Advances in Mathematics: Scientific Journal **9** (2020), no.7, 4955–4970
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.7.61

A NEW MODIFICATION OF NPRP CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION

MAULANA MALIK[1], MUSTAFA MAMAT, SITI S. ABAS, IBRAHIM M. SULAIMAN, AND SUKONO

ABSTRACT. The conjugate gradient method is among the efficient method for solving unconstrained optimization problems. In this paper, we propose a new formula for the conjugate gradient method based on the modification of the NPRP formula (Zhang, 2009). The proposed method satisfies the sufficient descent condition, and global convergence proof was established under some assumptions and strong Wolfe line search. Numerical results based on 98 test problems show that the new method very efficient as compared with the classical conjugate gradient method.

1. INTRODUCTION

We consider the following unconstrained optimization problems

$$\min\{f(\mathbf{x})|\mathbf{x}\in\mathbb{R}^n\} \tag{1.1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ is continuous and differentiable function. The conjugate gradient method is of an iterative method to solving (1.1) with formula as follows:

$$\mathbf{x}_k=\mathbf{x}_k+\alpha_k\mathbf{d}_k\ ,k=0,1,2,..., \tag{1.2}$$

where $\mathbf{x}_0$ is initial point, $\mathbf{x}_k$ is the point in kth iterative, $\mathbf{d}_k$ is the search direction, and α_k is the stepsize [1]. There are numerous methods used for calculation the

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 49M37, 65K10, 90C06.
Key words and phrases. Conjugate gradient method, Sufficient descent condition, Global convergence, Strong Wolfe line search.

stepsize, including the exact and inexact line search. In this paper, we apply the inexact line search to compute the stepsize α_k. The inexact line search that is often used in practice is the strong Wolfe line search. The strong Wolfe line search is defined as

$$\begin{cases} f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq f(\mathbf{x}_k) + \delta \alpha_k \mathbf{g}_k^T \mathbf{d}_k, \\ \left| \mathbf{g}\left(\mathbf{x}_k + \alpha_k \mathbf{d}_k\right)^T \mathbf{d}_k \right| \leq -\sigma \mathbf{g}_k^T \mathbf{d}_k \end{cases} \tag{1.3}$$

where $\mathbf{g}_k = \nabla f(\mathbf{x}_k)$ is a gradient of f at point $\mathbf{x}_k$, $\mathbf{g}_k^T$ is transpose of $\mathbf{g}_k$, and δ, σ are the parameters with value $0 < \delta < \sigma < 1$, see [2]. In conjugate gradient method, the search direction is defined by formula as follows:

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k, & k = 0 \\ -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, & k \geq 1 \end{cases} \tag{1.4}$$

where β_k is a scalar; we often say as conjugate formula. There are many parameters β_k known to date, including them

$$\beta_k^{FR} = \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2}, \ \beta_k^{CD} = -\frac{\|\mathbf{g}_k\|^2}{\mathbf{d}_k^T \mathbf{g}_{k-1}}, \ \beta_k^{DY} = \frac{\|\mathbf{g}_k\|^2}{\mathbf{d}_{k-1}^T \left(\mathbf{g}_k - \mathbf{g}_{k-1}\right)},$$

$$\beta_k^{PRP} = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \mathbf{g}_{k-1}\right)}{\|\mathbf{g}_{k-1}\|^2}, \ \beta_k^{WYL} = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_{k-1}\|} \mathbf{g}_{k-1}\right)}{\|\mathbf{g}_{k-1}\|^2},$$

$$\beta_k^{RMIL} = \frac{\mathbf{g}_k^T \left(\mathbf{g}_k - \mathbf{g}_{k-1}\right)}{\|\mathbf{d}_{k-1}\|^2},$$

where $\|.\|$ is the Euclidean norm of vectors. The above corresponding parameter are known as Fletcher-Reeves (FR) [3] method, Conjugate Descent (CD) method [4], Dai-Yuan (DY) method [5], Polak-Ribiére-Polyak (PRP) method [6], Wei-Yao-Liu (WYL) method [7], and Rivaie-Mustafa-Ismail-Leong (RMIL) method [8].

