*Proceedings of the V Congress of Mathematicians of Macedonia,*
*September 24–27, 2014, Ohrid, R. Macedonia*
Vol.2 (2015) 15–25

UDC: ???

# A LINE SEARCH METHOD WITH MEMORY FOR UNCONSTRAINED OPTIMIZATION OF NOISY FUNCTIONS

NATAŠA KREJIĆ [1], ZORANA LUŽANIN [2], FILIP NIKOLOVSKI [3],
AND IRENA STOJKOVSKA [4]

**Abstract.** We propose a new line search method for unconstrained optimization of noisy functions. The nonmonotone line search rule is based on Ulbrich's nonmonotone component [SIAM J. on Optimiz., 11 (4) (2001), 889–917]. The method uses only nosy functional values. Convergence under standard assumptions is established. Computational results show a good performance of the method compared with the monotone one.

## 1. Introduction

Let us consider the unconstrained minimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1.1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ has continuous partial derivatives. Assume that only noisy measurements $F(x)$ are available,

$$F(x) = f(x) + \delta(x), \tag{1.2}$$

at every $x \in \mathbb{R}^n$, where $\delta(x)$ represents the noise at $x$.

There are several approaches for solving the problem (1.1). One approach is to collect several function evaluations $F(x)$ at each value $x$ generated in the optimization process, then take the average of these values as an estimate for $f(x)$, Andradottir [1]. While this approach is well justified for a number of problems, it is not always possible to get an arbitrary number of function evaluation at the same point.

Another approach is the random search method along a random search direction $d_k$ that finds the new point $x_k + d_k$ such that the following inequality is satisfied

$$F(x_k + d_k) < F(x_k) - \tau_k, \tag{1.3}$$

where $\tau_k > 0$ is a threshold value. The drawback of this approach is that an inappropriate threshold value may results in rejecting many iterates, see [19].

---

2010 *Mathematics Subject Classification.* 90C56, 65K05.

*Key words and phrases.* Line search method, nonmonotone line search, unconstrained optimization, noisy function.

Among direct search methods for optimization in presence of noise is the coordinate search method, Lucidi and Sciandrone [15], that searches along a coordinate direction $d_k$ for a stepsize $\alpha_k$ that satisfies the inequality

$$F(x_k + \alpha_k d_k) < F(x_k) - \gamma \alpha_k^2, \tag{1.4}$$

where $\gamma > 0$. As smaller steps are allowed the method is more immune to the noise influence than the threshold approach. Some of the recent methods for optimization of noisy functions are considered in [2, 11, 12, 21, 23].

Nonmonotone line search strategies are a well developed class of methods for classical optimization. The dominant three nonmonotone rules are originally presented in Grippo et al. [9], Li and Fukushima [14] and Zhang and Hager [24]. All of these three strategies are successfully used for solving different problems in either derivative based or derivative-free methods, [3, 4, 7, 13, 17]. There are several important properties of nonmonotone line search methods. First of all, one can consider search directions which are not necessarily descent in all iterations, further more these methods are applicable even if the gradient is not available. And an additional property of nonmonotone line search rules that makes them attractive is the ability of converging to a global solution of problems with multiple local and global solutions. This property is reported in several papers, see for example [24].

There are several papers dealing with unconstrained optimization problems within line search framework and nonmonotone methods in particular, with results that are applicable on noisy problems even if the noise is not explicitly assumed, [4, 5, 7, 10, 15, 17]. In [10], authors proposed two nonmonotone line search strategies for solving the problem (1.1) using only functional noisy values (1.2). The line-search rules are of the following form

$$F(x_k + \alpha_k d_k) \leq \overline{F}_k + \eta_k - \alpha_k^2 \beta_k. \tag{1.5}$$

where $\alpha_k$ is the step size, $d_k$ is the search direction. The sequences $\{\eta_k\}$ and $\{\beta_k\}$ are sequences of positive numbers such that

$$\sum_{k=0}^{\infty} \eta_k = \eta < \infty, \tag{1.6}$$

while $\beta_k$ is bounded and

$$\lim_{k \in K} \beta_k = 0 \quad \Rightarrow \quad \lim_{k \in K} \nabla f(x_k) = 0 \tag{1.7}$$

