myopic approach apparently seems to be just as good in targeting the robust peak as the other approaches. Or, the myopic instances seem to be largely attracted by the robust peaks. For the pickelhaube problem, this is even more surprising given the fact that both peaks have the same basin of attraction areas and cover the same volume.

A possible explanation is the following: The attraction of a peak on an Evolution Strategy is determined by two matters: 1) the probability of generating individuals in that particular peak, and 2) the probability that a random solution of that peak is better than a random solution of the other peaks. For the pickelhaube problem, the probability of generating individuals in both peaks is equal, but the probability that a random solution of the robust peak is better than a random solution of the non-robust peak is larger. This could explain the inherent attraction of the robust peak on the myopic approaches.

On the other problems, however, the myopic approach is clearly not a good alternative. Furthermore, we observe that an optimally tuned $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ evaluation approach yields better results than an optimally tuned $\mathrm{MEM}^{-}_{\mathrm{MS}}$ approach. Also for the uncertainty handling schemes, the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ is to be recommended. When looking at the advanced approaches for finding robust optima, we see that the ABRSS yields remarkably good results on the Heaviside sphere problem and the sawtooth problem, but yields comparably poor results on the other test problems. The Kriging approach, on the other hand, yields good results on the multipeak f1 problem and the multipeak f2 problem, and average results on the other problems. An approach that yields good results across the spectrum is the UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach, which always ranks as one of the three best approaches. The $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ also yields good results across the set of test problems.

To summarize, besides the myopic approach, the $\mathrm{MEM}^{-}_{\mathrm{MS}}$, and the UH-$\mathrm{MEM}^{-}_{\mathrm{MS}}$, all approaches seem to yield comparable results. The ABRSS and the Kriging approach yield particularly good results on specific test problems, the UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$ yields good results across the full set of test problems, and also a well-tuned static $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ seems to work very well, however, slightly worse than the UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$.

8.3.1 Results $(5/2_{DI}, 35)$-σSA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	**5.16**	**9.09**	**3.33**	**1913**	**1**
MEM MS-	5.63	9.02	3.84	14775	7
MEM LHS+	4.59	6.13	3.35	5126	2
UH MS-	3.64	0.15	3.60	12523	6
UH LHS+	4.24	6.21	3.35	5538	3
ABRSS	4.56	7.23	3.49	10593	4
Kriging	5.56	8.43	3.52	10957	5

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.79	0.03	0.78	16075	7
MEM MS-	0.19	0.24	0.12	11193	6
MEM LHS+	0.14	0.26	0.09	9555	4
UH MS-	0.07	0.02	0.07	6324	2
UH LHS+	0.08	0.02	0.07	6520	3
ABRSS	**0.09**	**0.28**	**0.05**	**1721**	**1**
Kriging	0.15	0.24	0.11	10037	5

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.55	0.01	0.55	15370	7
MEM MS-	0.27	0.09	0.25	11067	6
MEM LHS+	0.27	0.11	0.23	7484	3
UH MS-	0.26	0.07	0.25	10538	5
UH LHS+	0.24	0.09	0.22	5655	2
ABRSS	**0.21**	**0.06**	**0.20**	**1573**	**1**
Kriging	0.28	0.11	0.24	9738	4

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.72	0.31	0.68	11078	5
MEM MS-	0.82	0.43	0.71	14590	7
MEM LHS+	0.75	0.42	0.65	3989	2
UH MS-	0.86	0.55	0.68	11705	6
UH LHS+	**0.65**	**0.00**	**0.65**	**2709**	**1**
ABRSS	0.77	0.40	0.66	8660	3
Kriging	0.67	0.01	0.67	8694	4

RO PICKELHAUBE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	**0.25**	**0.02**	**0.23**	**4075**	**1**
MEM MS-	0.29	0.18	0.28	12691	7
MEM LHS+	0.28	0.14	0.27	7743	3
UH MS-	0.28	0.13	0.28	10969	6
UH LHS+	0.25	0.02	0.23	6462	2
ABRSS	0.26	0.02	0.27	10026	5
Kriging	0.25	0.02	0.23	9459	4