In the conjugate gradient method research, many researchers focus on its global convergence properties and descent condition. For the FR method, Zoutendijk has proved the global convergence properties under the exact line search [9]. As well as, Al-Baali also shows the FR method fulfill global convergence properties under inexact line search [10]. The CD method generated a descent search direction in each iteration for the parameter $\sigma < 1$ under the strong Wolfe line search, but its global convergence properties are not excellent. The DY method is a modification of the FR method, under the strong Wolfe line search the DY method fulfills the descent condition, but this method has bad numerical results.

Under strong Wolfe line search, Yuan and Stoer in [11] proved the PRP method has the global convergence properties and fulfills the descent condition. The WYL method is a modification of the PRP method; this method satisfies the descent condition and global convergence properties under an exact line search and strong Wolfe line search. The RMIL method is a modification of the PRP method by changing its denominator. Rivaie et al. proved the convergence properties of the RMIL method using an exact line search.

Based on the illustration above, in this article, we propose a new formula of conjugate gradient method β_k based on modification of NPRP method and we will compare the performance with other classic methods. The sufficient descent condition and global convergence properties of our new method are proved using the strong Wolfe line search.

2. New Formula and Algorithm

In 2009, Zhang [12] proposed a new conjugate gradient formula as follows:

$$\beta_k^{NPRP} = \frac{\|\mathbf{g}_k\|^2 - \frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_{k-1}\|}\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|}{\|\mathbf{g}_{k-1}\|^2} \tag{2.1}$$

that is, modification of the WYL method. In this section, we form a new formula with replace the term $\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_{k-1}\|}$ in the numerator (2.1) by $\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k-\mathbf{g}_{k-1}\|}$, add a negative $\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|$, extend the denominator by $(1-\mu)\|\mathbf{d}_{k-1}\|^2 + \mu\|\mathbf{g}_{k-1}\|^2$, and prevent negative value, so we define the new formulas as

$$\beta_k^{MMSSS2} = \begin{cases} A, & \text{if } B \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

where

$$A = \frac{\|\mathbf{g}_k\|^2 - \frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k-\mathbf{g}_{k-1}\|}\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right| - \left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|}{(1-\mu)\|\mathbf{d}_{k-1}\|^2 + \mu\|\mathbf{g}_{k-1}\|^2},$$

$$B = \|\mathbf{g}_k\|^2 > \left(\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k - \mathbf{g}_{k-1}\|} + 1\right)\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|,$$

$\mu = 0.6$, and MMSSS2 denotes Malik-Mustafa-Sabariah-Sulaiman-Sukono-the second.

Based on (1.2), (1.3), (1.4), and (2.2), we establish an algorithm of the MMSSS2 method as follows:

Algorithm 1. *(MMSSS2 method)*

Step 1. *Given a initial point* $\mathbf{x}_0$. *Choose value for stopping criteria* ϵ, *and parameter* σ, δ. *Set* $k = 0$.
Step 2. *Compute* $\|\mathbf{g}_k\|$, *if* $\|\mathbf{g}_k\| \leq \epsilon$, *then* $\mathbf{x}_k$ *is optimal point. Else, go to next step.*
Step 3. *Compute* β_k *using (2.2).*
Step 4. *Compute search direction* $\mathbf{d}_k$ *using (1.4).*
Step 5. *Compute stepsize* α_k *using (1.3).*
Step 6. *Set* $k := k + 1$ *and generate the next iteration* $\mathbf{x}_{k+1}$ *using (1.2).*
Step 7. *Go to Step 2.*

3. Analysis Convergence Under Strong Wolfe Line Search

In this section, we will analyze the sufficient descent condition and global convergence properties of the MMSSS2 method under the strong Wolfe line search. Further, the definition of the sufficient descent condition and global convergence properties [13] as follows are needed.

