A modified conjugate gradient method based on a modified secant equation

Parvaneh Faramarzi and Keyvan Amini*

Department of Mathematics, Faculty of Science, Razi University, Kermanshah, Iran
Emails: faramarzi.parvaneh@razi.ac.ir, kamini@razi.ac.ir

Abstract. Quasi-Newton methods are one of the popular iterative schemes to solve unconstrained optimization problems. The high convergence rate and excellent precision are two prominent characteristics of the quasi-Newton methods. In this paper, according to the preferable properties of a modified secant condition, a modified conjugate gradient method is introduced. The new algorithm satisfies the sufficient descent property independent of the line search. The convergence properties of the proposed algorithm are investigated both for uniformly convex and general functions. Numerical experiments show the superiority of the proposed method.

Keywords: Conjugate gradient methods, Modified secant condition, Sufficient descent condition, Global convergence.

AMS Subject Classification: 90C30, 90C06, 65K05, 90C53.

1 Introduction

The conjugate gradient (CG) methods are one of the useful iterative methods to solve the unconstrained optimization problem as follows:

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

where \( f : \mathbb{R}^n \to \mathbb{R} \) is a continuously differentiable function, and its gradient \( g(x) = \nabla f(x) \) is available. In each iteration, the CG methods only need to evaluate the objective function and its gradient. For this reason, this
method is very desirable to solve large-scale unconstrained optimization problems. The iterative formula of these methods is as follows:

\[ x_{k+1} = x_k + \alpha_k d_k, \]  

(2)

where \( \alpha_k \) is a step length. The search direction \( d_k \) is calculated by

\[ d_k = \begin{cases} 
-g_k, & k = 0 \\
-g_k + \beta_k d_{k-1}, & k \geq 1
\end{cases} \]

(3)

where \( \beta_k \) is a scalar describing the characteristics of the CG methods and \( g_k = g(x_k) \). Various CG methods are deduced by the different choices of the conjugate parameter \( \beta_k \). Some of the most famous conjugate parameters are listed as follows:

\[ \beta_{FR}^k = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_{HS}^k = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \beta_{PRP}^k = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad \beta_{DY}^k = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \]

which are called the Fletcher-Reeves (FR) [9], Hestenes-Stiefel (HS) [13], Polak-Ribire-Polyak (PRP) [22, 23], and Dai-Yuan (DY) [6], respectively. Note that \( y_{k-1} = g_k - g_{k-1} \) and \( \|\cdot\| \) denotes the \( L_2 \) norm.

The convergence properties of these methods have been widely studied [8, 19]. Despite the appropriate numerical results of the HS and PRP methods, their convergence properties are not efficient. On the other hand, the DY and FR methods have some strong convergence characteristics, but their numerical results are weaker than the other methods [19].

Quasi-Newton methods are another prominent class of iterative techniques which have been marked due to the high convergence rate and their good accuracy. The secant equation plays a basic role in constructing the Hessian approximation, and quasi-Newton algorithms are built based on it. Despite the good properties of this equation, quasi-Newton schemes only use gradient information while their efficiency may be increased by using the function information. In this regard, many researchers modified the secant equation to improve convergence conditions of quasi-Newton schemes. They modified the secant equation to using both the objective function information and the gradient information. In the following, some of these methods are mentioned.

In 2001, Zhang and Xu [27] introduced the following modified secant equations:

\[ B_k s_{k-1} = z_{k-1}; \quad z_{k-1} = y_{k-1} + \frac{\theta_{k-1}}{s_{k-1}^T u_{k-1}} u_{k-1}, \]

(4)

where

\[ \theta_{k-1} = 6 (f_{k-1} - f_k) + 3 (g_{k-1} + g_k)^T s_{k-1}, \]

(5)
f_k = f(x_k), s_{k-1} = x_k - x_{k-1}, and \( u_{k-1} \in \mathbb{R}^n \) is an arbitrary vector satisfying \( s_{k-1}^T u_{k-1} \neq 0 \). Zhang and Xu [27] established the local and super-linear convergence of the modified method (4), under some suitable conditions. The numerical results presented by Zhang and Xu [27] showed that quasi-Newton methods based on the equation (4) are comparable to the standard quasi-Newton methods. If \( u_{k-1} = s_{k-1} \), then the modified secant equation (4) is an extension of the modified secant equation by Zhang et al. [26]. Yabe and Takano [25] incorporated a non-negative parameter \( \psi_k \) to the modified secant equation suggested by Zhang and Xu [27] and proposed the following equations:

\[
B_k s_{k-1} = z_{k-1}, \quad z_{k-1} = y_{k-1} + \psi_k \frac{\theta_{k-1}}{s_{k-1}^T u_{k-1}} u_{k-1},
\]

It is clear that the equation (6) maintains the desired properties of the equation (4).

The convergence properties of the BFGS method have been extensively investigated by some scholars for convex functions [2, 3, 16]. To improve the convergence conditions and the extension of the BFGS method for non-convex functions, Li and Fukushima [14] suggested a modified secant equation as follows

\[
B_k s_{k-1} = y^*_{k-1}; \quad y^*_{k-1} = y_{k-1} + t_{k-1} \|g_{k-1}\| s_{k-1},
\]

where

\[
t_{k-1} = 1 + \max \left\{ \frac{-y_{k-1}^T s_{k-1}}{\|s_{k-1}\|^2}, 0 \right\}.
\]

They established the local and super-linear convergence of the BFGS method for non-convex functions. In 2006, based on (7), Zhou and Zhang [28] introduced the following modified secant equation

\[
B_k s_{k-1} = z_{k-1}; \quad z_{k-1} = y_{k-1} + h_{k-1} \|g_{k-1}\| r s_{k-1},
\]

where

\[
h_{k-1} = C + \max \left\{ \frac{-y_{k-1}^T s_{k-1}}{\|s_{k-1}\|^2}, 0 \right\} \|g_{k-1}\|^{-r}.
\]

The modified secant equation (8) improves the numerical efficiency and convergence conditions of the quasi-Newton methods.