for some infinite set of indices $K \subseteq \mathbb{N}$. In [10], the term $\overline{F}_k$ in the first line search strategy is defined by

$$\overline{F}_k = \max\{F(x_k), \ldots, F(x_{\max\{k-M+1,0\}})\}, \tag{1.8}$$

for an arbitrary but fixed $M \in \mathbb{N}$, and in the second line search strategy $\overline{F}_k$ is defined by

$$Q_{k+1} = r_k Q_k + 1, \; \overline{F}_{k+1} = \frac{r_k Q_k (\overline{F}_k + \eta_k) + F(x_{k+1})}{Q_{k+1}}, \tag{1.9}$$

with $r_k \in [r_{\min}, r_{\max}]$, $0 \leq r_{\min} \leq r_{\max} \leq 1$, $\overline{F}_0 = F(x_0)$ and $Q_0 = 1$.

In this paper, we propose a nonmonotone line-serch rule (1.5), where the term $\overline{F}_k$ is inspired by the Ulbrich's nonmonotone component, firstly used in the noise free trust region framework, [20], and lately used in noise free line-search methods, [17, 22]. Here we adopt it for noisy functional values. The convergence analysis of the proposed method relays on results from [20, 22] but extends them in the sense that the presented statements cover the case of noisy functional values. Some results presented relay on the results from [10]. In practical implementation we consider the BFGS search direction that requires approximations of gradient and Hessian in presence of noise, and we give a gradient approximation procedure in presence of noise. So, in Section 2 we present the model algorithm and analyze its convergence. In Section 3 we present the numerical results. Some final remarks are given in Section 4.

## 2. A new line-search strategy and convergence results

Let $\{F_j\}$ be a sequence of the last $m_k$ noisy functional values, where $m_k = \min\{k+1, M\}$, for fixed $M \in \mathbb{N}$. We define the term $\overline{F}_k$, in the line-search rule (1.5), as following,

$$\overline{F}_k = \max \left\{ F_k, \sum_{r=0}^{m_k-1} \lambda_{kr} F_{k-r} \right\}, \tag{2.1}$$

where $\lambda_{kr} \geq \lambda$, $r = 0, 1, \ldots, m_k - 1$ are scalars such that $\sum_{r=0}^{m_k-1} \lambda_{kr} = 1$, for fixed $\lambda \in (0, 1]$. This nonmonotone component (2.1) is inspired by the one proposed in [20] which uses noise free functional values in the trust region framework.

It is easy to see that

$$\overline{F}_k \geq F_k, \tag{2.2}$$

where $\overline{F}_k$ is defined by (2.1).

The method that we propose is described by the following algorithm.

**ALGORITHM.** Given the sequence $\{\eta_k\}$ such that (1.6) holds, the sequence $\{\beta_k\}$ such that (1.7) holds, an initial iterate $x_0 \in \mathbb{R}^n$ and $D > 0$.

**Step 1.:** Compute $d_k$ such that $\|d_k\| \leq D$.
**Step 2.:** Compute $\overline{F}_k$ according to (2.1).
**Step 3.:** Choose $\alpha_k$ such that (1.5) is satisfied.
**Step 4.:** Set $x_{k+1} = x_k + \alpha_k d_k$ and $k = k + 1$.

The positive sequence $\{\eta_k\}$ ensures that the the line search (1.5) is well-defined as $\alpha_k > 0$ exists for an arbitrary direction $d_k$.

For establishing the convergence of the proposed method, the following assumptions on the objective function and noise are made.

**A1:** The objective function $f \in C^1(\mathbb{R}^n)$ is bounded from below i.e. there exists $m$ such that $f(x) \geq m$ for all $x \in \mathbb{R}^n$

**A2:** The realized noise is bounded from above i.e there exists a constant $\Delta > 0$ such that for every iterate $x_k$

$$|\delta(x_k)| \leq \Delta. \tag{2.3}$$

The boundedness of noise stated in **A2** might look as a strong assumption at first, but we are interested only in the realized noise and thus **A2** is not a big obstacle in practical implementation of the algorithm. The same set of assumptions is used in [15].

