A Nonmonotone Line Search Method for Noisy Minimization - Additional Results

Nataša Krejić* Zorana Lužanin* Filip Nikolovski†
Irena Stojkovska†

December 17, 2013

The algorithm stops when the criterium

$$|F(x_k)| < (1 + 2\sigma) \cdot |F(x_0)| \cdot 10^{-3} \tag{1}$$

is satisfied, where σ is the noise level. Alternatively, the algorithm terminates when the maximum number of $400 \cdot n$ function evaluations is exceeded.

For each problem, a series of 50 independent runs are conducted. We considered a test run *successful* if the stopping criterium (1) is satisfied before exceeding the maximal number of function evaluations.

To compare the performance of methods from the set of methods $\mathcal{M}$ over the set of problems $\mathcal{P}$, the following indices (see [3]), for each method i are used. Let us denote by $\widetilde{\mathcal{P}} = \{j \in \mathcal{P} | \exists i \in \mathcal{M},\ N_{ij} > 0\}$, $\mathcal{P}_i = \{j \in \mathcal{P} | N_{ij} > 0\}$, $\mathcal{M}_j = \{i \in \mathcal{M} | N_{ij} > 0\}$, with N_{ij} being the number of successful runs out of 50 for the method i solving the problem j. Then we have:

- The efficiency index,

$$E_i = \frac{1}{|\mathcal{P}_i|} \sum_{j \in \mathcal{P}_i} \frac{\min_{i \in \mathcal{M}_j} \varphi_{ij}}{\varphi_{ij}},$$

  where φ_{ij} is the average relative number of function evaluations needed for the method i to solve the problem j, in the successful runs.

*Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad, Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia, e-mail: {natasak@uns.ac.rs, zorana@dmi.uns.ac.rs}

†Department of Mathematics, Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Arhimedova 3, 1000 Skopje, Macedonia, e-mail: {filipnikolovski@gmail.com, irenatra@pmf.ukim.mk}

- The robustness index,

$$R_i = \frac{1}{|\widetilde{\mathcal{P}}|} \sum_{j \in \widetilde{\mathcal{P}}} \frac{N_{ij}}{50}.$$

- The combined index,

$$C_i = R_i \cdot E_i.$$

- The nonmonotonicity index, as in [2],

$$\nu_i = \frac{1}{|\mathcal{P}_i|} \sum_{j \in \mathcal{P}_i} \mu_{ij},$$

  where $\mu_{ij} = \frac{1}{N_{ij}} \sum_{r \in \mathcal{N}_{ij}} \mu_{ij}^r$, with $\mathcal{N}_{ij}$ the set of successful runs out of 50 for the method i solving the problem j, where $\mu_{ij}^r = \frac{s}{k}$ is the nonmonotonicity coefficient, defined as the ratio of the number of iterations s at which the accepted step size would not be accepted if the line search rule was LS1, and the number of all iterations k at rth run for the method i solving the problem j.

In Tables 2-7, the values of the above defined indexes for different noise levels are presented. The numbers of solved problems from the set of problems $\mathcal{P}$ given in Table 1, at each of the noise levels, are reported at the end of the tables.

To compare the performance of methods we present the performance profiles defined in [1]. The measure for the performance profile is defined as the number of function evaluations as common for noisy problems. More precisely, let us denote by N_{ij} the number of successful runs out of 50 for the method i solving the problem j and let φ_{ij} be the average number of function evaluations needed for the method i to solve the problem j, in the successful runs. Clearly the smaller φ_{ij} is the method i is more efficient. However in noisy environment one is also interested in the variation of φ_{ij} as it shows, roughly speaking, the robustness of the method. Thus we take the linear combination of φ_{ij} and the corresponding standard deviation $\sigma(\varphi)_{ij}$ for the performance measure i.e. the performance measure is