Definition 3.1. *Sufficient descent condition holds when there exist* $C > 0$ *such that*

$$\mathbf{g}_k^T \mathbf{d}_k \leq -C\|\mathbf{g}_k\|, \ \forall k \geq 0.$$

Definition 3.2. *The conjugate gradient method is global convergence if*

$$\lim_{k\to\infty} \inf \|\mathbf{g}_k\| = 0.$$

Next, we provide lemma and theorem, which are meaningful relationships to help prove the sufficient descent condition and global convergence properties of the MMSSS2 method.

Lemma 3.1. *The relation*

$$0 \leq \beta_k^{MMSSS2} \leq \frac{5}{2} \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{d}_{k-1}\|^2} \tag{3.1}$$

holds for all $k \geq 0$.

Proof. Based on (2.2), there are two cases.

Case 1. If $\|\mathbf{g}_k\|^2 \leq \left(\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k-\mathbf{g}_{k-1}\|}+1\right)\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|$, then $\beta_k^{MMSSS2}=0$.

Case 2. If $\|\mathbf{g}_k\|^2 > \left(\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k-\mathbf{g}_{k-1}\|}+1\right)\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|$, then $\beta_k^{MMSSS2}=A$. Since $\mu=0.6$ and $\mu\|\mathbf{g}_{k-1}\|^2>0$, then

$$\beta_k^{MMSSS2}=\frac{\|\mathbf{g}_k\|^2-\left(\frac{\|\mathbf{g}_k\|}{\|\mathbf{g}_k-\mathbf{g}_{k-1}\|}+1\right)\left|\mathbf{g}_k^T\mathbf{g}_{k-1}\right|}{0.4\|\mathbf{d}_{k-1}\|^2+0.6\|\mathbf{g}_{k-1}\|^2}\leq\frac{5}{2}\frac{\|\mathbf{g}_k\|^2}{\|\mathbf{d}_{k-1}\|^2},$$

and $\beta_k^{MMSSS2}>0$.

Hence, $0\leq\beta_k^{MMSSS2}\leq\frac{5}{2}\frac{\|\mathbf{g}_k\|^2}{\|\mathbf{d}_{k-1}\|^2}$. The proof is completed. □

Theorem 3.1. *Suppose sequences* $\{\mathbf{g}_k\}$ *and* $\{\mathbf{d}_k\}$ *be generated by Algorithm 1, with paramater* $0<\sigma<\frac{1}{80}$. *Then*

$$\frac{\|\mathbf{g}_k\|}{\|\mathbf{d}_k\|}<4,\ \forall k\geq 0\,. \tag{3.2}$$

Proof. By induction, for $k=0$, then from (1.4) we have $\mathbf{d}_0=-\mathbf{g}_0$ and obtained $\frac{\|\mathbf{g}_0\|}{\|\mathbf{d}_0\|}=-1<4$. So, (3.2) is true. Further, suppose (3.2) is true for $k=n$, so we have

$$\frac{\|\mathbf{g}_n\|}{\|\mathbf{d}_n\|}<4. \tag{3.3}$$

We will be proven for $k=n+1$, the (3.2) is true. Based on (1.4) and multiply by $\mathbf{g}_{n+1}^T$ we have

$$\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}+\mathbf{g}_{n+1}^T\mathbf{g}_{n+1}=\beta_{n+1}^{MMSSS2}\mathbf{g}_{n+1}^T\mathbf{d}_n$$

and then we will get it

$$\begin{aligned}\|\mathbf{g}_{n+1}\|^2 &= -\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}+\beta_k^{MMSSS2}\mathbf{g}_{n+1}^T\mathbf{d}_n\\ &\leq \left|\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}\right|+\left|\beta_k^{MMSSS2}\mathbf{g}_{n+1}^T\mathbf{d}_n\right|.\end{aligned} \tag{3.4}$$

