Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 3, pp. 489–505, 2020

# A LIU-STOREY CONJUGATE GRADIENT METHOD FOR SOLVING LARGE-SCALE NONLINEAR SYSTEM OF EQUATIONS WITH GLOBAL CONVERGENCE*

Min Li

**Abstract:** In this paper, a derivative-free method is developed to solve a large-scale nonlinear system of equations. The method is an extension of a descent Liu-Storey conjugate gradient method for solving unconstrained optimization problems. Under mild conditions, the global convergence of the proposed method is established with a nonmonotone line search. The method is suitable to large-scale problems for the low memory requirement. It is shown from the numerical results that the proposed method is effective in practical computation.

**Key words:** *Liu-Storey method, onlinear system of equations, nonmonotone line search, global convergence*

**Mathematics Subject Classification:** *65H10, 90C30*

---

## 1 Introduction

In this paper, we consider the nonlinear system of equations

$$g(x) = 0, \tag{1.1}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable nonlinear mapping. Given an initial point $x_0$, an iterative scheme for solving (1.1) generally generates a sequence of iterates $\{x_k\}$ by

$$x_{k+1} = x_k + \alpha_k d_k, \qquad k = 0, 1, \ldots, \tag{1.2}$$

where the steplength $\alpha_k$ is determined by a line search and $d_k$ is a search direction. Among numerous algorithms for solving (1.1), the most popular schemes are based on Newton's method or quasi-Newton methods [1,2,4,11,15,16,19,24,32,33]. These methods are attractive because of their locally fast convergence rates. However, they are typically unsuitable for large-scale problems because they need to solve a linear system of equations at each iteration using the Jacobian matrix or an approximation of it.

Recently, some spectral gradient methods have been developed to solve problem (1.1), see [5, 7, 8, 29], etc. La Cruz and Raydan [8] introduced the *Spectral Algorithm for Nonlinear Equations* (SANE) and established the global convergence with a nonmonotone line search,

*This work is supported by the NSF (11401242) of China, This work was supported in part by the Huaihua University Double First-Class initiative Applied Characteristic Discipline of Control Science and Engineering.

which is a modification of the Grippo-Lampariello-Lucidi (GLL) [12] line search. If $f(x)$ is a merit function such that $f(x) = 0$ if and only if $\|g(x)\| = 0$, then, a simplified GLL condition can be rewritten as follows:

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq \min\{k, M-1\}} f(x_{k-j}) + \delta \alpha_k \nabla f(x_k)^T d_k,$$

where $M$ is a positive integer, $0 < \delta < 1$ is a constant and $\nabla f(x_k)$ denotes the gradient of $f$ at $x_k$. If we set $M = 1$, then the nonmonotone line search above will reduce to the standard Armijo line search.

La Cruz, Martínez and Raydan [7] proposed the *Derivative-Free SANE* (DF-SANE) method to solve equation (1.1). They analyzed the global convergence of the method with a nonmonotone line search, which is based on the GLL [12] and Li-Fukushima (LF) [16] schemes. The LF line search requires the steplength $\alpha_k$ to satisfy

$$\|g(x_k + \alpha_k d_k)\| - \|g(x_k)\| \leq -\sigma_1 \alpha_k^2 \|d_k\|^2 + \epsilon_k \|g(x_k)\|, \tag{1.3}$$

where $\|\cdot\|$ denotes the Euclidean norm, $\sigma_1$ is a positive constant and $\{\epsilon_k\}$ is a positive sequence satisfying

$$\sum_{k=0}^{\infty} \epsilon_k < \infty. \tag{1.4}$$

It is clear that as $\alpha_k \to 0^+$, the left-hand side of (1.3) tends to zero, while the right-hand side goes to the positive constant $\epsilon_k \|g(x_k)\|$. Therefore, the line search is well defined.

Nonlinear conjugate gradient methods are welcome for unconstrained optimization problems for their low memory requirements and strong local and global convergence properties. Recently, some sufficient descent conjugate gradient methods were proposed, e.g., [13, 17, 20, 21, 26, 27, 30, 31]. In the survey paper, Narushima and Yabe [20] gave a comprehensive review of the development of different versions of descent conjugate gradient methods, with special attention given to their global convergence properties. Readers can refer to the paper for more details. Very recently, there are several studies of conjugate gradient methods for solving large-scale nonlinear system of equations [6, 25]. Yu [25] extended the Polak-Ribière-Polyak (PRP) [22, 23] method to solve problem (1.1). Based on the three-term conjugate gradient method by Zhang, Zhou and Li [30], Cheng, Xiao and Hu [6] developed a family of derivative-free methods for solving large-scale nonlinear system of equations. Inspired by these studies, in this paper, we attempt to extend the sufficient descent Liu-Storey method proposed by Li and Feng [17] to solve problem (1.1).

The paper is organized as follows. In the next section, we propose the method. In Section 3, we establish the global convergence of the proposed algorithm. Some numerical results are reported to test the efficiency of the method in the last section.

Throughout the paper, $\|\cdot\|$ denotes the Euclidean norm, $J(x)$ denotes the Jacobian matrix of $g(x)$, i.e., $J(x) = \nabla g(x)^T \in \mathbb{R}^{n \times n}$, $\mathbb{N}$ denotes the set of all nonnegative integers, i.e., $\mathbb{N} = \{0, 1, 2, \ldots\}$.

## 2 Algorithm

In this section, we first recall the nonlinear conjugate gradient method for solving the following unconstrained optimization problem

$$\min f(x), \quad x \in \mathbb{R}^n, \tag{2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and its gradient $\nabla f(x)$ is available. A nonlinear conjugate gradient method always generates a sequence $\{x_k\}$ by

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \ldots, \tag{2.2}$$

where the steplength $\alpha_k$ is determined by a line search and the search direction $d_k$ is defined by

$$d_k = \begin{cases} -\nabla f(x_0), & \text{if} \quad k = 0, \\ -\nabla f(x_k) + \beta_k d_{k-1}, & \text{if} \quad k \geq 1. \end{cases} \tag{2.3}$$

Here, $\beta_k$ is the conjugate gradient update parameter. Different choices for the scalar $\beta_k$ correspond to different conjugate gradient methods.

Recently, based on the Liu-Storey method [18], Li and Feng [17] proposed a sufficient descent conjugate gradient method (MLS method), in which the scalar $\beta_k$ is defined by

$$\beta_k^{\text{MLS}} = -\frac{\nabla f(x_k)^T(\nabla f(x_k) - \nabla f(x_{k-1}))}{\nabla f(x_{k-1})^T d_{k-1}} - t\frac{\|\nabla f(x_k) - \nabla f(x_{k-1})\|^2 \nabla f(x_k)^T d_{k-1}}{(\nabla f(x_{k-1})^T d_{k-1})^2}, \tag{2.4}$$

where $t > 1/4$ is a constant. They proved that the method can generate sufficient descent directions for the objective function and established the global convergence with the strong Wolfe line search. The reported numerical results showed that the MLS method is efficient for the unconstrained optimization problems in the CUTEr library [3].

Now, we attempt to extend the MLS method [17] to solve nonlinear system of equations. If $g(x)$ is the gradient of $f(x)$, then equation $g(x) = 0$ is the first order necessary optimality condition of unconstrained optimization problem (2.1). Therefore, based on the MLS method and the GLL [12] and LF [16] nonmonotone line search methods, we design the following algorithm to solve problem (1.1).

**Algorithm 2.1.** (Derivative-Free MLS method: DF-MLS method)

**Step 0.** Given an initial point $x_0 \in \mathbb{R}^n$ and a positive integer $M$. Let $0 < \rho < 1$, $\lambda_1, \lambda_2, \lambda_3 > 0$ and $0 < \alpha_{min} < \alpha_{max}$ be given positive constants. Select a positive sequence $\{\eta_k\}$ such that

$$\sum_{k=0}^{\infty} \eta_k < \infty. \tag{2.5}$$

Set $k = 0$.