$$\pi_{ij} = \varphi_{ij} + \sigma(\varphi)_{ij}.$$

Problem	n	x_0
Helical valley function	3	$(-1, 0, 0)$
Biggs EXP6 function	6	$(10, 20, 10, 10, 10, 10)$
Gaussian function	3	$(4, 10, 0)$
Powell badly scaled function	2	$(0, 5)$
Box three-dimensionaly function	3	$(0, 10, 20)$
Variably dimensioned function	10	$(9/10, 8/10, ..., 0)$
Watson function	6	$(0, 0, ..., 0)$
Penalty function I	4	$(1, 2, 3, 4)$
Penalty function II	4	$(5/2, 5/2, 5/2, 5/2)$
Brown badly scaled function	2	$(1, 1)$
Brown and Dennis function	4	$(25, 5, -5, 1)$
Gulf research and development function	3	$(5, 2.5, 0.15)$
Trigonometric function	10	$(1, 1, ..., 1)$
Extended Rosenbrock function	10	$(-1.2, 1, ..., -1.2, 1)$
Extended Powell singular function	12	$(3, -1, 0, 1, ..., 3, -1, 0, 1)$
Beale function	2	$(1, 1)$
Wood function	4	$(-3, -1, -3, -1)$
Chebyquad function	10	$(5/11, 10/11, ..., 50/11)$

Table 1: Test problems

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.534046	0.631476	0.632701	0.601909
BFGS	0.763998	0.695993	0.642208	0.660295
SR1	0.845075	0.815584	0.793245	0.770265
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.622667	0.705333	0.698667	0.712
BFGS	0.605333	0.742667	0.752	0.76
SR1	0.352	0.466667	0.473333	0.476
Combined index (C_i)				
	LS1	LS2	LS3	LS4
SGR	0.332533	0.445401	0.442047	0.428559
BFGS	0.462473	0.516891	0.482941	0.501824
SR1	0.297466	0.380606	0.375469	0.366646
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.325829	0.332914	0.345153
BFGS	0	0.380896	0.376952	0.427847
SR1	0	0.325919	0.33615	0.331305

Table 2: The values of the indexes for $\sigma = 0.01$ (15 problems solved)

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.607516	0.590386	0.541389	0.532668
BFGS	0.675353	0.730964	0.616066	0.564471
SR1	0.802454	0.82221	0.757181	0.736553
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.45	0.547143	0.585714	0.588571
BFGS	0.447143	0.54	0.6	0.601429
SR1	0.348571	0.358571	0.365714	0.364286
Combined index (C_i)				
	LS1	LS2	LS3	LS4
SGR	0.273382	0.323025	0.317099	0.313513
BFGS	0.301979	0.394721	0.36964	0.339489
SR1	0.279713	0.294821	0.276912	0.268316
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.295866	0.339607	0.369187
BFGS	0	0.311726	0.383475	0.401783
SR1	0	0.281227	0.333832	0.354457

Table 3: The values of the indexes for $\sigma = 0.1$ (14 problems solved)

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.835858	0.79462	0.840589	0.83085
BFGS	0.489351	0.461739	0.458317	0.437478
SR1	0.422568	0.426319	0.487814	0.492092
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.212222	0.262222	0.277778	0.288889
BFGS	0.386667	0.421111	0.477778	0.463333
SR1	0.278889	0.258889	0.286667	0.274444
Combined index (C_i)				
	LS1	LS2	LS3	LS4
SGR	0.177388	0.208367	0.233497	0.240023
BFGS	0.189216	0.194443	0.218974	0.202698
SR1	0.117849	0.110369	0.13984	0.135052
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.233679	0.31913	0.30813
BFGS	0	0.28032	0.410535	0.421942
SR1	0	0.260494	0.443768	0.419285

Table 4: The values of the indexes for $\sigma = 0.5$ (18 problems solved)

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.715657	0.767285	0.668126	0.696932
BFGS	0.417906	0.376875	0.405383	0.385202
SR1	0.436637	0.454467	0.451175	0.625057
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.08	0.089412	0.104706	0.103529
BFGS	0.234118	0.238824	0.32	0.284706
SR1	0.125882	0.138824	0.164706	0.143529
Combined index (C_i)				
	LS1	LS2	LS3	LS4
SGR	0.057253	0.068604	0.069957	0.072153
BFGS	0.097839	0.090007	0.129723	0.109669
SR1	0.054965	0.063091	0.074311	0.089714
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.203899	0.330797	0.296876
BFGS	0	0.16416	0.391475	0.403974
SR1	0	0.187273	0.344953	0.267892

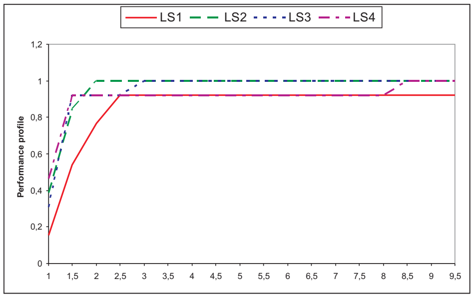
Table 5: The values of the indexes for $\sigma = 1$ (17 problems solved)