RO BRANKE MULTIPEAK PROBLEM

	Mean	Std	Med	∑#	#
Myopic	**0.39**	**0.02**	**0.38**	**3177**	**1**
MEM MS-	0.41	0.02	0.40	13167	7
MEM LHS+	0.38	0.01	0.38	4032	2
UH MS-	0.41	0.03	0.40	13033	6
UH LHS+	0.39	0.04	0.38	6185	3
ABRSS	0.44	0.11	0.40	12109	5
Kriging	0.39	0.02	0.39	9722	4

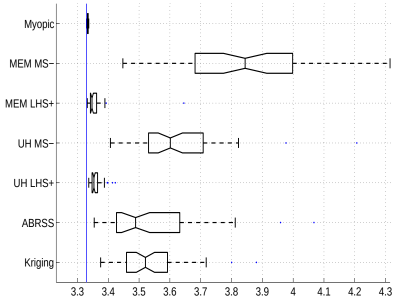
RO MULTIPEAK F1 PROBLEM

	Mean	Std	Med	∑#	#
Myopic	-0.53	0.09	-0.57	9958	6
MEM MS-	-0.55	0.08	-0.57	9572	4
MEM LHS+	**-0.60**	**0.06**	**-0.62**	**4363**	**1**
UH MS-	-0.46	0.06	-0.47	13553	7
UH LHS+	-0.56	0.09	-0.60	7828	3
ABRSS	-0.51	0.12	-0.57	9926	5
Kriging	-0.59	0.04	-0.60	6225	2

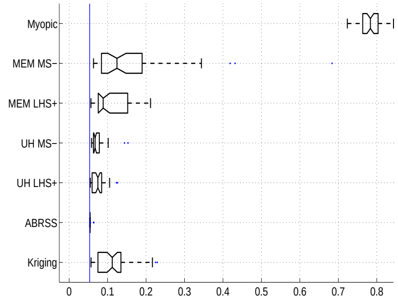
RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
Myopic	-0.29	0.17	-0.27	13511	7
MEM MS-	-0.57	0.18	-0.66	7408	4
MEM LHS+	-0.56	0.21	-0.68	6590	3
UH MS-	-0.45	0.14	-0.45	10167	5
UH LHS+	**-0.61**	**0.16**	**-0.67**	**6026**	**1**
ABRSS	-0.38	0.23	-0.30	11583	6
Kriging	-0.63	0.16	-0.67	6140	2