Now we give the convergence analysis of the proposed method.

**Lemma 2.1.** *Let $\{x_k\}_{k\in\mathbb{N}}$ is a sequence generated by the Algorithm. Then, for every $k \in \mathbb{N}$,*

$$F_k \leq F_0 + \sum_{j=0}^{k-1} \eta_j - \lambda \sum_{j=0}^{k-2} \alpha_j^2 \beta_j - \alpha_{k-1}^2 \beta_{k-1} \leq F_0 + \sum_{j=0}^{k-1} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j.$$

*Proof.* We will prove the assertion by induction. For $k = 1$, since $0 < \lambda \leq 1$, from (1.5) and (1.6) it follows that

$$F_1 \leq F_0 + \eta_0 - \alpha_0^2 \beta_0 \leq F_0 + \eta_0 - \lambda \alpha_0^2 \beta_0.$$

Let us assume that the proposition is true for all $j$, $1 \leq j \leq k$. We discuss two cases.

*Case 1.* Let

$$\max\left\{F_k, \sum_{r=0}^{m_k-1} \lambda_{kr} F_{k-r}\right\} = F_k.$$

Then, we have

$$\begin{aligned}
F_{k+1} &\leq F_k + \eta_k - \alpha_k^2 \beta_k \leq \\
&\leq F_0 + \sum_{j=0}^{k-1} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j + \eta_k - \alpha_k^2 \beta_k \leq \\
&\leq F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j - \alpha_k^2 \beta_k \leq \\
&\leq F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j - \lambda \alpha_k^2 \beta_k = \\
&= F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k} \alpha_j^2 \beta_j.
\end{aligned}$$

*Case 2.* Let

$$\max\left\{F_k, \sum_{r=0}^{m_k-1} \lambda_{kr} F_{k-r}\right\} = \sum_{r=0}^{m_k-1} \lambda_{kr} F_{k-r},$$

and let $q = m_k - 1$. Then

$$F_{k+1} \leq \sum_{p=0}^{q} \lambda_{kp} F_{k-p} + \eta_k - \alpha_k^2 \beta_k. \tag{2.4}$$

From the inductive step we have

$$F_{k-p} \leq F_0 + \sum_{j=0}^{k-p-1} \eta_j - \lambda \sum_{j=0}^{k-p-2} \alpha_j^2 \beta_j - \alpha_{k-p-1}^2 \beta_{k-p-1}. \tag{2.5}$$

Substituting (2.5) into (2.4), we have

$$F_{k+1} \leq \sum_{p=0}^{q} \lambda_{kp} \left( F_0 + \sum_{j=0}^{k-p-1} \eta_j - \lambda \sum_{j=0}^{k-p-2} \alpha_j^2 \beta_j - \alpha_{k-p-1}^2 \beta_{k-p-1} \right) + \eta_k - \alpha_k^2 \beta_k \leq$$

$$\leq \left( \sum_{p=0}^{q} \lambda_{kp} \right) F_0 + \left( \sum_{p=0}^{q} \lambda_{kp} \right) \left( \sum_{j=0}^{k-p-1} \eta_j \right) -$$

$$- \lambda \left( \sum_{p=0}^{q} \lambda_{kp} \right) \sum_{j=0}^{k-p-2} \alpha_j^2 \beta_j - \sum_{p=0}^{q} \lambda_{kp} \alpha_{k-p-1}^2 \beta_{k-p-1} + \eta_k - \alpha_k^2 \beta_k \leq$$

$$\leq F_0 + \sum_{j=0}^{k-q-1} \eta_j - \lambda \sum_{j=0}^{k-q-2} \left( \sum_{p=0}^{q} \lambda_{kp} \right) \alpha_j^2 \beta_j -$$

$$- \sum_{p=0}^{q} \lambda_{kp} \alpha_{k-p-1}^2 \beta_{k-p-1} + \eta_k - \alpha_k^2 \beta_k \leq$$