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.616157	0.791697	0.772204	0.785307
BFGS	0.677091	0.664173	0.675023	0.721942
SR1	0.471902	0.520049	0.553782	0.597666
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.050588	0.048235	0.057647	0.069412
BFGS	0.138824	0.147059	0.231765	0.18
SR1	0.063529	0.063529	0.082353	0.076471
Combined index (C_i)				
	LS1	LS2	LS3	LS4
SGR	0.03117	0.038188	0.044515	0.05451
BFGS	0.093996	0.097673	0.156447	0.12995
SR1	0.02998	0.033038	0.045606	0.045704
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.115811	0.357973	0.255958
BFGS	0	0.068474	0.413639	0.38227
SR1	0	0.1016	0.377572	0.333091

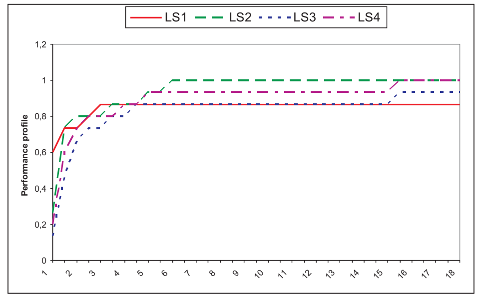
Table 6: The values of the indexes for $\sigma = 5$ (17 problems solved)

Efficiency index (E_i)				
	LS1	LS2	LS3	LS4
SGR	0.739957	0.663176	0.81703	0.744727
BFGS	0.694691	0.61851	0.728504	0.698589
SR1	0.511933	0.402908	0.685891	0.650842
Robustness index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.061176	0.062353	0.078824	0.075294
BFGS	0.189412	0.194118	0.282353	0.201176
SR1	0.083529	0.081176	0.12	0.102353
Combined index (R_i)				
	LS1	LS2	LS3	LS4
SGR	0.045268	0.041351	0.064401	0.056074
BFGS	0.131583	0.120064	0.205695	0.14054
SR1	0.042761	0.032707	0.082307	0.066616
Nonmonotonicity index (ν_i)				
	LS1	LS2	LS3	LS4
SGR	0	0.154676	0.349745	0.281658
BFGS	0	0.069054	0.408121	0.37676
SR1	0	0.107819	0.360692	0.304992

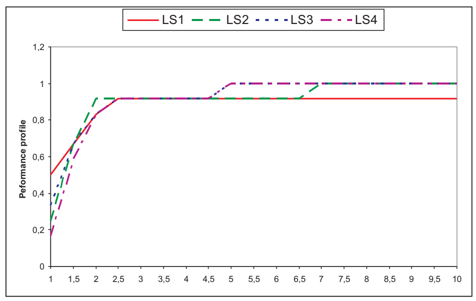
Table 7: The values of the indexes for $\sigma = 10$ (17 problems solved)



a. SGR methods
(13 problems solved)



b. BFGS methods
(15 problems solved)



c. SR1 methods
(12 problems solved)

Figure 1: SGR, BFGS, SR1 methods at the noise level $\sigma = 0.01$

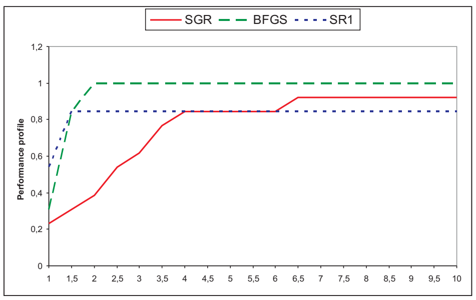
References

[1] E. D. Dolan, J. J. Moré, *Benchmarking optimization software with performance profiles,* Math. Program., Ser. A, Vol. 91 (2002), pp.201-213

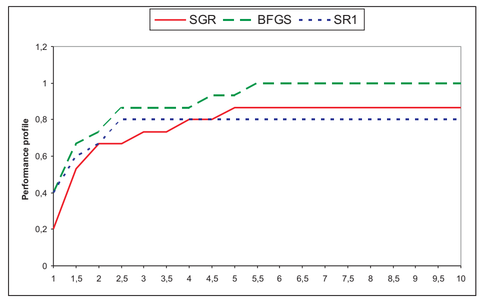
[2] N. Krejić, N. Krklec Jerinkić *Nonmonotone line search methods with variable sample size,*