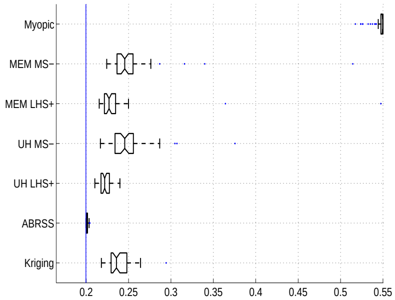
RO SPHERE PROBLEM



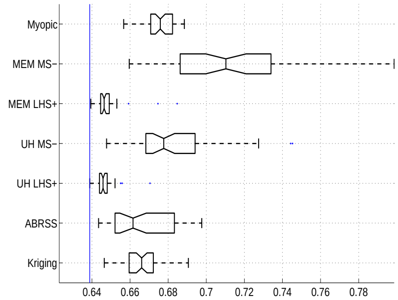
RO HEAVISIDE SPHERE PROBLEM



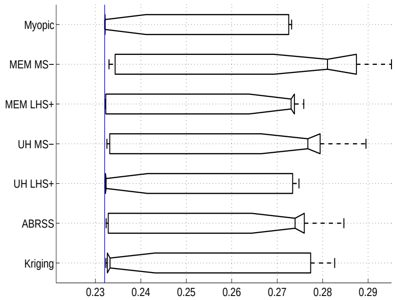
RO SAWTOOTH PROBLEM



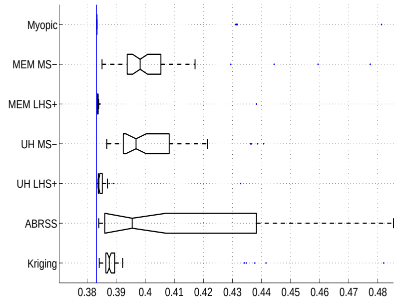
RO VOLCANO PROBLEM



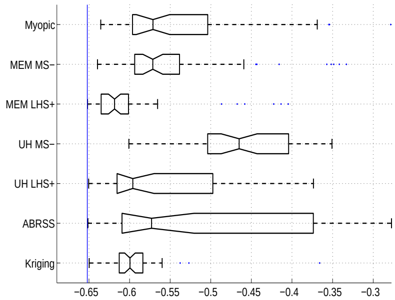
RO PICKELHAUBE PROBLEM



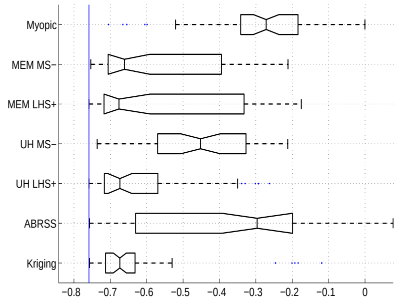
RO BRANKE MULTIPEAK PROBLEM



RO MULTIPEAK F1 PROBLEM



RO MULTIPEAK F2 PROBLEM



8.3.2 Results CMA-ES

RO SPHERE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	**3.33**	**0.00**	**3.33**	**1924**	**1**
MEM MS-	3.61	0.11	3.60	13651	7
MEM LHS+	4.39	7.47	3.34	3690	2
UH MS-	5.80	11.43	3.49	10832	4
UH LHS+	4.45	7.80	3.34	6111	3
ABRSS	3.61	0.16	3.57	12941	6
Kriging	3.55	0.09	3.54	12276	5

RO HEAVISIDE SPHERE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.78	0.03	0.78	16225	7
MEM MS-	0.13	0.28	0.08	9209	4
MEM LHS+	0.14	0.06	0.13	11649	6
UH MS-	0.07	0.01	0.06	6364	3
UH LHS+	0.06	0.01	0.06	5305	2
ABRSS	**0.06**	**0.00**	**0.05**	**1720**	**1**
Kriging	0.11	0.04	0.11	10953	5

RO SAWTOOTH PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.56	0.02	0.55	16182	7
MEM MS-	0.24	0.02	0.23	9923	4
MEM LHS+	0.22	0.02	0.21	5212	2
UH MS-	0.25	0.06	0.24	11216	6
UH LHS+	0.22	0.02	0.21	5514	3
ABRSS	**0.21**	**0.02**	**0.20**	**2267**	**1**
Kriging	0.25	0.05	0.24	11111	5

RO VOLCANO PROBLEM

	Mean	Std	Med	∑#	#
Myopic	0.68	0.01	0.68	14242	7
MEM MS-	0.72	0.31	0.67	12347	6
MEM LHS+	0.70	0.37	0.64	3046	2
UH MS-	0.78	0.47	0.66	9410	3
UH LHS+	**0.64**	**0.00**	**0.64**	**2610**	**1**
ABRSS	0.67	0.02	0.66	9546	4
Kriging	0.67	0.01	0.67	10224	5

RO PICKELHAUBE PROBLEM

	Mean	Std	Med	∑#	#
Myopic	**0.26**	**0.02**	**0.27**	**6018**	**1**
MEM MS-	0.28	0.12	0.23	11225	7
MEM LHS+	0.30	0.22	0.27	7912	3
UH MS-	0.25	0.02	0.23	9392	4
UH LHS+	0.28	0.22	0.23	6287	2
ABRSS	0.29	0.22	0.27	10244	5
Kriging	0.32	0.27	0.27	10347	6