**Step 1.** Stop if $\|g(x_k)\| = 0$.

**Step 2.** Compute $d_k$ by

$$d_k = \begin{cases} -g_0, & \text{if} \quad k = 0, \\ -g_k + \beta_k^{\text{MLS}} d_{k-1}, & \text{if} \quad k \geq 1, \end{cases} \tag{2.6}$$

$$\beta_k^{\text{MLS}} = -\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}} - t\frac{\|y_{k-1}\|^2 g_k^T d_{k-1}}{(g_{k-1}^T d_{k-1})^2}, \tag{2.7}$$

where $t > \frac{1}{4}$ is a constant, $g_k = g(x_k)$ and $y_{k-1} = g_k - g_{k-1}$.

**Step 3.** Choose an initial steplength $\alpha_{0,k} \in [\alpha_{min}, \alpha_{max}]$, and set $\alpha_k = \alpha_{0,k}$.

**Step 4.** Nonmonotone line search.
If

$$\|g(x_k+\alpha_k d_k)\|^2 \leq \max_{0\leq j\leq \min\{k,M-1\}} \|g(x_{k-j})\|^2-\lambda_1\alpha_k^2\|d_k\|^2-\lambda_2\alpha_k^2\|d_k\|^4-\lambda_3\alpha_k^2\|g_k\|^2+\eta_k, \tag{2.8}$$

then set $x_{k+1} = x_k + \alpha_k d_k$ and goto Step 5.
Else if

$$\|g(x_k-\alpha_k d_k)\|^2 \leq \max_{0\leq j\leq \min\{k,M-1\}} \|g(x_{k-j})\|^2-\lambda_1\alpha_k^2\|d_k\|^2-\lambda_2\alpha_k^2\|d_k\|^4-\lambda_3\alpha_k^2\|g_k\|^2+\eta_k, \tag{2.9}$$

then set $x_{k+1} = x_k - \alpha_k d_k$ and goto Step 5.
Else set $\alpha_k = \rho\alpha_k$ and goto Step 4.

**Step 5.** Set $k = k + 1$ and goto Step 1.

**Remark.** (I) The formula for $\beta_k$ is derived from (2.4). From the definition of $d_k$, we have that if $g_k \neq 0$ and $g_{k-1}^T d_{k-1} \neq 0$ for some $k \geq 1$, then

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \beta_k^{\mathrm{MLS}} g_k^T d_{k-1} \\ &= -\|g_k\|^2 - \left(\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}} + t\frac{\|y_{k-1}\|^2 g_k^T d_{k-1}}{(g_{k-1}^T d_{k-1})^2}\right) g_k^T d_{k-1} \\ &= \frac{-\|g_k\|^2(g_{k-1}^T d_{k-1})^2 - (g_k^T y_{k-1})(g_k^T d_{k-1})(g_{k-1}^T d_{k-1}) - t\|y_{k-1}\|^2(g_k^T d_{k-1})^2}{(g_{k-1}^T d_{k-1})^2} \\ &\leq \frac{-\|g_k\|^2(g_{k-1}^T d_{k-1})^2 + \frac{1}{2}\frac{1}{2t}\|g_k\|^2(g_{k-1}^T d_{k-1})^2 + \frac{1}{2}2t\|y_{k-1}\|^2(g_k^T d_{k-1})^2}{(g_{k-1}^T d_{k-1})^2} \\ &\quad - \frac{t\|y_{k-1}\|^2(g_k^T d_{k-1})^2}{(g_{k-1}^T d_{k-1})^2} \\ &= \left(\frac{1}{4t} - 1\right)\|g_k\|^2, \end{aligned} \tag{2.10}$$

where the inequality follows from an upper bound for the second term of the third equation, which is obtained from $u^T v \leq \frac{1}{2}(\|u\|^2 + \|v\|^2)$ with $u = \frac{-1}{\sqrt{2t}}(g_{k-1}^T d_{k-1})g_k$ and $v = \sqrt{2t}(g_k^T d_{k-1})y_{k-1}$. Obviously, inequality (2.10) also holds for $k = 0$. Therefore, the direction generated at Step 2 will always satisfy

$$g_k^T d_k \leq \left(\frac{1}{4t} - 1\right)\|g_k\|^2 < 0. \tag{2.11}$$

This means that the scalar $\beta_k^{\mathrm{MLS}}$ in (2.7) is well defined as long as $g_{k-1} \neq 0$ for $k \geq 1$, since $g_{k-1}^T d_{k-1} \neq 0$ by (2.11). In this paper, we focus on extending the MLS method [17] to solve nonlinear system of equations not only because the method is efficient, but also because we can establish the convergence of the DF-MLS method in a simple way.

(II) At Step 4, the nonmonotone line search is a modification of the GLL [12] and LF [16] schemes. Since $d_k$ is not necessarily a descent direction of $\|g(x)\|^2$ at $x_k$, we determine the next iteration by two inequalities in the line search step. Noting that $\eta_k > 0$, (2.8) or (2.9) will be satisfied when $\alpha_k$ is small enough. Therefore, the algorithm is well defined. For convenience, we call this algorithm as DF-MLS method.

## $\boxed{3}$ Convergence Property

This section is devoted to the global convergence of the DF-MLS method. In the remainder of this paper, we always assume that the sequence $\{x_k\}$ is generated by the DF-MLS method and the finite termination never occurs, i.e., $g_k \neq 0$ for all $k \geq 0$, which implies that the generated $\{x_k\}$ is an infinite sequence. Moreover, we also assume that the function $g$ satisfies the following assumptions.

**Assumption 3.1.**

(i) The level set $\Omega := \{x \in \mathbb{R}^n \mid \|g(x)\| \leq \sqrt{\|g(x_0)\|^2 + \eta}\}$ is bounded, where $x_0 \in \mathbb{R}^n$ is an arbitrary initial point, and $\eta$ is a positive constant such that

$$\sum_{k=0}^{\infty} \eta_k \leq \eta. \tag{3.1}$$

(ii) The function $g(x)$ is continuously differentiable on some neighborhood $\Gamma$ of $\Omega$. Hence, $g(x)$ is Lipschitz continuous on $\Gamma$. That is, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in \Gamma. \tag{3.2}$$

From the definition of $\Omega$, it is obvious that, for $\gamma_1 = \sqrt{\|g(x_0)\|^2 + \eta} > 0$,

$$\|g(x)\| \leq \gamma_1, \quad \forall x \in \Omega. \tag{3.3}$$

Before we proceed with the convergence analysis, we give some preliminary definitions. Define $V_0 = \|g(x_0)\|^2$ and

$$V_k = \max_{0 \leq j \leq \min\{k, M-1\}} \|g(x_{k-j})\|^2, \quad \forall k = 1, 2, \ldots.$$

Let $v(k) \in \{k - \min\{k, M-1\}, \ldots, k\}$ be such that

$$\|g(x_{v(k)})\|^2 = V_k, \quad \forall k = 1, 2, \ldots.$$

**Lemma 3.1.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. Then $\{x_k\} \subset \Omega$.*

*Proof.* It follows from the line search step of the algorithm that, for any $k \geq 0$,

$$\begin{aligned} \|g(x_{k+1})\|^2 &\leq V_k - \lambda_1 \alpha_k^2 \|d_k\|^2 - \lambda_2 \alpha_k^2 \|d_k\|^4 - \lambda_3 \alpha_k^2 \|g_k\|^2 + \eta_k \\ &\leq V_k + \eta_k. \end{aligned} \tag{3.4}$$

From the definition of $V_k$, we have, for any $k \geq 0$,

$$\begin{aligned} V_{k+1} &= \max_{0 \leq j \leq \min\{k+1, M-1\}} \|g(x_{k+1-j})\|^2 \\ &= \max_{0 \leq j \leq \min\{k, M-2\}+1} \|g(x_{k+1-j})\|^2 \\ &\leq \max_{0 \leq j \leq \min\{k, M-1\}+1} \|g(x_{k+1-j})\|^2 \\ &= \max\{V_k, \|g(x_{k+1})\|^2\} \\ &\leq V_k + \eta_k. \end{aligned} \tag{3.5}$$

Clearly, $x_0 \in \Omega$. By (3.4) and (3.5), we have, for all $k \geq 0$,

$$\begin{aligned}\|g(x_{k+1})\|^2 &\leq V_k + \eta_k \\ &\leq V_{k-1} + \eta_{k-1} + \eta_k \\ &\leq \dots \\ &\leq V_0 + \sum_{j=0}^{k} \eta_j \\ &\leq \|g_0\|^2 + \eta.\end{aligned}$$

Thus, $x_{k+1} \in \Omega$, which completes the proof. □

Note that if $k \geq M - 1$, then $k - M + 1 \leq v(k) \leq k$, so that $\lim_{k\to\infty} v(k) = \infty$. Then we have the following lemma.