`http://www.optimization-online.org/DB_HTML/2013/05/3902.html`

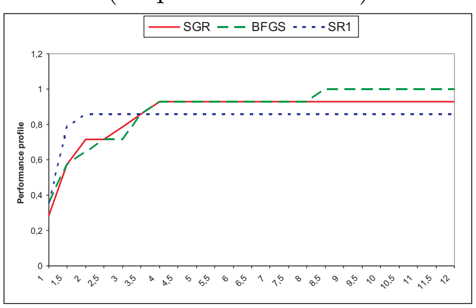
[3] N. Krejić, S. Rapajić *Gllobaly convergent Jacobian smoothing inexact Newton methods for NCP,* Cpmputational Optimization and Applications, Vol. 41, No. 2 (2008), pp.243-261



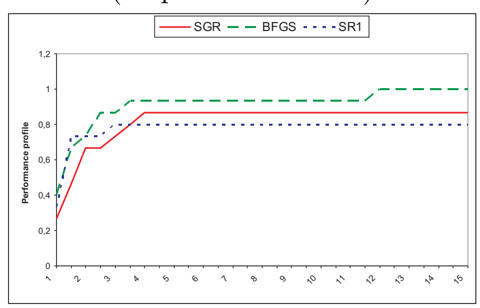
a. Line-search rule LS1
(13 problems solved)



b. Line-search rule LS2
(15 problems solved)

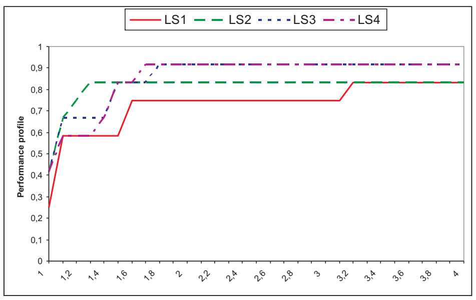


c. Line-search rule LS3
(14 problems solved)

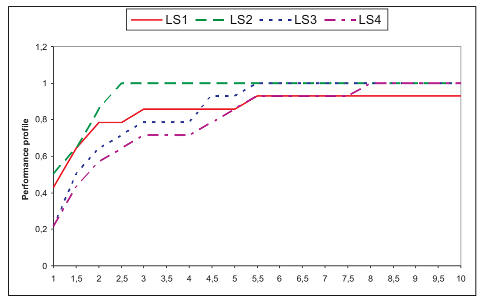


d. Line-search rule LS4
(15 problems solved)

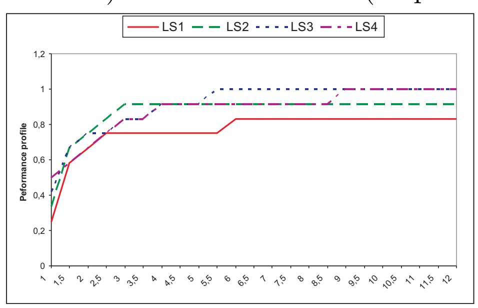
Figure 2: The line-search rules at the noise level $\sigma = 0.01$



a. SGR methods
(12 problems solved)

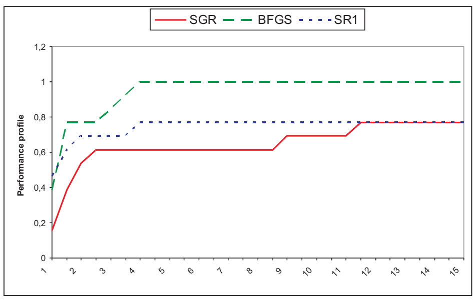


b. BFGS methods
(14 problems solved)

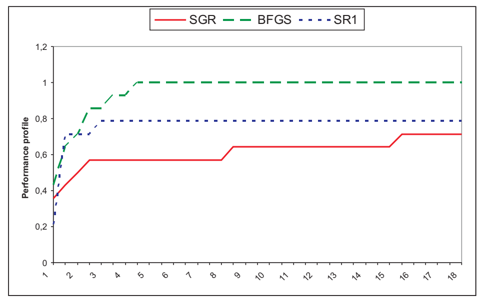


c. SR1 methods
(12 problems solved)