On the other hand, the efficiency of the CG methods can be improved by using the second-order information of the objective function. Dai and
P. Faramarzi, K. Amini

Liao [5] was a pioneer this idea and presented the CG method with the following parameter based on the standard secant equation:

\[ \beta_{DL}^k = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \]

where \( t > 0 \) is a parameter. They proposed the following modification of \( \beta_{DL}^k \) to ensure global convergence for general functions:

\[ \beta_{DL+}^k = \max \left\{ \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, 0 \right\} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}. \]

They showed that DL+ method has suitable convergence properties and generates appropriate results in comparison to PRP and HS methods. According to the modified secant equation (8), similar to the idea of Dai and Liao [5], Zhou and Zhang [28] presented the CG with the following parameter as follows:

\[ \beta_{ZZ}^k = \frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T z_{k-1}}. \]

Many researchers have been presented the CG methods satisfying the sufficient descent property, see for example [1, 4, 12, 17, 18]. Although sufficient descent condition is an extremely important property in proving the global convergence of the CG methods, some CG methods lack it. The search direction \( d_k \) fulfills the sufficient descent property whenever there exists a constant \( c_1 > 0 \) such that

\[ g_k^T d_k \leq -c_1 \|g_k\|^2, \quad \text{for all} \quad k \geq 0. \]

Hager and Zhang [12] modified the well-known conjugate parameter \( \beta_{HS}^k \) as follows:

\[ \beta_{HZ}^k = \beta_{HS}^k - 2 \left( \frac{\|y_{k-1}\|}{d_{k-1}^T y_{k-1}} \right)^2 g_k^T d_{k-1}, \]

which called CG-DESCENT method and is one of the best-known CG methods both in terms of theoretical features and practical performances. They established the global convergence for general functions by considering the conjugate parameter as follows:

\[ \beta_{HZ+}^k = \max \left\{ \beta_{HZ}^k, \eta_k \right\}; \quad \eta_k = \frac{-1}{\|d_k\| \min \{\eta, \|g_k\|\}}, \quad (9) \]

where \( \eta > 0 \) is a constant.
A modified conjugate gradient method

To deduce an effective family of the CG methods, Dai and Kou [4] paid attention to the self-scaling memoryless BFGS method presented by Perry [21] and Shanno [24], in which the search direction is defined by

\[ d_{k+1} = -H_{k+1}g_{k+1}, \]  

where

\[ H_{k+1} = \frac{1}{\tau_k} \left( I - \frac{s_k y_k^T + y_k s_k^T}{s_k^T y_k} \right) + \left( 1 + \frac{\|y_k\|^2}{\tau_k s_k^T y_k} \right) \frac{s_k s_k^T}{s_k^T y_k}, \]  

and \( \tau_k \) is a scaling parameter. Setting (11) in (10), the search direction with a different multiplier is computed as follows:

\[ d_{PS}^{k+1} = -g_{k+1} + \left[ \frac{g_{k+1}^T y_k}{s_{k-1}^T y_k} - \left( \tau_k + \frac{\|y_k\|^2}{s_{k-1}^T y_k} \right) \frac{g_{k+1}^T s_{k-1}}{s_{k-1}^T y_k} \right] s_k + \frac{g_{k+1}^T s_{k-1}}{s_{k-1}^T y_k} y_k. \]  

The idea of Dai and Kou [4] was to seek the closest two-term direction to the search direction (12), and they obtained the following conjugate parameter:

\[ \beta_{DK} = \frac{g_{k+1}^T y_k}{s_{k-1}^T y_k} - \left( \tau_k + \frac{\|y_k\|^2}{s_{k-1}^T y_k} \right) \frac{g_{k+1}^T s_{k-1}}{s_{k-1}^T y_k}. \]  

Dai and Kou [4] used the following scaling parameter by Oren and Luenberger [20]

\[ \tau_k = \frac{s_{k-1}^T y_k}{s_{k-1}^T s_{k-1}}. \]  

By this, the following conjugate parameter results

\[ \beta_{DK} = \frac{g_{k+1}^T y_k}{s_{k-1}^T y_k} - \left( \tau_k + \frac{\|y_k\|^2}{s_{k-1}^T y_k} \right) \frac{g_{k+1}^T s_{k-1}}{s_{k-1}^T y_k}. \]  

Based on the following conjugate parameter, they proved the global convergence of this method

\[ \beta_{DK+} = \max \left\{ \beta_{DK}^{\tau_k}, \eta \frac{g_{k+1}^T d_{k-1}}{\|d_{k-1}\|^2} \right\}, \]  

where \( \eta \in [0, 1) \) is a parameter. Numerical experiments showed that this algorithm is one of the most efficient CG methods.

In this regard, similar to the idea of Dai and Kou [4], Amini et al. [1] sought the closest two-term direction to three-term direction of Narushima et al. [18] and introduced the following parameter:

\[ \beta_{MHS} = \beta_{HS} \left( 1 - \frac{(g_{k+1}^T d_{k-1})^2}{\|g_k\|^2 \|d_{k-1}\|^2} \right). \]
They proved the global convergence of this method for general functions by considering the following truncated form:

\[ \tilde{\beta}^\text{MHS+}_k = \max\left\{ \beta^\text{MHS}_k, 0 \right\}, \]

with

\[ \tilde{\beta}^\text{MHS}_k = \beta^\text{MHS}_k - \lambda \left( \frac{\|y_{k-1}\|\theta_k}{d_{k-1}^T y_{k-1}} \right)^2 g_k^T d_{k-1}; \quad \theta_k = 1 - \frac{(g_k^T d_{k-1})^2}{\|g_k\|^2 \|d_{k-1}\|^2}, \]

where \( \lambda > \frac{1}{4} \) is a parameter. The numerical experiments showed this algorithm is effective in comparison with other algorithms.