**Lemma 3.2.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. Then*

$$\begin{cases} \lim_{k\to\infty} \alpha_{v(k)-1}\|d_{v(k)-1}\| = 0, \\ \lim_{k\to\infty} \alpha_{v(k)-1}\|d_{v(k)-1}\|^2 = 0, \\ \lim_{k\to\infty} \alpha_{v(k)-1}\|g_{v(k)-1}\| = 0. \end{cases} \tag{3.6}$$

*Proof.* It follows from (3.5) that for all $k \geq 0$,

$$V_{k+1} \leq V_k + \eta_k \leq (1 + \eta_k)V_k + \eta_k. \tag{3.7}$$

We get from Lemma 3.3 in [9] that the sequence $\{V_k\}$ is convergent. Moreover, we get from (3.4) that

$$\begin{aligned}V_k = \|g(x_{v(k)})\|^2 \leq V_{v(k)-1} &- \lambda_1\alpha_{v(k)-1}^2\|d_{v(k)-1}\|^2 - \lambda_2\alpha_{v(k)-1}^2\|d_{v(k)-1}\|^4 \\ &- \lambda_3\alpha_{v(k)-1}^2\|g_{v(k)-1}\|^2 + \eta_{v(k)-1}.\end{aligned} \tag{3.8}$$

This inequality implies

$$\begin{aligned}&\lambda_1\alpha_{v(k)-1}^2\|d_{v(k)-1}\|^2 + \lambda_2\alpha_{v(k)-1}^2\|d_{v(k)-1}\|^4 + \lambda_3\alpha_{v(k)-1}^2\|g_{v(k)-1}\|^2 \\ &\leq V_{v(k)-1} - V_k + \eta_{v(k)-1}.\end{aligned} \tag{3.9}$$

Taking limits for $k \to \infty$, we will obtain (3.6) since the sequence $\{V_k\}$ is convergent and $\lim_{k\to\infty} \eta_k = 0$. □

Based on the result given by Grippo, Lampariello and Lucidi [12], we get the following lemma, which shows that the infinite sequence $\{\|g(x_k)\|\}$ is convergent.

**Lemma 3.3.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. Then the sequence $\{\|g(x_k)\|\}$ is convergent. Furthermore,*

$$\begin{cases} \lim_{k\to\infty} \alpha_k\|d_k\| = 0, \\ \lim_{k\to\infty} \alpha_k\|d_k\|^2 = 0, \\ \lim_{k\to\infty} \alpha_k\|g_k\| = 0. \end{cases} \tag{3.10}$$

*Proof.* Here and in the sequel we assume that the iteration index $k$ is large enough and $k \geq j$ for a given $j \geq 1$. It follows from the Assumption 3.1 that $\|g(x)\|$ is uniformly continuous on the level set $\Omega$. Let $\overline{v}(k) = v(k + M)$. Note that

$$1 \leq k - j + 1 \leq \overline{v}(k) - j \leq k - j + M \tag{3.11}$$

since $k-M+1 \leq v(k) \leq k$ and $k \geq j$. First, we will prove by induction that, for any given $j \geq 1$,

$$\lim_{k\to\infty} \|x_{\overline{v}(k)-j+1} - x_{\overline{v}(k)-j}\| = 0 \tag{3.12}$$

and

$$\lim_{k\to\infty} \|g(x_{\overline{v}(k)-j})\| = \lim_{k\to\infty} \|g(x_{v(k)})\|. \tag{3.13}$$

It follows from (3.6) that $\lim_{k\to\infty} \|x_{v(k)} - x_{v(k)-1}\| = 0$. Therefore, (3.12) holds for $j=1$ since $\{x_{\overline{v}(k)}\} \subset \{x_{v(k)}\}$. Using the uniform continuity of $\|g(x)\|$ on the level set $\Omega$ and the fact that the sequence $\{\|g(x_{v(k)})\|\}$ is convergent from the proof of Lemma 3.2, we have that (3.13) is satisfied for $j=1$. Assume now that (3.12) and (3.13) hold for a given $j$ . Then we get from (3.4) that

$$\begin{aligned}\|g(x_{\overline{v}(k)-j})\|^2 \leq \|g(x_{v(\overline{v}(k)-j-1)})\|^2 - \lambda_1\alpha^2_{\overline{v}(k)-j-1}\|d_{\overline{v}(k)-j-1}\|^2 - \lambda_2\alpha^2_{\overline{v}(k)-j-1}\|d_{\overline{v}(k)-j-1}\|^4 \\ - \lambda_3\alpha^2_{\overline{v}(k)-j-1}\|g_{\overline{v}(k)-j-1}\|^2 + \eta_{\overline{v}(k)-j-1}.\end{aligned}$$

This implies

$$\lambda_1\alpha^2_{\overline{v}(k)-j-1}\|d_{\overline{v}(k)-j-1}\|^2 \leq \|g(x_{v(\overline{v}(k)-j-1)})\|^2 - \|g(x_{\overline{v}(k)-j})\|^2 + \eta_{\overline{v}(k)-j-1}.$$

Taking limits for $k \to \infty$, it follows from the assumption (3.13) that

$$\lim_{k\to\infty} \|x_{\overline{v}(k)-j} - x_{\overline{v}(k)-j-1}\| = \lim_{k\to\infty} \alpha_{\overline{v}(k)-j-1}\|d_{\overline{v}(k)-j-1}\| = 0. \tag{3.14}$$

Since $\|g(x)\|$ is uniformly continuous on the level set $\Omega$, we get from the assumption (3.13) that

$$\lim_{k\to\infty} \|g_{\overline{v}(k)-j-1}\| = \lim_{k\to\infty} \|g_{\overline{v}(k)-j}\| = \lim_{k\to\infty} \|g_{v(k)}\|. \tag{3.15}$$

Hence, (3.12) and (3.13) hold for any given $j \geq 1$.

Moreover, for every $k$,

$$\begin{aligned} x_{\overline{v}(k)} &= x_k + (x_{k+1} - x_k) + \cdots + (x_{\overline{v}(k)} - x_{\overline{v}(k)-1}) \\ &= x_k + \sum_{j=1}^{\overline{v}(k)-k} (x_{\overline{v}(k)-j+1} - x_{\overline{v}(k)-j}). \end{aligned} \tag{3.16}$$

Noting from (3.11) for $j=k$ that $1 \leq \overline{v}(k) - k \leq M$, we get from (3.12) and (3.16) that

$$\lim_{k\to\infty} \|x_{\overline{v}(k)} - x_k\| = 0. \tag{3.17}$$

Therefore, based on (3.17) and the uniform continuity of $\|g(x)\|$, we conclude that the sequence $\{\|g(x_k)\|\}$ is convergent and

$$\lim_{k\to\infty} \|g(x_k)\| = \lim_{k\to\infty} \|g(x_{v(k)})\|. \tag{3.18}$$

Combining this with (3.4), we get (3.10) and complete the proof. □

The following lemma shows that the sequence of directions $\{d_k\}$ is bounded.