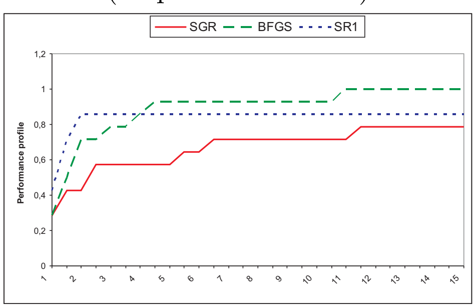
Figure 3: SGR, BFGS, SR1 methods at the noise level $\sigma = 0.1$



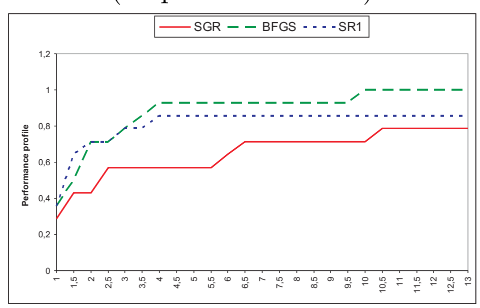
a. Line-search rule LS1
(13 problems solved)



b. Line-search rule LS2
(14 problems solved)

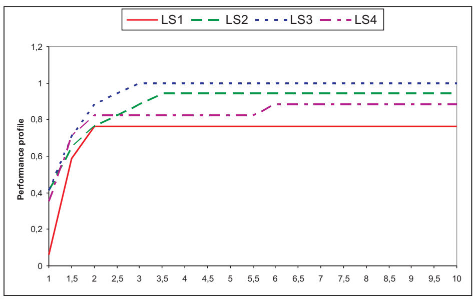


c. Line-search rule LS3
(14 problems solved)

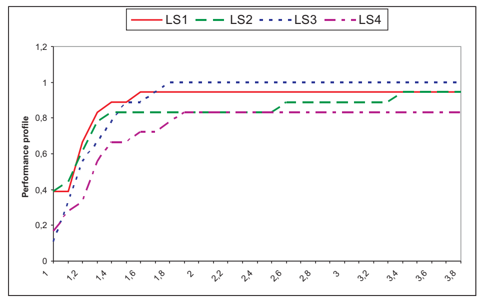


d. Line-search rule LS4
(14 problems solved)

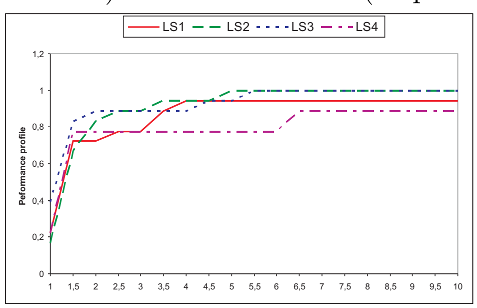
Figure 4: The line-search rules at the noise level $\sigma = 0.1$



a. SGR methods
(17 problems solved)

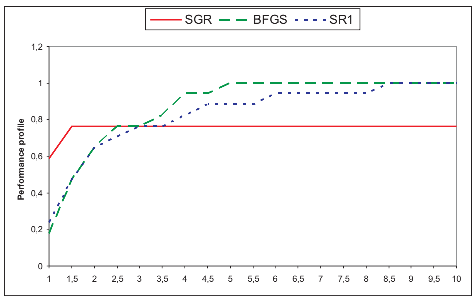


b. BFGS methods
(18 problems solved)

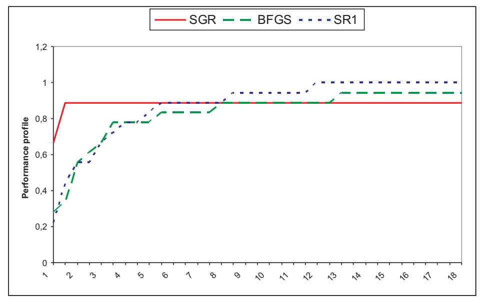


c. SR1 methods
(18 problems solved)