In this study, using a modified secant equation, we improve the CG method by Dai and Kou [4] and present a modified CG method. The new search direction satisfies the sufficient descent property independent of the line search and the convexity assumption on the objective function. The new algorithm has the global convergence for general functions, under the standard assumptions. Numerical experiments indicated that the new algorithm is preferable, dealing with unconstrained optimization problems.

This paper is organized as follows. In the next section, after scheming the new idea, some basic properties of the proposed method are proven. In the third section, global convergence of the new algorithm is investigated both for general functions and uniformly convex functions. In Section 4, numerical experiments of the new method are evaluated. Some conclusions are given in the final section.

## 2 New algorithm

In this section, according to a modified secant equation, a modified CG method is proposed.

Quasi-Newton methods are one of the famous classes to solve the problem (1). We recall that both the modified secant equations (4) and (6) approximate the Hessian matrix with the high accuracy in comparison with the standard secant equation. Inspired by the idea of Dai and Liao [5], Yabe and Takano [25] proposed the following conjugate parameter:

\[ \beta^{\gamma T}_k = \frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T z_{k-1}}, \]

where \( z_{k-1} \) is defined by (6). They showed that the numerical results of this method are improved by using the modified secant equation (6).
Motivated by the idea of Yabe and Takano [25], strong theory features of the equation (6), and considering high efficiency of the DK method [4], we propose the new conjugate parameter $\beta_{MDK}^k$ as follows:

$$\beta_{MDK}^k = \frac{g_k^T y_{k-1}}{d_{k-1}^T z_{k-1}} - \frac{\|y_{k-1}\|^2}{d_{k-1}^T z_{k-1}} \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}},$$

(15)

where $z_{k-1}$ is defined by (6). This formula is an extension of DK CG parameter [4]. The relation (15) is resulted by replacing vector $y_{k-1}$ with $z_{k-1}$ in the denominator of the relation (13).

Because $d_{k-1}^T z_{k-1}$ may turn out to zero for a general function $f$, so the formula (15) may not well-defined. To overcome this drawback, similar to Li et al. [15]; we consider a modified version of (6) as follows:

$$B_k s_k = z_{k-1}^+; \quad z_{k-1}^+ = y_{k-1} + \psi_k \max\{0, \theta_{k-1}\} \frac{s_k^T u_{k-1}}{s_k^T u_{k-1}} u_{k-1},$$

(16)

with $\theta_{k-1}$ defined by (5).

The following lemma shows that the new direction satisfies the sufficient descent property independent of the line search.

**Lemma 1.** Consider the CG method (2), (3) with (15). If $d_{k-1}^T z_{k-1} \neq 0$, then we have

$$g_k^T d_k \leq -\frac{3}{4} \|g_k\|^2.$$  

(17)

Proof. Since $d_0 = -g_0$, we have

$$g_0^T d_0 = -\|g_0\|^2 \leq -\frac{3}{4} \|g_0\|^2,$$

which implies that (17) holds for $k = 0$. For $k \geq 1$, multiplying (3) by $g_k^T$, we have

$$g_k^T d_k = -\|g_k\|^2 + \frac{g_k^T y_{k-1}}{d_{k-1}^T z_{k-1}} - \frac{\|y_{k-1}\|^2}{d_{k-1}^T z_{k-1}} \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}} d_{k-1}^T z_{k-1}

- \frac{(g_k^T y_{k-1}) (d_{k-1}^T z_{k-1}) (g_k^T d_{k-1}) - \|g_k\|^2 (d_{k-1}^T z_{k-1})^2 - \|y_{k-1}\|^2 (g_k^T d_{k-1})^2}{(d_{k-1}^T z_{k-1})^2}.$$

(18)

Applying the inequality $u_k^T v_k \leq \frac{1}{2} (\|u_k\|^2 + \|v_k\|^2)$ to the first term (18) with

$$u_k = \frac{1}{\sqrt{2}}(d_{k-1}^T z_{k-1}) g_k, \quad v_k = \sqrt{2} (g_k^T d_{k-1}) y_{k-1},$$

we obtain (17).

$\square$
The new method with $\theta_{k-1} = 0$ reduces to the HS method. It is expected that the convergence properties of the proposed parameter (15) be similar to the HS method. Therefore, according to the idea of Gilbert and Nocedal [10] to prove the global convergence of the HS method for general objective functions, we consider the following truncating

$$\beta_k^{MDK+} = \max\{0, \beta_k^{MDK}\}. \quad (19)$$

Using this truncating, the sufficient descent condition still holds for the new search direction. Based on the above, the new algorithm is presented for the proposed method.

**Algorithm 1 (MDK algorithm)**

**Input.** Given constant $\epsilon > 0$. Chosen $x_0 \in \mathbb{R}^n$ and set $k = 0$, $d_0 = -g_0$.

**Step 1.** If $\|g_k\|_\infty \leq \epsilon$, stop.

**Step 2.** Compute $d_k$ by (3) and (19).

**Step 3.** Determine the step length $\alpha_k$ by an appropriate line search.

**Step 4.** Set $x_{k+1} = x_k + \alpha_k d_k$.

**Step 5.** Set $k = k + 1$ and go to Step 1.

### 3 Global convergence properties

In this part, the global convergence of the MDK algorithm is analyzed. Therefore, the following assumptions will be needed through this study. Without loss of generality, suppose $g_k \neq 0$ for all $k \geq 0$; otherwise, we assume that a stationary point can be found.