From (3.4) and applied together (1.3), Lemma 3.1, and using Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}\|\mathbf{g}_{n+1}\|^2 &\leq \left|\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}\right|+\beta_k^{MMSSS2}\left|\mathbf{g}_n^T\mathbf{d}_n\right|\\ &\leq \|\mathbf{g}_{n+1}\|\|\mathbf{d}_{n+1}\|+\sigma\frac{5}{2}\frac{\|\mathbf{g}_{n+1}\|^2}{\|\mathbf{d}_n\|^2}\|\mathbf{g}_n\|\|\mathbf{d}_n\|.\end{aligned} \tag{3.5}$$

Dividing both sides of (3.5) by $\|\mathbf{g}_{n+1}\|$ and using (3.3), we get

$$\|\mathbf{g}_{n+1}\|<\|\mathbf{d}_{n+1}\|+10\sigma\|\mathbf{g}_{n+1}\|,$$

which means that $(1-10\sigma)\|\mathbf{g}_{n+1}\| < \|\mathbf{d}_{n+1}\|$. Since $0 < \sigma < \frac{1}{80}$, then $1-10\sigma > 0$, so we have

$$\frac{\|\mathbf{g}_{n+1}\|}{\|\mathbf{d}_{n+1}\|} < \frac{1}{1-10\sigma} < 4.$$

This shows that (3.2) is true for $k = n+1$. The proof is finished. □

The theorem below states that the MMSSS2 method satisfies the sufficient descent condition.

Theorem 3.2. *Suppose the sequences $\{\mathbf{g}_k\}$ and $\{\mathbf{d}_k\}$ be generated by Algorithm 1, with parameter $0 < \sigma < \frac{1}{80}$. Then*

$$\frac{-1}{1-40\sigma} < \frac{\mathbf{g}_k^T\mathbf{d}_k}{\|\mathbf{g}_k\|^2} < \frac{80\sigma-1}{1-40\sigma}, \ \forall k \geq 0. \tag{3.6}$$

Hence, the sufficient descent condition in Definition 3.1 holds.

Proof. The proof is by induction. For $k=0$ and using (1.4), we have $\mathbf{d}_0 = -\mathbf{g}_0$, and further

$$\frac{-1}{1-40\sigma} < \frac{\mathbf{g}_0^T\mathbf{d}_0}{\|\mathbf{g}_0\|^2} = \frac{-\mathbf{g}_0^T\mathbf{g}_0}{\|\mathbf{g}_0\|^2} = -1 < \frac{80\sigma-1}{1-40\sigma},$$

it states that (3.6) is true for $k=0$; furthermore, we get $\mathbf{g}_0^T\mathbf{d}_0 = -\|\mathbf{g}_0\|^2$, so based on Definition 3.1, the sufficient descent condition holds. Next, we prove for $k \geq 1$. Suppose that (3.6) is true for $k=n$, so we have

$$\frac{-1}{1-40\sigma} < \frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2} < \frac{80\sigma-1}{1-40\sigma}. \tag{3.7}$$

Furthermore, we need to proof that (3.6) is true for $k=n+1$. Rewriting (1.4) for $k=n+1$, we get

$$\mathbf{d}_{n+1} = -\mathbf{g}_{n+1} + \beta_{n+1}^{MMSSS2}\mathbf{d}_n.$$

Multiply the both sides by $\mathbf{g}_{n+1}^T$, we have

$$\mathbf{g}_{n+1}^T\mathbf{d}_{n+1} = -\|\mathbf{g}_{n+1}\|^2 + \beta_k^{MMSSS2}\mathbf{g}_{n+1}^T\mathbf{d}_n.$$

By dividing the both sides by $\|\mathbf{g}_{n+1}\|^2$, we obtain

$$\frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} = -1 + \beta_k^{MMSSS2}\frac{\mathbf{g}_{n+1}^T\mathbf{d}_n}{\|\mathbf{g}_{n+1}\|^2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{g}_n\|^2}. \tag{3.8}$$

Based on relation the strong Wolfe line search in (1.3), we get

$$\sigma\mathbf{g}_n^T\mathbf{d}_n \leq \mathbf{g}_{n+1}^T\mathbf{d}_n \leq -\sigma\mathbf{g}_n^T\mathbf{d}_n.$$