**Lemma 3.4.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. If there exists a constant $\gamma > 0$ such that*

$$\|g_k\| \geq \gamma, \quad \forall k \geq 0, \tag{3.19}$$

*then there exists a constant $B > 0$ such that*

$$\|d_k\| \le B, \quad \forall k \ge 0 \tag{3.20}$$

*and*

$$\lim_{k\to\infty} \beta_k^{\mathrm{MLS}} \|d_{k-1}\| = 0. \tag{3.21}$$

*Proof.* From (2.6), (2.11), (3.2), (3.3), (3.10) and (3.19), we have

$$\begin{aligned}
\|d_k\| &\le \|g_k\| + |\beta_k^{\mathrm{MLS}}|\|d_{k-1}\| \\
&\le \|g_k\| + \frac{\|g_k\|\|y_{k-1}\|\|d_{k-1}\|}{|g_{k-1}^T d_{k-1}|} + t\frac{\|y_{k-1}\|^2\|g_k\|\|d_{k-1}\|}{|g_{k-1}^T d_{k-1}|^2}\|d_{k-1}\| \\
&\le \gamma_1 + \frac{4t\gamma_1 L\alpha_{k-1}\|d_{k-1}\|^2}{(4t-1)\gamma^2} + \frac{t(4t)^2 2\gamma_1^2 L\alpha_{k-1}\|d_{k-1}\|^2}{(4t-1)^2\gamma^4}\|d_{k-1}\| \\
&= \gamma_1 + \frac{4t\gamma_1 L}{(4t-1)\gamma^2}\alpha_{k-1}\|d_{k-1}\|^2 + \frac{32t^3\gamma_1^2 L}{(4t-1)^2\gamma^4}\alpha_{k-1}\|d_{k-1}\|^2\|d_{k-1}\|.
\end{aligned} \tag{3.22}$$

We get from (3.10) that, for any constant $b \in (0,1)$, there exists an index $k_0$ such that

$$\frac{32t^3\gamma_1^2 L}{(4t-1)^2\gamma^4}\alpha_{k-1}\|d_{k-1}\|^2 < b, \quad \forall k > k_0. \tag{3.23}$$

Then

$$\|d_k\| \le \gamma_1 + \frac{(4t-1)\gamma^2 b}{8t^2\gamma_1} + b\|d_{k-1}\| = c + b\|d_{k-1}\|, \tag{3.24}$$

where $c = \gamma_1 + (4t-1)\gamma^2 b/(8t^2\gamma_1)$ is a constant. For any $k > k_0$ we have

$$\|d_k\| \le c(1 + b + b^2 + \cdots + b^{k-k_0+1}) + b^{k-k_0}\|d_{k_0}\| \le \frac{c}{1-b} + \|d_{k_0}\|. \tag{3.25}$$

Therefore, (3.20) holds with $B = \max\{\|d_1\|, \|d_2\|, \cdots, \|d_{k_0}\|, \frac{c}{1-b} + \|d_{k_0}\|\}$.

By (3.2), (3.19) and (3.20), we have

$$\begin{aligned}
|\beta_k^{\mathrm{MLS}}|\|d_{k-1}\| &\le \frac{\|g_k\|\|y_{k-1}\|\|d_{k-1}\|}{|g_{k-1}^T d_{k-1}|} + t\frac{\|y_{k-1}\|^2\|g_k\|\|d_{k-1}\|}{|g_{k-1}^T d_{k-1}|^2}\|d_{k-1}\| \\
&\le \left(\frac{4t\gamma_1 L}{(4t-1)\gamma^2} + \frac{32t^3\gamma_1^2 L}{(4t-1)^2\gamma^4}\|d_{k-1}\|\right)\alpha_{k-1}\|d_{k-1}\|^2.
\end{aligned} \tag{3.26}$$

Combining this with (3.10) gives (3.21). □

The following theorem establishes the global convergence of Algorithm 2.1. It is similar to the Theorem 1 in reference [7].

**Theorem 3.5.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. Then we have*

$$\lim_{k\to\infty} \|g_k\| = 0, \tag{3.27}$$

*or every limit point $x^*$ of $\{x_k\}$ satisfies*

$$g(x^*)^T J(x^*) g(x^*) = 0. \tag{3.28}$$

*Proof.* From Lemma 3.3, it follows that $\lim_{k\to\infty}\|g_k\| = \zeta \geq 0$. We only need to prove the case of $\zeta > 0$. Then, since by (3.10) $\lim_{k\to\infty}\alpha_k\|g_k\| = 0$, we have $\lim_{k\to\infty}\alpha_k = 0$. It follows from the line search step of the algorithm that, when $k$ is sufficiently large, $\rho^{-1}\alpha_k$ satisfies neither (2.8) nor (2.9). So we have,

$$\begin{aligned}
&\|g(x_k+\rho^{-1}\alpha_k d_k)\|^2 - \|g_k\|^2 \\
\geq&\|g(x_k+\rho^{-1}\alpha_k d_k)\|^2 - \max_{0\leq j\leq \min\{k,M-1\}}\|g(x_{k-j})\|^2 \\
>&-\lambda_1\rho^{-2}\alpha_k^2\|d_k\|^2 - \lambda_2\rho^{-2}\alpha_k^2\|d_k\|^4 - \lambda_3\rho^{-2}\alpha_k^2\|g_k\|^2 + \eta_k \\
>&-\lambda_1\rho^{-2}\alpha_k^2\|d_k\|^2 - \lambda_2\rho^{-2}\alpha_k^2\|d_k\|^4 - \lambda_3\rho^{-2}\alpha_k^2\|g_k\|^2
\end{aligned} \tag{3.29}$$

and

$$\begin{aligned}
&\|g(x_k-\rho^{-1}\alpha_k d_k)\|^2 - \|g_k\|^2 \\
\geq&\|g(x_k-\rho^{-1}\alpha_k d_k)\|^2 - \max_{0\leq j\leq \min\{k,M-1\}}\|g(x_{k-j})\|^2 \\
>&-\lambda_1\rho^{-2}\alpha_k^2\|d_k\|^2 - \lambda_2\rho^{-2}\alpha_k^2\|d_k\|^4 - \lambda_3\rho^{-2}\alpha_k^2\|g_k\|^2 + \eta_k \\
>&-\lambda_1\rho^{-2}\alpha_k^2\|d_k\|^2 - \lambda_2\rho^{-2}\alpha_k^2\|d_k\|^4 - \lambda_3\rho^{-2}\alpha_k^2\|g_k\|^2.
\end{aligned} \tag{3.30}$$

Using (3.3) and (3.20), we have

$$\|g(x_k+\rho^{-1}\alpha_k d_k)\|^2 - \|g_k\|^2 > -C\alpha_k^2 \tag{3.31}$$

and

$$\|g(x_k-\rho^{-1}\alpha_k d_k)\|^2 - \|g_k\|^2 > -C\alpha_k^2, \tag{3.32}$$

where $C = \lambda_1\rho^{-2}B^2 + \lambda_2\rho^{-2}B^4 + \lambda_3\rho^{-2}\gamma_1^2$.
By (3.31),

$$\frac{\|g(x_k+\rho^{-1}\alpha_k d_k)\|^2 - \|g_k\|^2}{\alpha_k} > -C\alpha_k. \tag{3.33}$$

By the mean-value theorem and (3.33), there exists a $\xi_k \in (0,1)$ such that

$$2\rho^{-1}g(x_k+\xi_k\rho^{-1}\alpha_k d_k)^T J(x_k+\xi_k\rho^{-1}\alpha_k d_k)d_k > -C\alpha_k. \tag{3.34}$$

Namely,

$$2\rho^{-1}g(x_k+\xi_k\rho^{-1}\alpha_k d_k)^T J(x_k+\xi_k\rho^{-1}\alpha_k d_k)(-g_k+\beta_k^{\mathrm{MLS}}d_{k-1}) > -C\alpha_k. \tag{3.35}$$

Combining this with (3.10) and (3.21) and taking limits in (3.35) give

$$g(x^*)^T J(x^*)g(x^*) \leq 0. \tag{3.36}$$

Using (3.32) and proceeding in the same way, we obtain

$$g(x^*)^T J(x^*)g(x^*) \geq 0. \tag{3.37}$$

The last two inequalities imply (3.28). The proof is complete. □

By Theorem 3.5 and the continuity of $\|g(x)\|$, we immediately have the following corollary.

**Corollary 3.6.** *Suppose that $\{x_k\}$ is an infinite sequence generated by the DF-MLS method and Assumption 3.1 holds. Suppose also that $x^*$ is a limit point of $\{x_k\}$ and*

$$y^T J(x^*)y \neq 0, \quad \forall y \in \mathbb{R}^n, y \neq 0. \tag{3.38}$$

*Then $g(x^*) = 0$.*

For example, if a mapping $g(x)$ is uniformly monotone, then its Jacobian matrix $J(x)$ satisfies $y^T J(x)y \neq 0$ for all $x, y \in \mathbb{R}^n, y \neq 0$.