**Assumption 1**

**(A1)** The level set $L_0 = \{x | f(x) \leq f(x_0)\}$ is bounded, namely, there exists a constant $B > 0$ such that

$$\|x\| \leq B, \quad \text{for all } x \in L_0. \quad (20)$$

**(A2)** In some open neighborhood $N \in L_0$, $f$ is continuously differentiable and its gradient $g$ is Lipschitz continuous, namely, there exists a positive constant $L$ such that

$$\|g(x) - g(y)\| \leq L \|x - y\|, \quad \text{for all } x, y \in N. \quad (21)$$

Assumption 1 conclude that there is a constant $\gamma > 0$ such that

$$\|g(x)\| \leq \gamma, \quad \text{for all } x \in L_0. \quad (22)$$
To guarantee the global convergence of a nonlinear CG method, we need to impose a line search on it. Suppose that the step length $\alpha_k$ fulfills the Wolfe conditions, i.e.,

\begin{align}
\left(1\right) f(x_k + \alpha_k d_k) & \leq f(x_k) + \delta \alpha_k g_k^T d_k, \\
\left(2\right) g(x_k + \alpha_k d_k)^T d_k & \geq \sigma g_k^T d_k,
\end{align}

where $0 < \delta \leq \sigma < 1$. Note that the Wolfe conditions and (17) yield

\begin{equation}
d_{k-1}^T z_{k-1} \geq d_{k-1}^T y_{k-1} \geq (\sigma - 1) g_{k-1}^T d_{k-1} \geq \frac{3}{4} (1 - \sigma) \| g_{k-1} \|^2 > 0,
\end{equation}

which implies that (15) and (19) are well defined.

The next lemma, called Zoutendijk condition [29], plays an important role to establish the MDK algorithm.

**Lemma 2.** [29] Suppose that Assumption 1 holds. Consider any CG method in the forms (2)-(3), where step length $\alpha_k$ satisfies the Wolfe line search and the search direction $d_k$ is decreasing. If

\begin{equation}
\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} = \infty,
\end{equation}

then,

\begin{equation}
\lim_{n \to \infty} \inf \|g_k\| = 0.
\end{equation}

**Lemma 3.** [25] Suppose that Assumption 1 holds. For $\theta_{k-1}$ given by (5), we have

\begin{equation}
|\theta_{k-1}| \leq 3L\|s_{k-1}\|^2.
\end{equation}

**Proof.** This relation immediately is resulted from the relation (5.12) from [25].

Subsequently, we establish the global convergence of the MDK algorithm for uniformly convex functions.

**Theorem 1.** Suppose that $f$ is uniformly convex, i.e., there exists a constant $\mu > 0$, such that

\begin{equation}
(g(x) - g(y))^T (x - y) \geq \mu \|x - y\|^2, \quad \forall x, y \in \mathbb{R}^n.
\end{equation}

Also, Assumption 1 holds and $\psi_k$ satisfies $0 \leq \psi_k \leq \bar{\rho}$ where $\bar{\rho}$ is a positive constant such that $\bar{\rho} < \frac{\mu}{3L}$. Let $\{x_k\}$ be generated by Algorithm 1, then either $\|g_k\| = 0$ for some $k$ or

\begin{equation}
\lim_{n \to \infty} \inf \|g_k\| = 0.
\end{equation}
Proof. Due to the sufficient descent condition (17), we get $d_k \neq 0$. By considering Lemma 2, it suffices to show that $\|d_k\|$ is bounded above. The relations (6), (27), and (28) conclude that

$$|d_{k-1}^T z_{k-1}| = |d_{k-1}^T y_{k-1} + \psi_k \frac{\theta_{k-1}}{s_{k-1}^T u_{k-1}} d_{k-1}^T u_{k-1}|$$

$$\geq |d_{k-1}^T y_{k-1}| - 3\psi_k L \frac{\|s_{k-1}\|^2}{\alpha_{k-1}}$$

$$\geq (\mu - 3\overline{\rho} L) \alpha_{k-1} \|d_{k-1}\|^2.$$

Because $\mu - 3\overline{\rho} L > 0$, we obtain by this, (15), (21), and (29) that

$$|\beta_{MDK}| = \left| \frac{g_k^T y_{k-1}}{d_{k-1}^T z_{k-1}} - \frac{\|y_{k-1}\|^2}{\|d_{k-1}^T z_{k-1}\|^2} \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}} \right|$$

$$\leq \frac{L \|g_k\| \|s_{k-1}\|}{(\mu - 3\overline{\rho} L) \alpha_{k-1} \|d_{k-1}\|^2} + \frac{L^2 \|s_{k-1}\|^2 \|g_k\| \|d_{k-1}\|}{(\mu - 3\overline{\rho} L)^2 \alpha_{k-1}^2 \|d_{k-1}\|^4}$$

$$= A \frac{\|g_k\|}{\|d_{k-1}\|},$$

where

$$A = \frac{L}{\mu - 3\overline{\rho} L} + \frac{L^2}{(\mu - 3\overline{\rho} L)^2}.$$

This relation together with (3) and (22) yield

$$\|d_k\| \leq \|g_k\| + |\beta_{MDK}| \|d_{k-1}\| \leq (1 + A) \gamma < \infty.$$

This completes the proof.

Property (*) plays an important role to prove the global convergence of the CG methods. In the following, this property is expressed.

Property (*). Consider the CG method (2)-(3) and suppose that there exist positive constants $\gamma$ and $\overline{\gamma}$ such that $0 < \overline{\gamma} \leq \|g_k\| \leq \gamma$, for all $k \geq 1$. Then, the method has property (*) if there exist constants $b > 1$ and $\nu > 0$ such that for all $k$

$$|\beta_k| \leq b,$$

and

$$\|s_{k-1}\| \leq \nu \implies |\beta_k| \leq \frac{1}{2b}.$$

Property (*) has been first formally stated by Gilbert and Nocedal [10].
Lemma 4. Suppose that Assumption 1 holds. Also, $\psi_k$ satisfies $0 \leq \psi_k \leq \bar{\rho}$ where $\bar{\rho}$ is a fixed positive constant. Then, the new algorithm with (19) satisfies property (*).