RO BRANKE MULTIPEAK PROBLEM

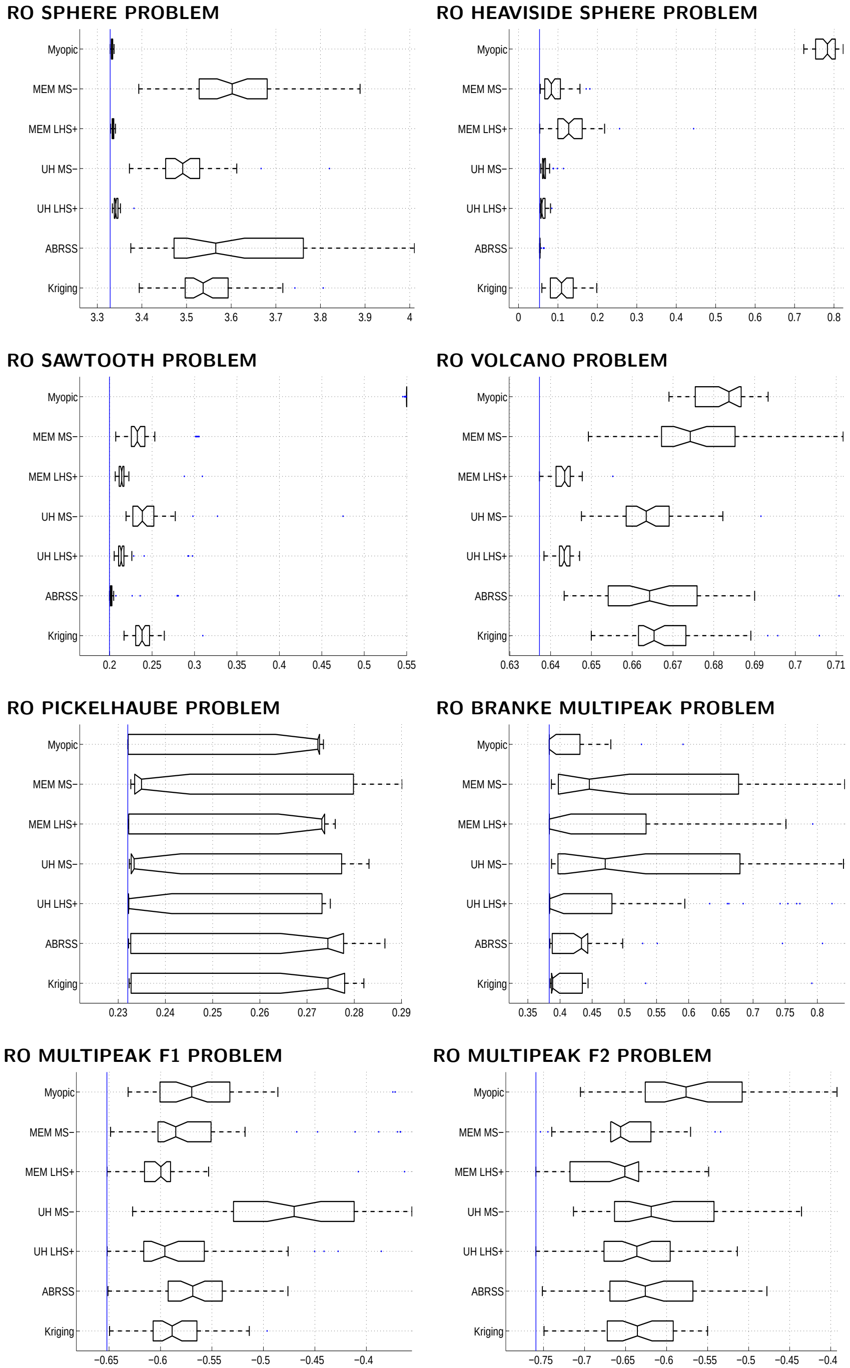
	Mean	Std	Med	∑#	#
Myopic	**0.41**	**0.04**	**0.38**	**4411**	**1**
MEM MS-	0.52	0.15	0.45	12012	6
MEM LHS+	0.47	0.13	0.38	7088	2
UH MS-	0.54	0.15	0.47	12378	7
UH LHS+	0.47	0.13	0.38	7925	4
ABRSS	0.44	0.08	0.43	9804	5
Kriging	0.41	0.06	0.39	7807	3