## 4 Numerical Results

In this section, we compare numerical performance of the DF-MLS method with that of the DF-SANE method [7], the N-DF-SANE method [5] and the DF-CGNE method [25]. The test problems are the unconstrained minimization problems in the CUTEr [3] library. Letting $f$ be an objective function of such problems the stationary point is a solution of the equation $g(x) = 0$, where $g$ is the gradient of $f$. We used the above four methods to find the stationary points of the test problems by using only their gradients. We tested 142 problems with dimensions varying from 10 to 10000. We often ran two different versions of the problems randomly for which the dimensions can be chosen arbitrarily. Since different methods may converge to different stationary points of the same function, we selected the available results by using the rule in [14]. Table 1 lists all the numerical results, which include, for each problem, the name of the problem (Prob), the dimension of the problem (Dim), the total number of iterations (Iter), the total number of function evaluations (Nfun) and the CPU time in seconds (Time), respectively. We use the symbol "$-$" to specify either of the following two cases:

(a) The number of iterations exceeded 5000.

(b) The number of backtracking iterations at some line search step exceeded 100.

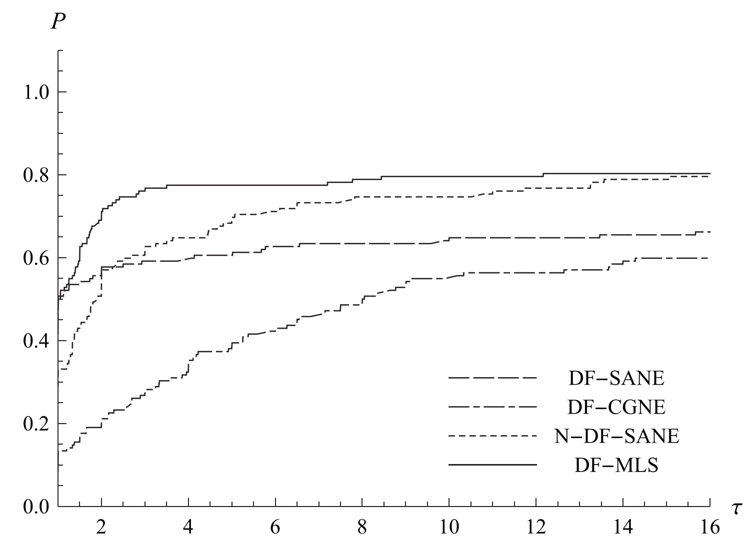


Figure 1: Performance profile relative to CPU time

The methods were coded in Fortran and run on a PC with 3.7 GHz CPU processor and 4 GB RAM and Linux operating system. The code for the DF-SANE method was obtained from professor Raydan's home page at `http://kuainasi.ciens.ucv.ve/mraydan/mraydan_pub.html`. In all these four methods, we stopped the process when

$$\frac{\|g(x_k)\|}{\sqrt{n}} \leq e_a + e_r \frac{\|g(x_0)\|}{\sqrt{n}}, \tag{4.1}$$

where $e_a = 10^{-5}$ and $e_r = 10^{-4}$. This termination condition comes from [7].

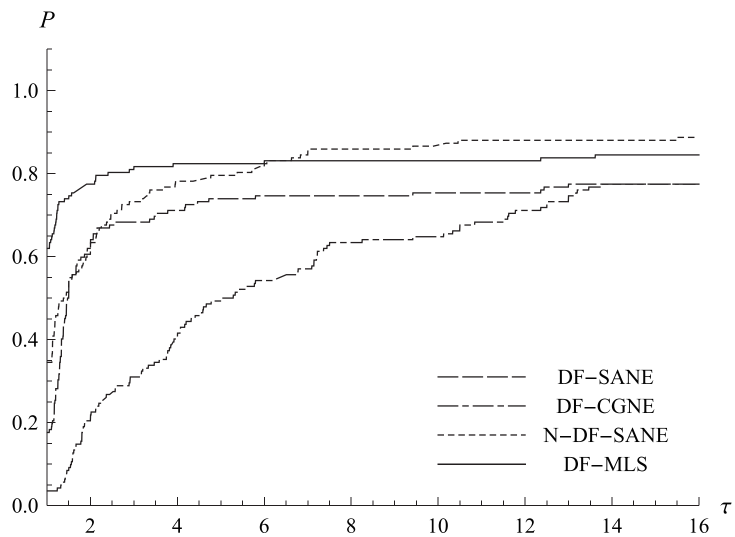


Figure 2: Performance profile relative to the number of iterations

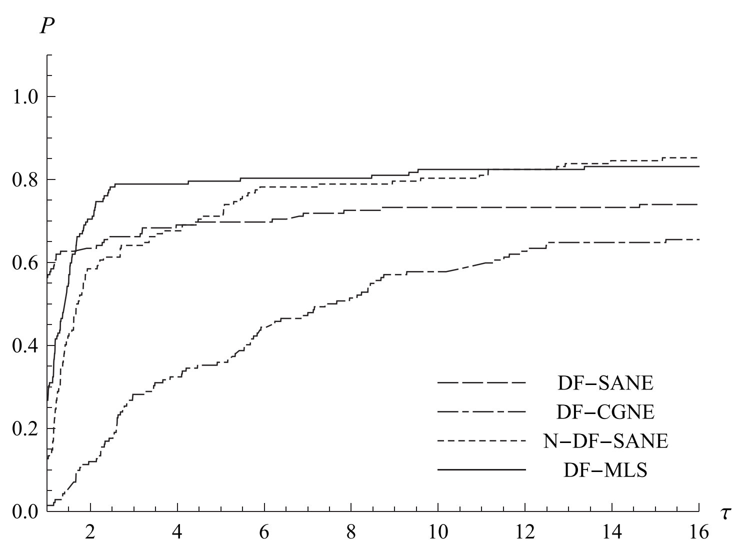


Figure 3: Performance profile relative to the number of function evaluations

We used the performance profiles by Dolan and Moré [10] to compare the performance of these four methods. Figures 1−3 show the performance of the above methods relative to the CPU time (in seconds), the total numbers of iterations and the total numbers of function evaluations, respectively. The curves in the figures have the following meanings:

"DF-SANE " denotes the DF-SANE method [7] with $nexp = 2, \sigma_{min} = 10^{-10}, \sigma_{max} = 10^{10}, \sigma_0 = 1, \tau_{min} = 0.1, \tau_{max} = 0.5, \gamma = 10^{-4}, M = 10$ and $\eta_k = \frac{\|g(x_0)\|}{2^k}$ for all $k$.

"N-DF-SANE" stands for the N-DF-SANE method [5] with $\eta_k = 0.6$, $\sigma_{min} = 10^{-10}$, $\sigma_{max} = 10^{10}$, $\rho_{min} = 0.1$, $\rho_{max} = 0.5$, $\gamma = 10^{-4}$ and $\epsilon_k = \frac{\|g(x_0)\|}{2^k}$ for all $k$.

"DF-MLS" stands for the DF-MLS method. The parameters are specified as follows: $M = 10$, $\rho = 0.5$, $t = 1$, $\lambda_1 = \lambda_2 = \lambda_3 = 10^{-4}$, $\alpha_{min} = 10^{-10}$, $\alpha_{max} = 10^{10}$ and $\eta_k = \frac{\|g(x_0)\|}{2^k}$ for all $k$. At Step 3, the choice of the initial steplength $\alpha_{0,k}$ is

$$\alpha_{0,k} = \begin{cases} \alpha_{max}, & \text{if} \quad \sigma > \alpha_{max}, \\ \sigma, & \text{if} \quad \sigma \in [\alpha_{min}, \alpha_{max}], \\ \alpha_{min}, & \text{if} \quad \sigma < \alpha_{min}, \end{cases} \tag{4.2}$$

where

$$\sigma = \frac{-g_k^T d_k}{d_k^T z_k}, \quad z_k = \frac{g(x_k + \epsilon d_k) - g(x_k)}{\epsilon}, \quad \epsilon = 10^{-8}. \tag{4.3}$$

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable function and its gradient $\nabla f(x) = g(x)$. Consider the following quadratic model

$$f(x_k + td_k) \approx q_k(t) \triangleq f(x_k) + tg_k^T d_k + \frac{t^2}{2} d_k^T \nabla^2 f(x_k) d_k. \tag{4.4}$$

If $\epsilon > 0$ is sufficiently small, then we have the following approximate equation

$$\nabla^2 f(x_k) d_k \approx z_k \triangleq \frac{g(x_k + \epsilon d_k) - g(x_k)}{\epsilon}. \tag{4.5}$$

Then, we get

$$q_k(t) \approx f(x_k) + tg_k^T d_k + \frac{t^2}{2} d_k^T z_k \tag{4.6}$$

and

$$q_k'(t) \approx g_k^T d_k + t d_k^T z_k. \tag{4.7}$$

Let $g_k^T d_k + t d_k^T z_k = 0$. We get

$$t = -\frac{g_k^T d_k}{d_k^T z_k}. \tag{4.8}$$

Therefore, it is reasonable to determine the initial steplength $\alpha_{0,k}$ by (4.2).