$$\leq F_0 + \sum_{j=0}^{k-1} \eta_j - \lambda \sum_{j=0}^{k-q-2} \alpha_j^2 \beta_j - \lambda \sum_{p=0}^{q} \alpha_{k-p-1}^2 \beta_{k-p-1} + \eta_k - \alpha_k^2 \beta_k =$$

$$= F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k-q-2} \alpha_j^2 \beta_j - \lambda \sum_{p=k-q-1}^{k-1} \alpha_p^2 \beta_p - \alpha_k^2 \beta_k =$$

$$= F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j - \alpha_k^2 \beta_k \leq F_0 + \sum_{j=0}^{k} \eta_j - \lambda \sum_{j=0}^{k} \alpha_j^2 \beta_j,$$

which proves the lemma. $\square$

We need the result from Lemma 2.1 to prove the following theorem.

**Theorem 2.1.** *Let $\{x_k\}_{k \in \mathbb{N}}$ is a sequence of iterates generated by the Algorithm and let assumptions **A1** and **A2** hold. Then*

$$\lim_{k \to \infty} \alpha_k^2 \beta_k = 0.$$

*Proof.* From Lemma 2.1 and (1.6), for every $k \in \mathbb{N}$, we have

$$F_k \leq F_0 + \sum_{j=0}^{k-1} \eta_j - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j \leq F_0 + \eta - \lambda \sum_{j=0}^{k-1} \alpha_j^2 \beta_j.$$

From (1.2) and assumption **A2** we have

$$f_{k+1} \leq f_0 + \eta - \lambda \sum_{j=0}^{k} \alpha_j^2 \beta_j + 2\Delta.$$

Now, using assumption **A1** we have

$$\lambda \sum_{j=0}^{k} \alpha_j^2 \beta_j \leq f_0 + \eta + 2\Delta - f_{k+1} \leq f_0 + \eta + 2\Delta - m.$$

So,

$$\lambda \sum_{j=0}^{\infty} \alpha_j^2 \beta_j < \infty,$$

and because $\lambda > 0$ we conclude that

$$\lim_{k\to\infty} \alpha_k^2 \beta_k = 0.$$

□

Let us define a sequence $\{\delta_k\}$ with

$$\delta_k = \sup_{x \in B_k} |\delta(x)|,$$

where $B_k$ is the closed ball

$$B_k = \{x \in \mathbb{R}^n \mid \|x_k - x\| \leq D\},$$

and $D$ is the upper bound on the length of the search directions $d_k$.

Following two theorems that are valid for nonmonotone terms (1.8) and (1.9), see [10], can be also proved for the nonmonotone term (2.1), using Theorem 2.1 and the inequality (2.2).

**Theorem 2.2.** *Let $\{x_k\}$ be a sequence generated with the Algorithm and let the assumptions **A1** and **A2** hold. Let $(x^*, d)$ be a limit point of the sequence $\{(x_k, d_k)\}_{k \in K}$ where $K \subseteq \mathbb{N}$ is an infinite set od indices such that $\lim_{k \in K} \alpha_k^2 \beta_k = 0$. Additionally, let us assume that*

$$\lim_{k \in K} \frac{\delta_k}{\alpha_k} = 0.$$

*Then*

$$g(x^*)^T d \geq 0.$$

*Proof.* See proof of Theorem 3.3 in [10]. □

**Theorem 2.3.** *Assume that all conditions from Theorem 2.2 hold. Let $0 < \theta < 1$ and $0 < d < D < \infty$. Let us assume that the level set $\Omega = \{x \in \mathbb{R}^n \,\big|\, f(x) \leq f_0 + \eta + 2\Delta\}$ is bounded and that $K_1 \subseteq K$ is an infinite set of indices such that for any $k \in K_1$, there exists a search direction $d_k$ such that*

$$d \leq \|d_k\| \leq D \qquad \text{and} \qquad g(x_k)^T d_k \leq -\theta \|g(x_k)\| \|d_k\|. \tag{2.6}$$

*Then, for all $\varepsilon > 0$, there exists $k \in \mathbb{N}$ such that $\|g(x_k)\| \leq \varepsilon$.*

*Proof.* See proof of Theorem 3.4 in [10]. □

## 3. Numerical results

The algorithm has been tested on a set of 18 problems from the Moré-Garbow-Hillstrom collection [16]. All problems have objective function of the form $f(x) = \sum_{i=1}^{m} f_i^2(x)$. The test functions as well as the dimensions $n$ and the initial points $x_0$ are given in Appendix A.