RO MULTIPEAK F1 PROBLEM

	Mean	Std	Med	∑#	#
Myopic	-0.56	0.05	-0.57	9338	6
MEM MS-	-0.55	0.09	-0.59	8548	4
MEM LHS+	**-0.60**	**0.05**	**-0.60**	**5373**	**1**
UH MS-	-0.48	0.07	-0.47	13621	7
UH LHS+	-0.56	0.10	-0.60	7431	2
ABRSS	-0.56	0.07	-0.57	9186	5
Kriging	-0.57	0.08	-0.59	7928	3

RO MULTIPEAK F2 PROBLEM

	Mean	Std	Med	∑#	#
Myopic	-0.56	0.10	-0.58	12192	7
MEM MS-	-0.62	0.10	-0.66	8117	4
MEM LHS+	**-0.64**	**0.11**	**-0.65**	**5955**	**1**
UH MS-	-0.58	0.12	-0.62	10281	6
UH LHS+	-0.62	0.10	-0.64	7501	2
ABRSS	-0.61	0.08	-0.63	9394	5
Kriging	-0.63	0.09	-0.64	7985	3


RO SPHERE PROBLEM
RO HEAVISIDE SPHERE PROBLEM
RO SAWTOOTH PROBLEM
RO VOLCANO PROBLEM
RO PICKELHAUBE PROBLEM
RO BRANKE MULTIPEAK PROBLEM
RO MULTIPEAK F1 PROBLEM
RO MULTIPEAK F2 PROBLEM
Myopic
MEM MS−
MEM LHS+
UH MS−
UH LHS+
ABRSS
Kriging

8.4 Summary and Discussion

This chapter has presented the results of an empirical comparison on different techniques that can be used within Evolution Strategies for finding robust optima. From this empirical study we conclude that the myopic approach can be a very risky approach when aiming to find robust optima. It highly depends on the particularities of the objective function landscape, whether this approach will work. However, because the particularities of the objective function landscape are not known beforehand, this approach is not recommendable.

When considering optimally tuned resampling approaches, the results clearly show that the $\mathrm{MEM}^{-}_{\mathrm{MS}}$ is, overall, outperformed by the $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach. Hence, it is recommended to use Latin Hypercube Sampling within a resampling approach as well as to use the same sampling disturbances/perturbations for all individuals in the population.

When considering the more advanced methods, we see that the ABRSS and the Kriging approach yield particularly good results on specific test problems, but do not yield exceptionally good performance on other test problems. Of the two, the Kriging approach seems to yield slightly more stable results across all test problems, but ABRSS yields exceptionally good results on the Heaviside sphere and the sawtooth problem.

Two methods that yield a good performance across the full set of test problems are the UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach and a well tuned $\mathrm{MEM}^{+}_{\mathrm{LHS}}$ approach. However, the empirical results in Section 8.2.1 and Section 8.2.2 have shown that the setting of the sample size highly affects the performance of Evolution Strategies for finding robust optima. Therefore, the adaptive averaging approach, the UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$ seems to be a promising method for finding robust optima. Based on these empirical results, we conclude therefore by recommending UH-$\mathrm{MEM}^{+}_{\mathrm{LHS}}$ as the most promising approach when aiming to find robust optima.