"DF-CGNE" means the DFCGNE method [25] with the same line search as the DF-MLS method. The original DFCGNE method consists of a search direction $d_k$ based on the PRP method and an approximately monotone line search. Yu [25] also proposed a modified version of the method, which uses a sufficiently nonmonotone line search. In order to compare the efficiency of the directions of the DFCGNE and DF-MLS methods, we used the same line search in practical computation. More precisely, the DF-CGNE method here means the DF-MLS method with $d_k$ replacing $\beta_k^{\text{MLS}}$ at Step 2 by the PRP parameter

$$\beta_k^{\text{PRP}} = \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}. \tag{4.9}$$

In Figure 1, for each method, we plot the fraction $P$ of problems for which the method is within a factor $\tau$ of the best time. It is shown that the DF-MLS method is the fastest for about 46% (66 out of 142) of the test problems. From Figure 2, we note that the DF-MLS method solves 62% (88 out of 142) of the test problems with the least number iterations. The performance of the DF-MLS method relative to the CPU time and the total number of iterations is better than that of other methods. However, we also see from Figure 3 that the DF-MLS method requires more function evaluations than the DF-SANE method. Maybe this is because that the two algorithms used different line search methods. Moreover, in the DF-MLS method, the formula for the initial steplength $\alpha_{0,k}$ also needs an additional function evaluation at each iteration. How to improve the practical efficiency of the DF-MLS method will be a future topic for us.

Table 1: Comparison of the four derivative-free methods

| | | DF-MLS | DF-SANE | N-DF-SANE | DF-CGNE |
|---|---|---|---|---|---|
| Prob | Dim | Iter/ Nfun/Time | Iter/Nfun/Time | Iter/Nfun/Time | Iter/Nfun/Time |
| ARGLINA | 100 | 2/5/0 | 3/3/0.001 | 3/5/0.001 | 2/5/0.001 |
| ARGLINA | 200 | 2/5/0.001 | 3/3/0.001 | 3/5/0.001 | 2/5/0.000999 |
| ARGLINB | 50 | 111/2597/0.044994 | 615/13715/0.18997 | 3/27/0 | 100/2478/0.032995 |
| ARGLINB | 200 | 146/4537/0.80888 | 867/24347/4.3563 | 3/33/0.006999 | 39/1209/0.21297 |
| ARGLINC | 50 | 173/4043/0.047993 | 550/12218/0.14198 | 3/27/0.001 | 5/74/0.001 |
| ARGLINC | 200 | 80/2455/0.43693 | 39/1071/0.19497 | 3/33/0.006999 | 121/4160/0.73189 |
| ARWHEAD | 100 | 5/15/0 | 10/16/0 | 6/17/0.001 | 36/81/0.000999 |
| ARWHEAD | 1000 | 4/15/0.001 | 5/13/0.001 | 4/15/0.001 | 52/250/0.013998 |
| BDQRTIC | 500 | 16/41/0.002 | 67/131/0.004999 | 86/297/0.014998 | 65/141/0.005999 |
| BDQRTIC | 5000 | 12/35/0.013998 | 50/110/0.040994 | 194/1019/0.40494 | 264/896/0.35795 |
| BOX | 100 | 51/195/0.002 | 66/198/0.002 | 309/873/0.006999 | 63/231/0.002 |
| BOX | 1000 | 70/455/0.032995 | 1820/7294/0.51692 | 706/4367/0.34895 | 285/1490/0.10698 |
| BROWNAL | 100 | 3/15/0.001 | 3/13/0.001 | 3/15/0.001 | 3/15/0.001 |
| BROWNAL | 200 | 3/15/0.001999 | 3/13/0.001999 | 3/15/0.001999 | 5/19/0.002 |
| BROYDN7D | 100 | 28/57/0.001999 | −/−/− | 91/183/0.005999 | 42/85/0.002 |
| BROYDN7D | 1000 | 25/51/0.014998 | −/−/− | 68/137/0.044993 | −/−/− |
| BRYBND | 500 | 19/43/0.001999 | 20/31/0.002 | 36/84/0.004999 | 76/161/0.007999 |
| BRYBND | 10000 | 15/35/0.036994 | 22/34/0.034994 | 26/55/0.079988 | 34/76/0.079987 |
| CHAINWOO | 4000 | 36/81/0.023997 | 170/355/0.095986 | 94/332/0.11198 | 49/109/0.031994 |
| CHAINWOO | 10000 | 22/51/0.038995 | 41/67/0.048992 | 2130/18738/13.534 | 162/406/0.31395 |
| CHNROSNB | 50 | −/−/− | −/−/− | −/−/− | −/−/− |
| COSINE | 1000 | −/−/− | −/−/− | −/−/− | 111/317/0.025996 |
| COSINE | 10000 | −/−/− | −/−/− | 527/3176/2.5826 | −/−/− |
| CRAGGLVY | 500 | 19/45/0.003999 | 28/44/0.004999 | 35/81/0.008999 | 82/181/0.020997 |
| CRAGGLVY | 5000 | 22/53/0.053992 | 31/51/0.050992 | 28/63/0.073989 | 84/187/0.16997 |
| CURLY10 | 100 | 702/2154/0.015997 | −/−/− | 435/2610/0.015997 | −/−/− |
| CURLY10 | 1000 | 722/1819/0.088987 | −/−/− | 185/752/0.045993 | 884/3698/0.18197 |
| CURLY20 | 100 | −/−/− | −/−/− | 593/3842/0.027996 | −/−/− |
| CURLY20 | 1000 | 527/1371/0.11298 | −/−/− | 182/942/0.080988 | 3784/20396/1.3248 |
| CURLY30 | 100 | −/−/− | −/−/− | 809/4720/0.049993 | −/−/− |
| CURLY30 | 1000 | 553/1439/0.12398 | −/−/− | 457/1520/0.15098 | −/−/− |
| DECONVU | 61 | 22/47/0.001 | 19/23/0 | 24/51/0 | 206/844/0.007999 |
| DIXMAANA | 1500 | 5/11/0.000999 | 6/8/0.001 | 5/11/0.001 | 23/47/0.003999 |
| DIXMAANA | 9000 | 7/17/0.008999 | 6/8/0.003999 | 5/11/0.006999 | 21/46/0.025996 |
| DIXMAANB | 300 | 6/13/0 | 7/9/0.001 | 6/13/0.001 | 144/294/0.004999 |
| DIXMAANB | 1500 | 7/17/0.000999 | 7/9/0.001 | 6/13/0.002 | 32/69/0.008999 |
| DIXMAANC | 90 | 7/17/0 | 9/13/0.001 | 7/17/0 | 50/109/0.001 |
| DIXMAANC | 9000 | 7/17/0.009997 | 9/13/0.007999 | 7/17/0.009999 | 35/77/0.042993 |
| DIXMAAND | 90 | 6/15/0 | 7/11/0.000999 | 6/15/0 | 62/131/0.001 |
| DIXMAAND | 3000 | 5/13/0.002999 | 7/11/0.003 | 6/15/0.003999 | 58/128/0.023997 |
| DIXMAANE | 90 | 19/39/0.000999 | 22/26/0.001 | 45/91/0.000999 | 25/51/0.001 |
| DIXMAANE | 9000 | 26/55/0.029995 | 22/26/0.014998 | 47/95/0.06699 | 32/68/0.037994 |
| DIXMAANF | 3000 | 13/29/0.005999 | 17/19/0.003999 | 25/51/0.012998 | 51/108/0.020997 |