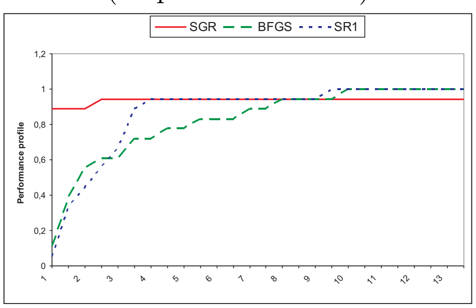
Figure 5: SGR, BFGS, SR1 methods at the noise level $\sigma = 0.5$



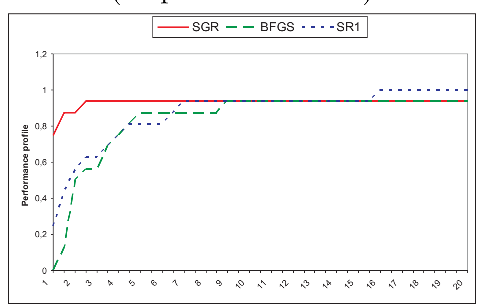
a. Line-search rule LS1
(17 problems solved)



b. Line-search rule LS2
(18 problems solved)

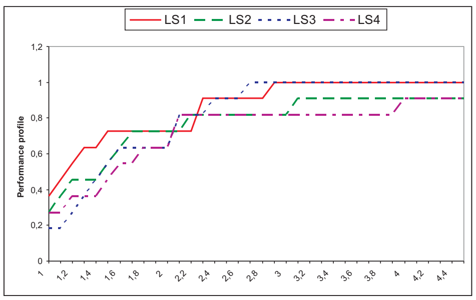


c. Line-search rule LS3
(18 problems solved)

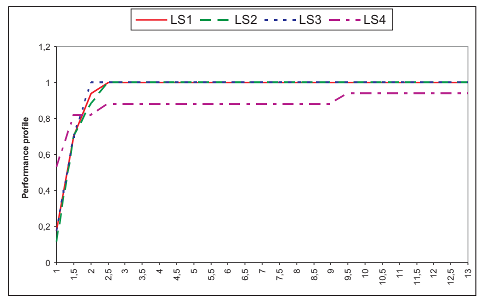


d. Line-search rule LS4
(16 problems solved)

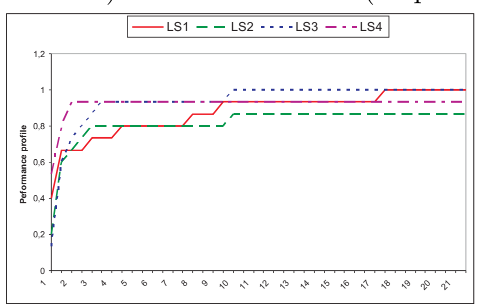
Figure 6: The line-search rules at the noise level $\sigma = 0.5$



a. SGR methods
(11 problems solved)

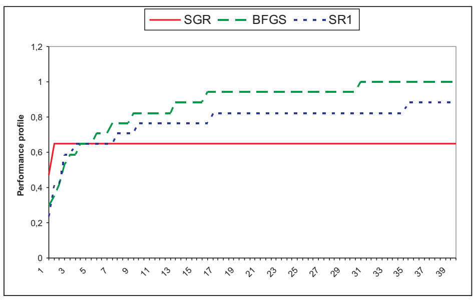


b. BFGS methods
(17 problems solved)

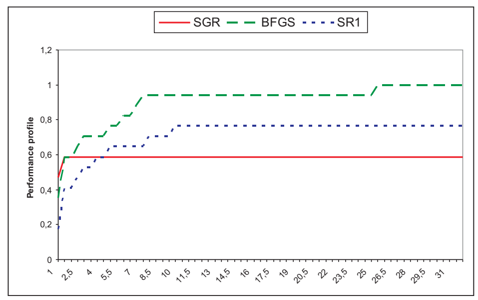


c. SR1 methods
(15 problems solved)

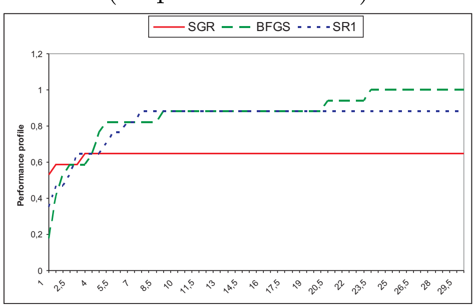
Figure 7: SGR, BFGS, SR1 methods at the noise level $\sigma = 1$