The noisy measurements of the objective function are obtained with the simulated normally distributed noise $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ that is multiplied with the exact functional value to obtain $F(x) = f(x)(1+\varepsilon)$ at each point $x_k$ generated by the algorithm. Two different noise levels are tested $\sigma = 1$ and $\sigma = 10$.

The gradient approximation with centered differences is implemented using the noisy function values as follows. For a positive sequence $\{h_k\}$, the gradient of $f$ at $x_k$ is approximated with $\hat{g}_k$, given by

$$[\hat{g}_k]_j = \frac{F(x_k + h_k e_j) - F(x_k - h_k e_j)}{2h_k}, \; j = 1, 2, ..., n, \tag{3.1}$$

where $e_j$ is $j$th coordinate vector. The choice of $h_k$ is crucial for the approximation of the gradient in presence of noise, see [10]. We have established empirically that the choice $h_k = 3\sigma$, where $\sigma$ is the noise level, is appropriate for the test collection we considered.

For testing the algorithm proposed here, we choose the BFGS search direction $d_k$ of the form

$$d_k = -H_k \hat{g}_k,$$

where the inverse Hessian approximation is updated by the formula

$$H_{k+1} = \left(I - \rho s_k y_k^T\right) H_k \left(I - \rho y_k s_k^T\right) + \rho s_k s_k^T \tag{3.2}$$

for all $k = 0, 1, \ldots$, with $H_0 = I$, $s_k = x_{k+1} - x_k$, $y_k = \hat{g}_{k+1} - \hat{g}_k$, $\rho = 1/(y_k^T s_k)$ and rescaling the initial approximation $H_0$, see [18]. If the positive curvature condition $y_k^T s_k > 0$ is not satisfied, we set $H_{k+1} = H_k$.

We set $M = 4$, $\lambda = 0.01$ and we choose scalars $\lambda_{kr}$ as following: first we find the index $p$, $0 \leq p < m_k$, such that $F_{k-p} = \max\{F_{k-r} \mid 0 \leq r < m_k\}$ and set

$$\lambda_{kr} = \begin{cases} \lambda, & \text{if } r \neq p \\ 1 - (m_k - 1)\lambda, & \text{if } r = p \end{cases}$$

The sequences $\eta_k$ and $\beta_k$ in the line-search rule (1.5) are chosen as $\eta_k = |F(x_0)|/k^{1.1}$ and $\beta_k \equiv 1$, for all $k \in \mathbb{N}$.

The initial step length tested in Step 3 of Algorithm is $\alpha = 1$. If the line search rule (1.5) is not satisfied then a smaller step is computed using the safeguarded quadratic interpolation, [6].

We have compared the nonmonotone line-search rule (1.5) to the monotone one defined by