*Continued*

| Prob | Dim | DF-MLS Iter/ Nfun/Time | DF-SANE Iter/ Nfun/Time | N-DF-SANE Iter/ Nfun/Time | DF-CGNE Iter/Nfun/Time |
|---|---|---|---|---|---|
| DIXMAANF | 9000 | 13/29/0.015997 | 17/19/0.010998 | 25/51/0.039994 | 50/106/0.071989 |
| DIXMAANG | 3000 | 11/25/0.004999 | 14/18/0.003 | 13/29/0.005999 | 78/167/0.030995 |
| DIXMAANG | 9000 | 11/25/0.013998 | 14/18/0.008998 | 13/29/0.017997 | 32/69/0.037994 |
| DIXMAANH | 90 | 8/19/0 | 9/13/0.001 | 9/21/0 | 51/109/0.000999 |
| DIXMAANH | 9000 | 11/27/0.019997 | 9/13/0.006999 | 9/21/0.011998 | 42/89/0.049993 |
| DIXMAANI | 90 | 28/57/0.000999 | 25/31/0.001 | 83/169/0.001 | 39/79/0 |
| DIXMAANI | 9000 | 37/77/0.042993 | 29/37/0.019997 | 84/169/0.12198 | 46/96/0.053992 |
| DIXMAANJ | 90 | 18/37/0.001 | 18/20/0.001 | 33/67/0.001 | 72/148/0.001 |
| DIXMAANJ | 9000 | 15/33/0.018997 | 18/20/0.010998 | 33/67/0.048992 | 39/82/0.045993 |
| DIXMAANK | 90 | 12/27/0.000999 | 14/18/0.001 | 15/33/0.000999 | 45/96/0.001 |
| DIXMAANK | 300 | 12/27/0.000999 | 14/18/0.001 | 15/33/0.001 | 50/105/0.002 |
| DIXMAANL | 90 | 8/19/0 | 12/16/0.001 | 9/21/0.001 | 54/114/0.000999 |
| DIXMAANL | 300 | 8/19/0.001 | 12/16/0.000999 | 9/21/0.000999 | 66/140/0.003 |
| DIXMAANL | 1500 | 8/19/0.001999 | 12/16/0.002 | 9/21/0.001999 | 81/174/0.015997 |
| DIXMAANL | 9000 | 11/27/0.015998 | 12/16/0.008999 | 9/21/0.011998 | 65/135/0.074988 |
| DIXON3DQ | 1000 | 1062/2185/0.070989 | −/−/− | −/−/− | −/−/− |
| DIXON3DQ | 10000 | 1725/3455/0.98785 | −/−/− | −/−/− | −/−/− |
| DQDRTIC | 1000 | 7/19/0.001 | 12/18/0.001 | 40/91/0.004999 | 18/42/0.002 |
| DQDRTIC | 5000 | 7/27/0.006999 | 5/11/0.002999 | 4/13/0.003999 | 8/29/0.007999 |
| DQRTIC | 50 | 10/29/0 | 11/19/0.000999 | 8/23/0 | 93/203/0.000999 |
| DQRTIC | 5000 | −/−/− | 11/27/0.002999 | 8/31/0.003999 | −/−/− |
| EDENSCH | 2000 | 13/33/0.005 | 19/25/0.004 | 12/29/0.004999 | 67/156/0.022996 |
| EG2 | 1000 | 6/13/0.001999 | 9/15/0.001999 | 6/13/0.001999 | 79/406/0.033995 |
| ENGVAL1 | 1000 | 7/17/0.001999 | 11/15/0.001 | 9/21/0.002 | 38/83/0.004999 |
| ENGVAL1 | 5000 | 6/15/0.004999 | 8/12/0.002999 | 7/17/0.004999 | 63/135/0.040993 |
| ERRINROS | 50 | 12/33/0.001 | 20/32/0.001 | 12/33/0 | 43/95/0.001 |
| EXTROSNB | 100 | 20/47/0.000999 | 40/55/0.000999 | 36/81/0.001 | 36/79/0.001 |
| EXTROSNB | 1000 | 18/47/0.002999 | 22/28/0.002 | 17/39/0.002 | 27/65/0.003 |
| FLETCBV2 | 100 | 2896/23901/0.31595 | 1851/9315/0.130 | 2171/16781/0.236 | 3730/32437/0.4349 |
| FLETCBV2 | 10000 | 1/1/0.002 | 1/1/0.001999 | 1/1/0.002 | 1/1/0.002 |
| FLETCBV3 | 100 | −/−/− | −/−/− | −/−/− | −/−/− |
| FMINSRF2 | 1024 | 594/3301/0.23996 | 1005/3883/0.27796 | 936/5999/0.45993 | 949/5491/0.39794 |
| FMINSRF2 | 10000 | 4229/37715/29.887 | 3434/17962/14.737 | 3839/30309/24.214 | −/−/− |
| FMINSURF | 5625 | −/−/− | 4723/26015/11.862 | −/−/− | −/−/− |
| FMINSURF | 10000 | −/−/− | −/−/− | −/−/− | −/−/− |
| FREUROTH | 50 | 32/77/0.001 | 694/2604/0.009998 | 87/446/0.001999 | 1019/5483/0.021996 |
| FREUROTH | 500 | 22/56/0.002 | 2165/11092/0.43493 | 74/315/0.012998 | 97/318/0.012998 |
| GENHUMPS | 1000 | −/−/− | −/−/− | −/−/− | −/−/− |
| GENHUMPS | 5000 | −/−/− | −/−/− | −/−/− | −/−/− |
| GENROSE | 100 | −/−/− | −/−/− | −/−/− | −/−/− |
| HILBERTA | 10 | 6/11/0 | 12/12/0.001 | 93/193/0.001 | 19/38/0 |
| HILBERTB | 50 | 5/11/0.001 | 5/7/0.001 | 5/11/0.001 | 9/20/0.001 |
| HYDC20LS | 99 | 42/103/0.015998 | 828/3346/0.15598 | 287/1439/0.07099 | −/−/− |
| INDEF | 1000 | −/−/− | −/−/− | −/−/− | −/−/− |
| INDEF | 5000 | −/−/− | −/−/− | −/−/− | −/−/− |
| LIARWHD | 5000 | 20/51/0.014998 | 43/97/0.025996 | 957/7417/1.9567 | 2979/21994/5.7451 |
| LIARWHD | 10000 | 14/49/0.027996 | −/−/− | 1313/10287/5.5132 | −/−/− |
| MANCINO | 50 | 5/23/0.013998 | 4/16/0.009998 | 4/19/0.011998 | 15/43/0.026995 |
| MANCINO | 100 | 30/161/0.40494 | 5/19/0.047992 | 5/23/0.06599 | 32/165/0.42094 |
| MODBEALE | 200 | 44/95/0.003 | 21/41/0.002 | 466/2913/0.10498 | 150/472/0.014998 |
| MODBEALE | 2000 | 31/78/0.022997 | 571/2332/0.6919 | 186/993/0.31195 | 179/667/0.20997 |
| MOREBV | 50 | 442/2133/0.005999 | 1078/4934/0.0120 | 742/4909/0.0120 | 805/4792/0.013 |
| MOREBV | 5000 | 3/5/0.002999 | 5/5/0.000999 | 6/11/0.003 | 7/13/0.004 |
| NONCVXUN | 100 | 41/103/0.001999 | 37/45/0.001 | 42/97/0.002 | 52/125/0.001 |
| NONDIA | 5000 | 3/17/0.004 | 3/15/0.003 | 3/17/0.007999 | 10/122/0.029995 |
| NONDIA | 10000 | 4/21/0.010999 | 4/18/0.007999 | 4/21/0.010998 | 30/499/0.24696 |
| NONDQUAR | 5000 | 18/49/0.006999 | 17/29/0.003999 | 17/43/0.006999 | 78/185/0.029996 |
| NONDQUAR | 10000 | 18/49/0.014997 | 15/27/0.007999 | 17/45/0.014997 | 202/460/0.15398 |
| NONMSQRT | 100 | −/−/− | −/−/− | 177/805/0.014997 | 338/1097/0.022997 |
| OSCIPATH | 100 | 9/19/0 | 12/14/0.001 | 13/27/0.001 | 16/34/0.000999 |
| OSCIPATH | 500 | 8/17/0.001 | 11/13/0.001 | 12/25/0.001 | 16/34/0.001 |
| PENALTY1 | 500 | 669/4897/0.087987 | 10/26/0.001 | 9/33/0.001 | 673/4907/0.080987 |
| PENALTY1 | 1000 | 4364/109573/3.4675 | 10/28/0.002 | 7/31/0.002 | 4364/109573/3.4555 |
| PENALTY2 | 50 | 9/27/0.001 | 12/22/0.001 | 9/27/0.000999 | 106/229/0.004999 |
| PENALTY2 | 100 | 9/29/0.002 | 13/25/0.002 | 10/31/0.001999 | 175/381/0.016998 |
| POWELLSG | 5000 | 24/53/0.007999 | 48/64/0.008999 | 251/591/0.10598 | 37/79/0.010998 |
| POWELLSG | 10000 | 24/53/0.016997 | 48/64/0.017997 | 159/685/0.19897 | 44/93/0.027996 |
| POWER | 100 | 10/31/0 | 12/24/0.001 | 10/31/0 | 132/289/0.000999 |
| POWER | 5000 | 17/401/0.035995 | 8/92/0.008999 | 11/43/0.004999 | 100/2432/0.22497 |
| POWER | 10000 | 17/391/0.06999 | 15/321/0.058991 | 9/41/0.008998 | 61/1606/0.30795 |
| QUARTC | 5000 | −/−/− | 11/27/0.003 | 8/31/0.003999 | −/−/− |
| QUARTC | 10000 | −/−/− | 13/31/0.007 | 9/35/0.008998 | −/−/− |
| SENSORS | 100 | 177/441/1.1798 | 353/483/1.3058 | 13/33/0.096985 | 179/412/1.2258 |
| SINQUAD | 5000 | 66/161/0.11698 | −/−/− | 305/2441/1.7627 | −/−/− |