Proof. We assume that there exists a positive constant $\bar{\gamma}$ such that

$$\bar{\gamma} \leq \|g_k\|,$$

(30)

holds for any $k \geq 0$. Because $\{x_k\} \subset L_0$ and the level set $L_0$ is bounded, from (20), we have

$$\|s_{k-1}\| < 2B.$$  

(31)

From (17), (24), and (30), we conclude that

$$d_{k-1}^T y_{k-1} \geq \frac{3}{4} (1 - \sigma) \|g_{k-1}\|^2 \geq \frac{3}{4} (1 - \sigma) \bar{\gamma}^2.$$  

(32)

Using relation $g_{k-1}^T d_{k-1} < 0$ and (24), result

$$g_{k-1}^T d_{k-1} \leq g_{k-1}^T d_{k-1} - g_{k-1}^T d_{k-1} = d_{k-1}^T y_{k-1},$$

$$g_{k-1}^T d_{k-1} \leq \sigma g_{k-1}^T d_{k-1} = -\sigma d_{k-1}^T y_{k-1} + \sigma g_{k-1}^T d_{k-1}.$$  

Since $\sigma \in (0, 1)$ and $d_{k-1}^T y_{k-1} > 0$, we get

$$|g_{k-1}^T d_{k-1}| \leq \max \left\{1, \frac{\sigma}{1 - \sigma} \right\} d_{k-1}^T y_{k-1} = c_1 d_{k-1}^T y_{k-1},$$  

(33)

where $c_1 = \max \left\{1, \frac{\sigma}{1 - \sigma} \right\}$. Using (16), (32) and $\psi_k \geq 0$, we conclude that

$$|d_{k-1}^T z_{k-1}^+| = |d_{k-1}^T y_{k-1} + \frac{\psi_k}{\alpha_{k-1}} \max \{0, \theta_{k-1}\}|$$

$$\geq d_{k-1}^T y_{k-1} \geq \frac{3}{4} (1 - \sigma) \bar{\gamma}^2.$$  

(34)

From (33) and (34), we have

$$|g_{k-1}^T d_{k-1}| \leq c_1 |d_{k-1}^T y_{k-1}| \leq c_1 |d_{k-1}^T z_{k-1}^+|.$$  

(35)

Therefore, from (31), (34) and (35), we obtain

$$|\beta_{MK+}^k| \leq \frac{\|g_k\| \|y_{k-1}\|}{\|d_{k-1}^T z_{k-1}^+\|} + \frac{\|y_{k-1}\|^2 |g_{k-1}^T d_{k-1}|}{(d_{k-1}^T z_{k-1}^+)^2}$$

$$\leq \frac{\gamma L \|s_{k-1}\|}{\frac{3}{4} (1 - \sigma) \bar{\gamma}^2} + \frac{c_1 L^2 \|s_{k-1}\|^2}{\frac{3}{4} (1 - \sigma) \bar{\gamma}^2}$$

$$\leq \frac{\gamma L \|s_{k-1}\|}{\frac{3}{4} (1 - \sigma) \bar{\gamma}^2} + \frac{2Bc_1 L^2}{\frac{3}{4} (1 - \sigma) \bar{\gamma}^2} \|s_{k-1}\|$$

$$= c_2 \|s_{k-1}\|,$$  

where $c_2$ is a positive constant.
where 
\[ c_2 = \frac{4\gamma L}{3(1 - \sigma)\gamma^2} + \frac{8BL^2c_1}{3(1 - \sigma)\gamma^2}. \]
This relation implies that 
\[ |\beta_{MDK}^k + k| \leq c_2\|s_{k-1}\|, \] for all \( k \). By putting 
\[ \nu = \frac{1}{2bc^2}, \] we obtain the desired result. 

The following theorem corresponds to Theorem 4.3 in [10], and we mention it for readability.

**Theorem 2.** [10] Suppose that Assumption 1 holds. Consider the method (2)-(3) which satisfies the following conditions:

- \((C_1)\) \( \beta_k \geq 0 \) for all \( k \geq 0 \);
- \((C_2)\) the Zoutendijk condition;
- \((C_3)\) the sufficient descent condition;
- \((C_4)\) property (*);

then \( \lim inf_{k \to \infty} \|g_k\| = 0 \).

Now, we can provide the global convergence result for MDK algorithm.

**Theorem 3.** Suppose that all conditions of Lemma 4 hold. Then, MDK algorithm with (19) is globally convergent, i.e.,

\[ \lim inf_{n \to \infty} \|g_k\| = 0. \]

**Proof.** It is sufficient to show that conditions of Theorem 2 are held. From (19), clearly, \( \beta_{MDK}^k \geq 0 \) and the MDK algorithm with \( \beta_{MDK}^k \) satisfies the sufficient descent condition. The relation (21) and Wolfe conditions get the Zoutendijk condition. Hence, conditions \((C_1)\), \((C_2)\), and \((C_3)\) of the Theorem 2 are satisfied. The previous lemma provides property (*) for the MDK algorithm; hence \((C_4)\) is satisfied. This completes the proof. 