$$F(x_k + \alpha_k d_k) \leq F(x_k) - \alpha_k^2 \beta_k. \tag{3.3}$$

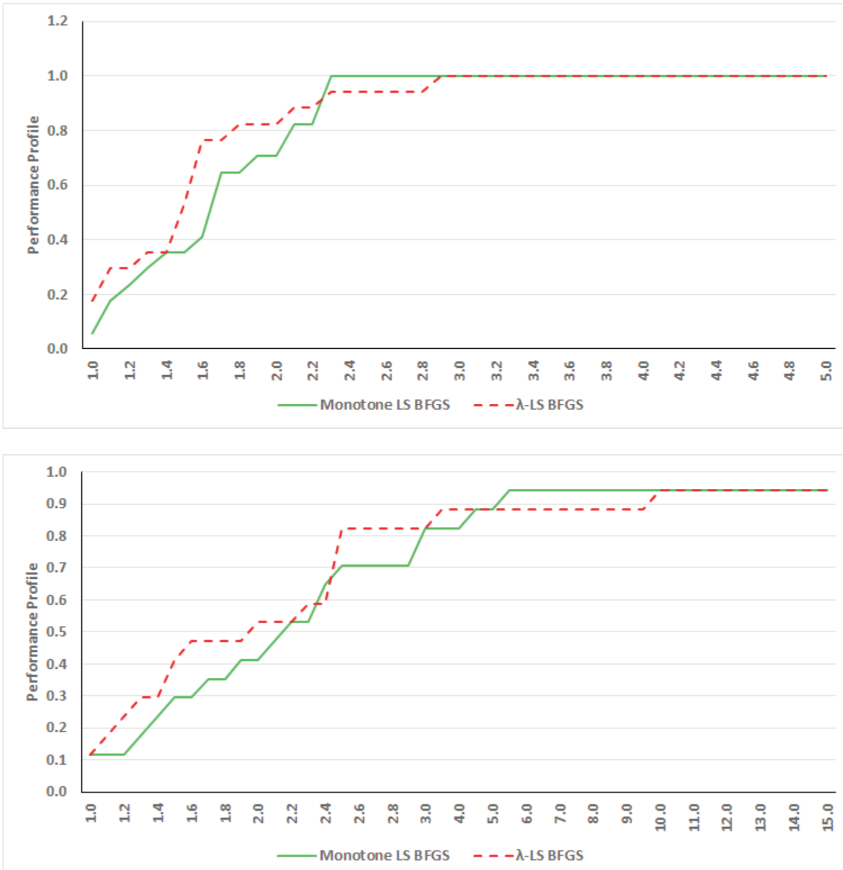


FIGURE 1. Performance profile for noise levels $\sigma = 1$ (top) and $\sigma = 10$ (bottom)

For each problem 50 independent test-runs have been made. We consider a test run successful if the stopping criterium

$$|F(x_k)| < (1 + 2\sigma) \cdot |F(x_0)| \cdot 10^{-3}, \tag{3.4}$$

is satisfied before exceeding the maximal number of $400 \cdot n$ function evaluations.

For each run we record the number of function evaluations, and we denote by $\varphi_{ij}$ the average number of function evaluations needed for the method $i$ to solve the problem $j$, in successful runs. To compare the performances of the nonmonotone and monotone method, we use the performance profiles defined in [8]. The performance measure that we use is

$$\pi_{ij} = 50 \cdot \varphi_{ij}/N_{ij},$$

where $N_{ij}$ is the number of successful runs out of the 50 for the method $i$ solving the problem $j$. The performance profiles are shown in Figure 1.

Based on results on Figure 1, we can conclude that the nonmonotone method we proposed is moderately better than the monotone one. This indicates that we have successfully implemented yet another nonmonotone and derivative-free method that performs better than the monotone method in presence of noise.

## 4. Conclusions

We defined a new nonmonotone line-search method for optimization in presence of noise. We analyzed the convergence of the proposed method and we tested it and compared it with the monotone one. Results show that the nonmonotone method is moderately better than the monotone one. There is some space for improvement however, if other parameter values are taken. For example, changing the value of $\lambda$, the choice of the scalars $\lambda_{kr}$ or the value of $M$ may result in a more efficient algorithm, and comparable to other nonmonotone methods. There is an opportunity this method to be used in a combination with stochastic approximation (SA), see for example [11, 12].

## References

[1] S. Andradottir, *A Scaled Stochastic Approximation Algorithm*, Management Science 42(4) (1996), 475–498.

[2] S. C. Billups, J. Larson, P. Graf, *Derivative-Free Optimization of Expensive Functions with Computational Error Using Weighted Regression,* SIAM J. Optim., Vol. 23, No. 1 (2013), 27–53.

[3] E. G. Birgin, N. Krejić, J. M. Martínez, *Globally convergent inexact quasi-Newton methods for solving nonlinear systems*, Numerical Algorithms 32, 2-4 (2003), 249–260.

[4] W. Cheng, D. H. Li, *A derivative-free nonmonotone line searh and its application to the spectral residual method,* IMA Journal of Numerical Analysis, Vol. 29 (2009), 814–825.