## $\boxed{5}$ Conclusion

A derivative-free algorithm for solving large-scale nonlinear system of equations has been introduced and analyzed. It was developed based on the modified Liu-Storey conjugate gradient method in [17] and the Grippo-Lampariello-Lucidi [12] and Li-Fukushima [16] nonmonotone line search methods. The global convergence of the proposed method has been analyzed. Extensive numerical results were reported. The performance profiles showed that the proposed method is very efficient.

## References

[1] S. Bellavia and B. Morini, A globally convergent Newton-GMRES subspace method for systems of nonlinear equations, *SIAM J. Sci. Comput.* 23 (2001) 940–960.

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

[3] I. Bongartz, A.R. Conn, N.I.M. Gould and P.L. Toint, CUTE: Constrained and Unconstrained Testing Environment, *ACM Trans. Math. Software* 21 (1995) 123–160.

[4] P.N. Brown and Y. Saad, Convergence theory of nonlinear Newton-Krylov algorithms, *SIAM J. Optim.* 4 (1994) 297–330.

[5] W. Cheng and D.H. Li, A derivative-free nonmonotone line search and its application to the spectral residual method, *IMA J. Numer. Anal.* 29 (2009) 814–825.

[6] W. Cheng, Y. Xiao and Q. Hu, A family of derivative-free conjugate gradient methods for large-scale nonlinear systems of equations, *J. Comput. Appl. Math.* 224 (2009) 11–19.

[7] W. La Cruz, J.M. Martínez and M. Raydan, Spectral residual method without gradient information for solving large-scale nonlinear systems of equations, *Math. Comp.* 75 (2006) 1429–1448.

[8] W. La Cruz and M. Raydan, Nonmonotone spectral methods for large-scale nonlinear systems, *Optim. Methods Softw.* 18 (2003) 583-599.

[9] J.E. Dennis, Jr. and J.J. Moré, A characterization of superlinear convergence and its application to quasi-Newton methods, *Math. Comp.* 28 (1974) 549–560.

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

[11] M.G. Gasparo, A nonmonotone hybrid method for nonlinear systems, *Optim. Methods Softw.* 13 (2000) 79–94.

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

[13] W.W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM J. Optim.* 16(2005) 170–192.

[14] W.W. Hager and H. Zhang, Algorithm 851: CG DESCENT, a conjugate gradient method with guaranteed descent, *ACM Trans. Math. Software*, 32 (2006), 113–137.

[15] D. Li and M. Fukushima, A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric equations, *SIAM J. Numer. Anal.* 37 (2000) 152–172.

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

[17] M. Li and H. Feng, A sufficient descent LS conjugate gradient method for unconstrained optimization problems, *Appl. Math. Comput.* 218 (2011) 1577–1586.

[18] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, Part 1: Theory, *J. Optim. Theory Appl.* 69 (1991) 129–137.

[19] J.M. Martínez, A family of quasi-Newton methods for nonlinear equations with direct secant updates of matrix factorizations, *SIAM J. Numer. Anal.* 27 (1990) 1034–1049.

[20] Y. Narushima and H. Yabe, A survey of sufficient descent conjugate gradient methods for unconstrained optimization, *SUT J. Math.* 50 (2014) 167–203.

[21] Y. Narushima, H. Yabe and J.A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, *SIAM J. Optim.* 21 (2011) 212–230.

[22] E. Polak and G. Ribière, *Note sur la convergence de méthodes de directions conjuguées*, Rev. Fr. Inform. Rech. Oper., Série rouge 3 (1969), pp. 35–43.

[23] B.T. Polyak, The conjugate gradient method in extremal problems, USSR Comp. Math. Math. Phys., 9(4) (1969), pp. 94–112.

[24] M.V. Solodov and B.F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi (eds.), Kluwer Academic Publishers, Dordrecht, 1999, pp. 355–369.

[25] G. Yu, A derivative-free method for solving large-scale nonlinear systems of equations, *J. Ind. Manag. Optim.* 6 (2010) 149–160.

[26] G. Yu, L. Guan and W. Chen, Spectral conjugate gradient methods with sufficient descent property for large-scale unconstrained optimization, *Optim. Methods Softw.*, 23 (2008) 275–293.

[27] L. Zhang, A new Liu-Storey type nonlinear conjugate gradient method for unconstrained optimization problems, *J. Comput. Appl. Math.* 225 (2009) 146–157.

[28] H. Zhang and W.W. Hager, A nonmonotone line search technique and its application to unconstrained optimization, *SIAM J. Optim.* 4 (2004) 1043–1056.

[29] L. Zhang and W. Zhou, Spectral gradient projection method for solving nonlinear monotone equations, *J. Comput. Appl. Math.* 196 (2006) 478–484.

[30] L. Zhang, W. Zhou and D. Li, A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence, *IMA J. Numer. Anal.* 26 (2006) 629–640.

[31] L. Zhang, W. Zhou and D. Li, Some descent three-term conjugate gradient methods and their global convergence, *Optim. Methods Softw.* 22 (2007) 697–711.

[32] W. Zhou and D. Li, Limited memory BFGS method for nonlinear monotone equations, *J. Comput. Math.* 25 (2007) 89–96.

[33] G. Zhou and K.C. Toh, Superlinear convergence of a Newton-type algorithm for monotone equations, *J. Optim. Theory Appl.* 125 (2005) 205–221.

---

*Manuscript received 22 October 2016*
*revised 2 March 2017, 2 September 2017, 17 December 2018, 9 August 2019*
*accepted for publication 4 September 2019*

Min Li
Department of Mathematics and Computational Science
Huaihua University, Huaihua, Hunan, 418008, P.R. China
Key Laboratory of Intelligent Control Technology for Wuling-Mountain Ecological Agriculture in Hunan Province, Huaihua, Hunan, 418008, P.R. China
E-mail address: liminmath@aliyun.com