### 4 Numerical experiments

In this section, we report the numerical performance of the new algorithm. To investigate the numerical efficiency of the MDK algorithm, we compare this algorithm with the HZ+ method, developed by Hager and Zhang [12], and the DK+ method, proposed by Dai and Kou [4], on a collection of the 119 test problems of the CUTEst library [11], with dimensions between 2 to 20000. The codes were written in MATLAB 8.1 and were implemented on a Vaio Laptop with 2.30 GHz of CPU, 4 GB of RAM. The line search in all the algorithms is the strong Wolfe line search proposed by Nocedal.
and Wright [19] with $\delta = 0.01$ and $\sigma = 0.1$, in which the initial step length selected as follows:

$$
\alpha_k^{(0)} = \begin{cases} 
1, & k = 1, \\
\alpha_{k-1} \frac{g_{k-1}^T d_{k-1}}{g_k^T d_k}, & k > 1.
\end{cases}
$$

All algorithms are terminated when

$$
\|g_k\|_\infty \leq 10^{-6},
$$

or the number of iterations exceeds 10000, and show it with the symbol “Failed” in Table 1. In the figures, the curves have the following meaning:

**MDK+**: The CG method with $\beta_k^{MDK+}$ given by (19), $\psi_k = 0.6$ and $u_{k-1} = y_{k-1}$.

**HZ+**: The CG method proposed by Hager and Zhang [12] with $\beta_k^{HZ+}$ given by (9), $\eta = 0.01$.

**DK+**: The CG method proposed by Dai and Kou [4] with $\beta_k^{DK+}$ given by (14), $\eta = 0.5$ and

$$
\tau_k = \frac{s_{k-1}^T y_{k-1}}{s_k^T s_{k-1}}.
$$

The numerical results of the mentioned algorithms are presented in Table 1. This table includes the name of the test problems (Prob), their dimensions (DIM), the total number of iterations ($k$), the total number of function evaluations (Nf), and the total number of gradient evaluations (Ng). Based on the performance profile proposed by Dolan and Moré [7],

Figure 1: Performance profile for the number of iterations.
the numerical results of the algorithms are compared, while including the number of iterations, the number of function evaluations, and the number of gradient evaluations, respectively.

The preliminary experiments demonstrate the efficiency of the new algorithm. In details, Figure 1 shows that “MDK+” has the best performance among all algorithms, which solves about 65% of the test problems with the least number of iterations, while “HZ+” and “DK+” solve almost 56% of the test problems. Figure 2 shows that “MDK+” solves 59% of the test problems with the least number of function evaluations, where “HZ+” and “DK+” solve almost 51% and 53% of the test problems, respectively. Figure 3 is used to report the number of gradient evaluations, where the ”MDK+” solves 59% of the test problems with the least number of gradient evaluations, while “HZ+” and “DK+” can solve about 54% and 52%, respectively. Hence, the numerical results indicate that the “MDK+” performs better than the other algorithms (HZ+, DK+).

Table 1. Numerical results.

<table>
<thead>
<tr>
<th>Prob</th>
<th>Dim</th>
<th>MDK+ k/Nf/Ng</th>
<th>HZ+ k/Nf/Ng</th>
<th>DK+ k/Nf/Ng</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRCRAFT B</td>
<td>8</td>
<td>43/201/168</td>
<td>45/203/165</td>
<td>35/171/138</td>
</tr>
<tr>
<td>ALLINITU</td>
<td>4</td>
<td>9/41/28</td>
<td>11/50/36</td>
<td>13/57/40</td>
</tr>
<tr>
<td>ARGLINA</td>
<td>200</td>
<td>1/4/3</td>
<td>1/4/3</td>
<td>1/4/3</td>
</tr>
<tr>
<td>ARGLINB</td>
<td>10</td>
<td>2/7/5</td>
<td>2/7/5</td>
<td>2/7/5</td>
</tr>
<tr>
<td>ARGLINCC</td>
<td>10</td>
<td>2/7/5</td>
<td>2/7/5</td>
<td>2/7/5</td>
</tr>
<tr>
<td>ARWHEAD</td>
<td>500</td>
<td>9/81/58</td>
<td>9/84/61</td>
<td>9/84/61</td>
</tr>
<tr>
<td>BARD</td>
<td>3</td>
<td>24/97/81</td>
<td>62/231/198</td>
<td>28/110/95</td>
</tr>
<tr>
<td>BDEXP</td>
<td>5000</td>
<td>2/6/5</td>
<td>2/6/5</td>
<td>2/6/5</td>
</tr>
<tr>
<td>BEALE</td>
<td>2</td>
<td>12/71/53</td>
<td>12/60/45</td>
<td>12/62/47</td>
</tr>
<tr>
<td>BIGGS3</td>
<td>6</td>
<td>165/673/600</td>
<td>368/1481/1349</td>
<td>510/2011/1836</td>
</tr>
<tr>
<td>BIGGS5</td>
<td>6</td>
<td>165/673/600</td>
<td>368/1481/1349</td>
<td>510/2011/1836</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Iterations</td>
<td>Time (s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIGGS6</td>
<td>6</td>
<td>133/590/533</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIGGSB1</td>
<td>5000</td>
<td>Failed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOX</td>
<td>100</td>
<td>9/41/35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOX2</td>
<td>3</td>
<td>14/55/43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOX3</td>
<td>3</td>
<td>14/55/43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRKMCC</td>
<td>2</td>
<td>7/22/15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BROWNBS</td>
<td>2</td>
<td>16/63/48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BROYD17ND</td>
<td>500</td>
<td>207/637/604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRYBND</td>
<td>5000</td>
<td>64/247/205</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHAINWOO</td>
<td>1000</td>
<td>410/164/1257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHNROSNB</td>
<td>50</td>
<td>298/976/934</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSINE</td>
<td>1000</td>
<td>9/40/25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUBED</td>
<td>2</td>
<td>21/171/133</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DECONVUL</td>
<td>63</td>
<td>1807/5602/5550</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSCHNA</td>
<td>2</td>
<td>9/50/36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSCHNB</td>
<td>2</td>
<td>7/39/27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSCHND</td>
<td>3</td>
<td>20/479/235</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSCHNE</td>
<td>3</td>
<td>12/116/85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSCHNF</td>
<td>2</td>
<td>10/67/45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANA</td>
<td>3000</td>
<td>6/41/28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANB</td>
<td>3000</td>
<td>5/54/37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANC</td>
<td>3000</td>
<td>6/67/45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAAND</td>
<td>3000</td>
<td>7/73/49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANE</td>
<td>3000</td>
<td>264/695/682</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANF</td>
<td>3000</td>
<td>205/571/551</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANG</td>
<td>3000</td>
<td>200/571/549</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANH</td>
<td>3000</td>
<td>200/583/555</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANI</td>
<td>3000</td>
<td>4206/10511/10499</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANJ</td>
<td>3000</td>
<td>360/902/883</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANK</td>
<td>3000</td>
<td>313/806/784</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANL</td>
<td>3000</td>
<td>290/766/741</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANM</td>
<td>15</td>
<td>57/192/166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANN</td>
<td>15</td>
<td>73/240/271</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANO</td>
<td>15</td>
<td>54/192/161</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXMAANP</td>
<td>3000</td>
<td>1704/4238/4219</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIXON3DQ</td>
<td>1000</td>
<td>6757/16526/16522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQDRTIC</td>
<td>5000</td>
<td>5/22/19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQRTIC</td>
<td>5000</td>
<td>16/281/188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDENSCH</td>
<td>36</td>
<td>23/117/82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG2</td>
<td>1000</td>
<td>2/20/7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENGVAL1</td>
<td>100</td>
<td>25/122/82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENGVAL2</td>
<td>3</td>
<td>45/265/226</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERRINROS</td>
<td>10</td>
<td>Failed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXTROSNB</td>
<td>1000</td>
<td>Failed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLETCHCR</td>
<td>1000</td>
<td>4454/11659/11650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMINSRF2</td>
<td>5625</td>
<td>360/902/883</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMINSURF</td>
<td>5625</td>
<td>557/1383/1383</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FREUROT1</td>
<td>2</td>
<td>12/90/70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENHUMPS</td>
<td>5000</td>
<td>9358/28365/28096</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENROSE</td>
<td>500</td>
<td>1091/3382/3325</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROWTHLS</td>
<td>3</td>
<td>1/21/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GULF</td>
<td>3</td>
<td>201/816/753</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HATFLDD</td>
<td>3</td>
<td>52/272/235</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HATFLDFL</td>
<td>3</td>
<td>31/147/128</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEART6LS</td>
<td>6</td>
<td>Failed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEART8LS</td>
<td>8</td>
<td>5030/15605/15467</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Numerical results. (continued)