Since $\beta_k^{MMSSS2} \geq 0$, the relation above becomes

$$\sigma\beta_{n+1}^{MMSSS2}\mathbf{g}_n^T\mathbf{d}_n \leq \beta_{n+1}^{MMSSS2}\mathbf{g}_{n+1}^T\mathbf{d}_n \leq -\sigma\beta_k^{MMSSS2}\mathbf{g}_n^T\mathbf{d}_n. \tag{3.9}$$

Apply (3.8) and (3.9) together, we obtain

$$-1+\sigma\beta_{n+1}^{MMSSS2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{g}_{n+1}\|^2} \leq \frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} \leq -1-\sigma\beta_{n+1}^{MMSSS2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{g}_{n+1}\|^2}.$$

From Lemma 3.1, the inequality above becomes

$$-1+\sigma\frac{5\|\mathbf{g}_{n+1}\|^2}{2\|\mathbf{d}_n\|^2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{g}_{n+1}\|^2} \leq \frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} \leq -1-\sigma\frac{5\|\mathbf{g}_{n+1}\|^2}{2\|\mathbf{d}_n\|^2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{g}_{n+1}\|^2},$$

which implies

$$-1+\sigma\frac{5}{2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{d}_n\|^2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2} \leq \frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} \leq -1-\sigma\frac{5}{2}\frac{\|\mathbf{g}_n\|^2}{\|\mathbf{d}_n\|^2}\frac{\mathbf{g}_n^T\mathbf{d}_n}{\|\mathbf{g}_n\|^2}. \tag{3.10}$$

Combining (3.2), (3.7), and (3.10), we obtain

$$-1+\sigma\frac{5}{2}4^2\left(\frac{-1}{1-40\sigma}\right) < \frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} < -1-\sigma\frac{5}{2}4^2\left(\frac{-1}{1-40\sigma}\right).$$

Hence,

$$\frac{-1}{1-40\sigma} < \frac{\mathbf{g}_{n+1}^T\mathbf{d}_{n+1}}{\|\mathbf{g}_{n+1}\|^2} < \frac{80\sigma-1}{1-40\sigma}.$$

This shows that (3.6) is true for $k = n+1$. Denotes $c = \frac{80\sigma-1}{40\sigma-1}$, since $0 < \sigma < \frac{1}{80}$ then c is positive number. So that we have relation $\mathbf{g}_{n+1}^T\mathbf{d}_{n+1} < -c\|\mathbf{g}_{n+1}\|^2$, which indicate the sufficient descent condition holds. The proof is completed. □

To prove the convergence properties of the MMSSS2 method, we need the following assumptions for objective functions.

Assumption 1. *(A1) The objective function f has lower bond on the level set $\omega = \{\mathbf{x}|f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ where $\mathbf{x}_0$ is the initial point. (A2) In neighbourhood ω_0 of ω, the objective function f is continuously differentiable , and its gradient is Lipschitz continuous; then there exists a constant L such that $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \ \forall\, \mathbf{x}, \mathbf{y} \in \omega_0$.*

We also needed the following lemma, which was known as Zoutendijk condition. Proof of this lemma can be seen in [9].

Lemma 3.2. *Suppose Assumption 1 holds. Let a conjugate gradient method of the form (1.2) and (1.4), where α_k is calculated by strong Wolfe line search and search direction $\mathbf{d}_k$ satisfies the sufficient descent condition. Then,*

$$\sum_{k=0}^{\infty} \frac{\left(\mathbf{g}_k^T \mathbf{d}_k\right)^2}{\|\mathbf{d}_k\|^2} < \infty. \tag{3.11}$$

The following theorem shows that the MMSSS2 method is global convergence.