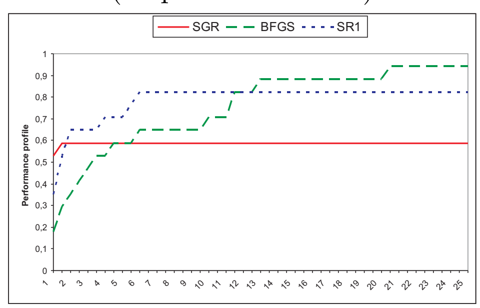
a. Line-search rule LS1
(17 problems solved)



b. Line-search rule LS2
(17 problems solved)

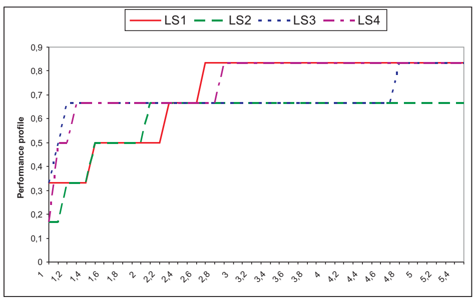


c. Line-search rule LS3
(17 problems solved)

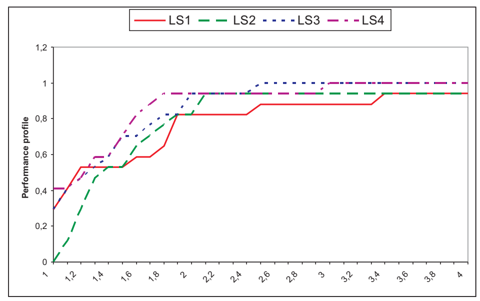


d. Line-search rule LS4
(17 problems solved)

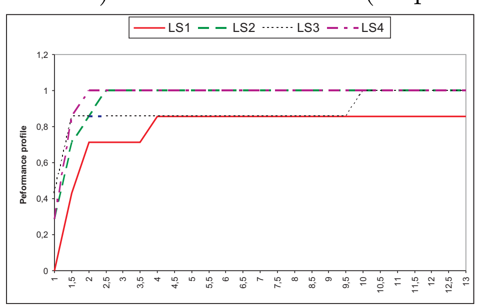
Figure 8: The line-search rules at the noise level $\sigma = 1$



a. SGR methods
(5 problems solved)

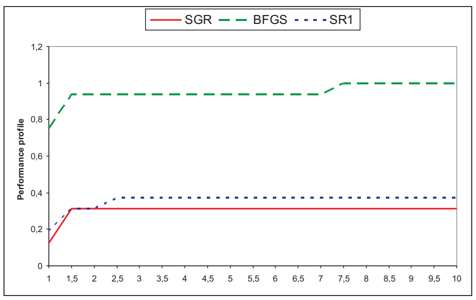


b. BFGS methods
(17 problems solved)

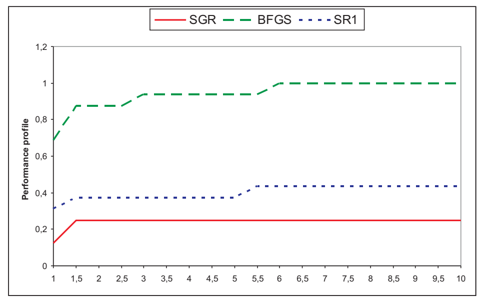


c. SR1 methods
(7 problems solved)

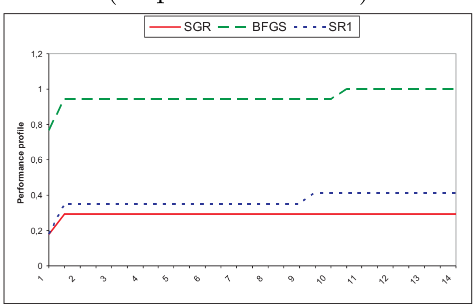
Figure 9: SGR, BFGS, SR1 methods at the noise level $\sigma = 5$



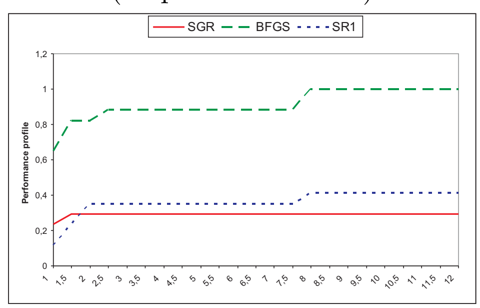
a. Line-search rule LS1
(16 problems solved)



b. Line-search rule LS2
(16 problems solved)