<table>
<thead>
<tr>
<th>Function</th>
<th>Mod</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HELIX 3</td>
<td>3</td>
<td>31/137/114</td>
<td>55/215/180</td>
<td>55/215/180</td>
</tr>
<tr>
<td>HILBERTA 2</td>
<td>2</td>
<td>2/7/6</td>
<td>2/7/6</td>
<td>2/7/6</td>
</tr>
<tr>
<td>HILBERTB 10</td>
<td>10</td>
<td>4/13/9</td>
<td>4/13/9</td>
<td>4/13/9</td>
</tr>
<tr>
<td>HIMMELBG 2</td>
<td>2</td>
<td>7/36/26</td>
<td>8/35/26</td>
<td>8/34/25</td>
</tr>
<tr>
<td>HIMMELBH 2</td>
<td>2</td>
<td>8/29/20</td>
<td>7/28/19</td>
<td>8/31/21</td>
</tr>
<tr>
<td>HUMPS 2</td>
<td>2</td>
<td>29/213/155</td>
<td>16/144/99</td>
<td>24/186/122</td>
</tr>
<tr>
<td>KOWOSB 4</td>
<td>4</td>
<td>65/269/244</td>
<td>148/587/530</td>
<td>51/215/196</td>
</tr>
<tr>
<td>LIARWHD 5000</td>
<td>5000</td>
<td>16/129/99</td>
<td>19/138/107</td>
<td>20/140/108</td>
</tr>
<tr>
<td>LOGHAIRY 2</td>
<td>2</td>
<td>53/420/308</td>
<td>29/314/210</td>
<td>35/293/221</td>
</tr>
<tr>
<td>LMINSURF 5625</td>
<td>5625</td>
<td>350/878/878</td>
<td>347/872/872</td>
<td>347/872/872</td>
</tr>
<tr>
<td>MANCINO 100</td>
<td>100</td>
<td>11/34/23</td>
<td>11/34/23</td>
<td>11/34/23</td>
</tr>
<tr>
<td>MODBEALE 20000</td>
<td>20000</td>
<td>554/2317/2025</td>
<td>798/2734/2584</td>
<td>Failed</td>
</tr>
<tr>
<td>MOREBV 5000</td>
<td>5000</td>
<td>107/392/390</td>
<td>167/392/390</td>
<td>167/392/390</td>
</tr>
<tr>
<td>MSQRITALS 1024</td>
<td>1024</td>
<td>4611/7345/7324</td>
<td>2970/7309/7208</td>
<td>2995/7359/7348</td>
</tr>
<tr>
<td>NLMsurf 5625</td>
<td>5625</td>
<td>392 985 983</td>
<td>404/1027/1024</td>
<td>404/1027/1024</td>
</tr>
<tr>
<td>NONCVXU2 5000</td>
<td>5000</td>
<td>Failed</td>
<td>Failed</td>
<td>Failed</td>
</tr>
<tr>
<td>NONDIA 5000</td>
<td>5000</td>
<td>10/96/72</td>
<td>11/102/78</td>
<td>11/102/78</td>
</tr>
<tr>
<td>NONDQUAR 5000</td>
<td>5000</td>
<td>7191/27269/25533</td>
<td>2724/10760/9838</td>
<td>1890/7493/6845</td>
</tr>
<tr>
<td>NONSCOMP 5000</td>
<td>5000</td>
<td>43/177/125</td>
<td>43/179/128</td>
<td>43/179/126</td>
</tr>
<tr>
<td>OSCIPATH 10</td>
<td>10</td>
<td>Failed</td>
<td>Failed</td>
<td>Failed</td>
</tr>
<tr>
<td>OSBORNEB 11</td>
<td>11</td>
<td>399/1384/1311</td>
<td>277/961/886</td>
<td>297/1042/950</td>
</tr>
<tr>
<td>PALMERSC 6</td>
<td>6</td>
<td>6/19/14</td>
<td>6/19/14</td>
<td>6/19/14</td>
</tr>
<tr>
<td>PENALTY1 100</td>
<td>100</td>
<td>27/287/212</td>
<td>29/278/204</td>
<td>26/269/198</td>
</tr>
<tr>
<td>PENALTY2 50</td>
<td>50</td>
<td>1608/6119/5679</td>
<td>384/1649/1485</td>
<td>879/3699/3184</td>
</tr>
<tr>
<td>POWELL 5000</td>
<td>5000</td>
<td>139/572/508</td>
<td>91/362/313</td>
<td>306/1317/1161</td>
</tr>
<tr>
<td>POWER 100</td>
<td>100</td>
<td>36/203/149</td>
<td>36/203/149</td>
<td>36/203/149</td>
</tr>
<tr>
<td>QUARTC 5000</td>
<td>5000</td>
<td>16/281/188</td>
<td>16/281/188</td>
<td>16/281/188</td>
</tr>
<tr>
<td>ROSENBR 2</td>
<td>2</td>
<td>27/158/127</td>
<td>33/186/150</td>
<td>31/171/135</td>
</tr>
<tr>
<td>S308 2</td>