[5] W. L. Cruz, J. M. Martinez, M. Raydan, *Spectral residual method without gradient information for solving large-scale nonlinear systems of equations,* Math. Comput, Vol. 75, No. 255 (2006), 1429–1448.

[6] J. E. Dennis Jr., R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Philadelphia, SIAM, 1996.

[7] M. A. Diniz-Ehrhardt, J. M. Martínez, M. Raydan, *A derivative free nonmonotone line-search technique for unconstrained optimization,* J. Comp. Appl. Math., Vol. 219, No. 2 (2008), 383–397.

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

[9] L. Grippo, F. Lampariello, S. Lucidi, *A nonmonotone line search technique for Newton's method,* SIAM. J. Numer. Anal., Vol. 23 (1986), 707–716.

[10] N. Krejić, Z. Lužanin, F. Nikolovski, I. Stojkovska *A Nonmonotone Line Search Method for Noisy Minimization,* Optimization Letters, Vol. 9, No. 7 (2015), 1371-1391, DOI: 10.1007/s11590-015-0848-9.

[11] N. Krejić, Z. Lužanin, Z. Ovcin, I. Stojkovska *Descent direction method with line search for unconstrained optimization in noisy environment*, Optimization Methods and Software, Vol. 30, No. 6 (2015), 1164-1184, DOI: 10.1080/10556788.2015.1025403

[12] N. Krejić, Z. Lužanin, I. Stojkovska *A gradient method for unconstrained optimization in noisy environment,* Applied Numerical Mathematics, Vol. 70 (2013), 1–21.

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

[14] D. H. Li, M. Fukushima, *A derivative-free line search and global convergence of Broyden-like method for nonlinear equations,* Opt. Methods Software, Vol. 13 (2000), 181–201.

[15] S. Lucidi, M. Sciandrone, *A derivative-free algorithm for bound constrained optimization,* Computational Optimization and Applications, Vol. 21 (2002), 119–142.

[16] J. J. Moré, B. S. Garbow, K. E. Hillstrom, *Testing unconstrained optimization software*, ACM Transactions on Mathematical Software (1981), 7(1) 17–41.

[17] F. Nikolovski, I. Stojkovska, *New derivative-free nonmonotone line search methods for unconstrained minimization*, Proceedings of the Fifth International Scientific Conference - FMNS2013, Vol.1 Mathematics and Informatics (2013), 47–53.

[18] J. Nocedal, S. J. Wright, *Numerical Optimization. Second edition*, New York, Springer-Verlag, 2006.

[19] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control,* John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.

[20] M. Ulbrich, *Non-Monotone Trust-Region Methods for Bound-Constrained Semismooth Equations with Applications to Nonlinear Mixed Complementarity Problems*, SIAM J. on Optimiz., 11 (4) (2001), 889–917.

[21] Z. Xu, Y-H. Dai, *New stochastic approximation algorithms with adaptive step size,* Optim. Lett., Vol. 6, No. 8 (2012), 1831–1846.

[22] Z. Yu, D. Pu, *A new nonmonotone line search technique for unconstrained optimization,* J. Comp. Appl. Math., Vol. 219 (2008), 134–144.

[23] B. Zhang, Z. Zhu, S. Li, *A modified spectral conjugate gradient projection algorithm for total variation image restoration,* Appl. Math. Lett., Vol. 27 (2014), 26–35.

[24] H. Zhang, W. W. Hager, *A nonmonotone line search technique and its application to unconstrained optimization,* SIAM J. Optim., Vol. 14, No. 4 (2004), 1043–1056.

[1],[2] Department of Mathematics and Informatics,
Faculty of Science, University of Novi Sad,
Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia
*E-mail address*: `natasak@uns.ac.rs`
*E-mail address*: `zorana@dmi.uns.ac.rs`

[3],[4] Institute of Mathematics,
Faculty of Natural Sciences and Mathematics,
Ss. Cyril and Methodius University,
Arhimedova 3, 1000 Skopje, Republic of Macedonia
*E-mail address*: `filipnikolovski@gmail.com`
*E-mail address*: `irenatra@pmf.ukim.mk`

## Appendix A.

| 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