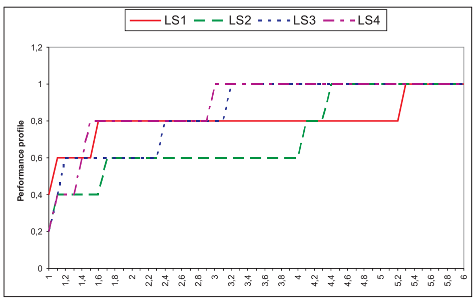


c. Line-search rule LS3
(17 problems solved)

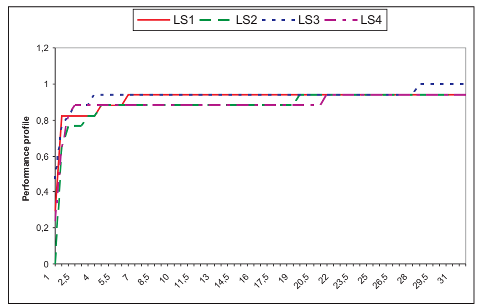


d. Line-search rule LS4
(17 problems solved)

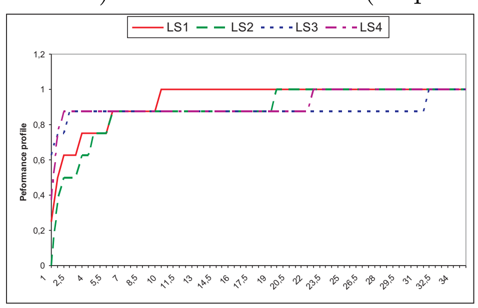
Figure 10: The line-search rules at the noise level $\sigma = 5$



a. SGR methods
(5 problems solved)

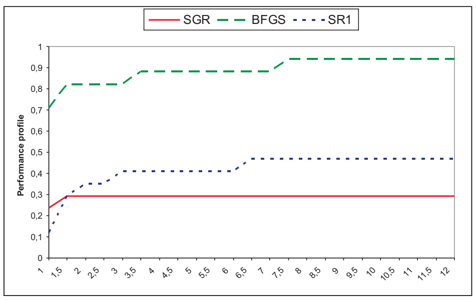


b. BFGS methods
(17 problems solved)

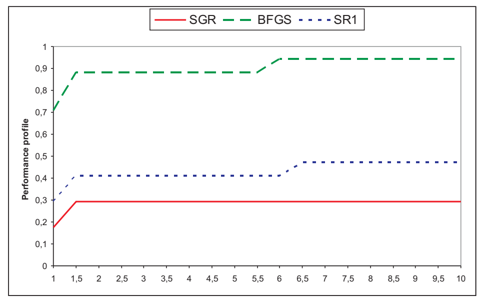


c. SR1 methods
(8 problems solved)

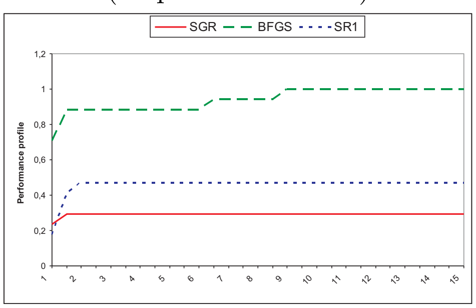
Figure 11: SGR, BFGS, SR1 methods at the noise level $\sigma = 10$



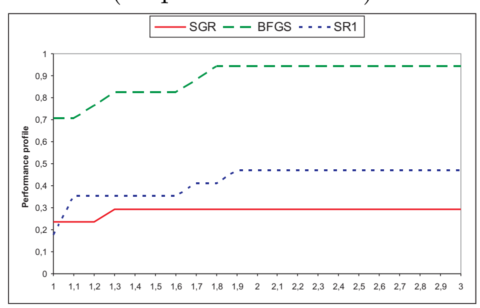
a. Line-search rule LS1
(17 problems solved)



b. Line-search rule LS2
(17 problems solved)



c. Line-search rule LS3
(17 problems solved)



d. Line-search rule LS4
(17 problems solved)

Figure 12: The line-search rules at the noise level $\sigma = 10$