<td>2</td>
<td>7/69/46</td>
<td>7/69/46</td>
<td>7/69/46</td>
</tr>
<tr>
<td>SCHM VETT 100</td>
<td>100</td>
<td>43/135/109</td>
<td>43/135/109</td>
<td>43/135/109</td>
</tr>
<tr>
<td>SENSORS 10</td>
<td>10</td>
<td>20/76/56</td>
<td>20/79/60</td>
<td>20/80/60</td>
</tr>
<tr>
<td>SINEVAL 2</td>
<td>2</td>
<td>51/246/212</td>
<td>52/282/235</td>
<td>50/271/225</td>
</tr>
<tr>
<td>SINEWAVE 5</td>
<td>5</td>
<td>31/115/95</td>
<td>22/91/74</td>
<td>34/126/104</td>
</tr>
<tr>
<td>SISER 2</td>
<td>2</td>
<td>5/76/51</td>
<td>5/76/51</td>
<td>5/76/51</td>
</tr>
<tr>
<td>SNAIL 2</td>
<td>2</td>
<td>5/27/16</td>
<td>6/28/18</td>
<td>6/26/16</td>
</tr>
<tr>
<td>SPARSE 5000</td>
<td>5000</td>
<td>Failed</td>
<td>Failed</td>
<td>Failed</td>
</tr>
<tr>
<td>SPARSQUR 10000</td>
<td>10000</td>
<td>15/199/136</td>
<td>15/199/136</td>
<td>15/199/136</td>
</tr>
<tr>
<td>SPMSQURTLS 4999</td>
<td>4999</td>
<td>224/609/601</td>
<td>234/647/638</td>
<td>234/647/638</td>
</tr>
<tr>
<td>SROSENBR 5000</td>
<td>5000</td>
<td>11/71/52</td>
<td>11/71/52</td>
<td>11/71/52</td>
</tr>
<tr>
<td>TESTQUAD 5000</td>
<td>5000</td>
<td>6794/15025/18018</td>
<td>7889/20705/20698</td>
<td>6336/16829/16822</td>
</tr>
<tr>
<td>TOINTGSS 5000</td>
<td>5000</td>
<td>5/16/11</td>
<td>5/16/11</td>
<td>4/13/9</td>
</tr>
<tr>
<td>TOINTQOR 50</td>
<td>50</td>
<td>29/88/64</td>
<td>30/89/64</td>
<td>29/88/64</td>
</tr>
<tr>
<td>TOUARTIC 5000</td>
<td>5000</td>
<td>15/88/72</td>
<td>16/106/90</td>
<td>12/98/84</td>
</tr>
<tr>
<td>TRIDAD 5000</td>
<td>5000</td>
<td>2398/5741/5736</td>
<td>2381/5717/5712</td>
<td>2415/5785/5780</td>
</tr>
<tr>
<td>VARDIM 200</td>
<td>200</td>
<td>16/586/325</td>
<td>16/586/325</td>
<td>16/586/325</td>
</tr>
<tr>
<td>WATSON 12</td>
<td>12</td>
<td>1475/5684/5212</td>
<td>1429/5786/5261</td>
<td>1739/6782/6203</td>
</tr>
<tr>
<td>WOODS 4000</td>
<td>4000</td>
<td>2263783/747</td>
<td>243/824/783</td>
<td>243/824/783</td>
</tr>
<tr>
<td>YFITU 3</td>
<td>3</td>
<td>1176/4416/4057</td>
<td>946/3560/3249</td>
<td>1322/4811/4514</td>
</tr>
<tr>
<td>ZANGWIL2 2</td>
<td>2</td>
<td>114/4</td>
<td>1/4/4</td>
<td>1/4/4</td>
</tr>
</tbody>
</table>
A modified conjugate gradient method

1.5
2
2.5
3
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
MDK+
HZ+
DK+

Figure 3: Performance profile based on the number of gradient evaluations.

5 Conclusion

The modified secant equations by Zhang et al. [26] and Zhang and Xu [27] possess strong theoretical properties. The modified secant equation approximates the Hessian matrix with high accuracy. In this study, regarding the strong theory features of a modified secant equation, we decided to modify the DK method [4] by using it, then a new CG method is presented. The new search direction always satisfies the sufficient descent condition, and the proposed method has the global convergence for general functions. Numerical experience on the CUTEst collection shows the MDK algorithm is capable.

References


A modified conjugate gradient method