Theorem 3.3. *Suppose that Assumption 1 holds true, $\mathbf{x}_k$ is generated by Algorithm 1, $\mathbf{d}_k$ is obtained by formula (1.4), α_k is calculated under strong Wolfe line search (1.3), β_k is calculated by β_k^{MMSSS2}, and the sufficient descent condition hold true. Then,*

$$\lim_{k\to\infty} \inf \|\mathbf{g}_k\| = 0 \tag{3.12}$$

Proof. We use contradiction. Let (3.12) be not true, then there exist a constant $M > 0$ such that $\|\mathbf{g}_k\| \geq M$, further

$$\frac{1}{\|\mathbf{g}_k\|^2} \leq \frac{1}{M^2}, \ \forall k \geq 0.$$

Rewrite (1.4) as

$$\mathbf{d}_k + \mathbf{g}_k = \beta_k^{MMSSS2} \mathbf{d}_{k-1},$$

and squaring the both sides, we get

$$\|\mathbf{d}_k\|^2 = \left(\beta_k^{MMSSS2}\right)^2 \|\mathbf{d}_{k-1}\|^2 - 2\mathbf{g}_k^T \mathbf{d}_k - \|\mathbf{g}_k\|^2. \tag{3.13}$$

Combining (3.1), (3.6), and (3.13), we obtain

$$\begin{aligned} \|\mathbf{d}_k\|^2 &< \left(\beta_k^{MMSSS2}\right)^2 \|\mathbf{d}_{k-1}\|^2 + \left(\frac{2}{1-40\sigma}\right) \|\mathbf{g}_k\|^2 - \|\mathbf{g}_k\|^2 \\ &< \frac{25}{4} \frac{\|\mathbf{g}_k\|^4}{\|\mathbf{d}_{k-1}\|^2} + \left(\frac{1+40\sigma}{1-40\sigma}\right) \|\mathbf{g}_k\|^2. \end{aligned}$$

By dividing the both sides of relation above by $\|\mathbf{g}_k\|^4$, and applying (3.2), (3.13) together, we have

$$\begin{aligned} \frac{\|\mathbf{d}_k\|^2}{\|\mathbf{g}_k\|^4} &< \frac{25}{4} \frac{1}{\|\mathbf{d}_{k-1}\|^2} + \left(\frac{1+40\sigma}{1-40\sigma}\right) \frac{1}{\|\mathbf{g}_k\|^2} < \frac{100}{\|\mathbf{g}_{k-1}\|^2} + \left(\frac{1+40\sigma}{1-40\sigma}\right) \frac{1}{\|\mathbf{g}_k\|^2} \\ &< \left(100 + \frac{1+40\sigma}{1-40\sigma}\right) \frac{1}{M^2} = \frac{101-360\sigma}{1-40\sigma} \frac{1}{M^2}, \end{aligned}$$

that implies,

$$\frac{\|\mathbf{g}_k\|^4}{\|\mathbf{d}_k\|^2} = \frac{\left(\mathbf{g}_k^T\mathbf{d}_k\right)^2}{\|\mathbf{d}_k\|^2} > \frac{1-40\sigma}{101-360\sigma}M^2.$$

Furthermore,

$$\left(\sum_{k=0}^{\infty}\frac{\left(\mathbf{g}_k^T\mathbf{d}_k\right)^2}{\|\mathbf{d}_k\|^2} = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{\left(\mathbf{g}_k^T\mathbf{d}_k\right)^2}{\|\mathbf{d}_k\|^2}\right) > \left(\lim_{n\to\infty}(n+1)\frac{1-40\sigma}{101-360\sigma}\frac{1}{M^2} = \infty\right).$$

Hence,

$$\sum_{k=0}^{\infty}\frac{\left(\mathbf{g}_k^T\mathbf{d}_k\right)^2}{\|\mathbf{d}_k\|^2} > \infty,$$

this contradict with Zoutendijk condition (3.11). So (3.12) is true. Based on Definition 3.2, the MMSSS2 is global convergent. The proof is completed. □

4. Numerical Experiments

In this section, we conducted some numerical experiments to test the performance of the MMSSS2 method. To see the MMSSS2 method's performance, we used some test functions mostly considered from Andrei [14]. Some initial point variations are also suggested by Andrei [14], and the dimensional variations used are the same as in paper Malik et al. [15, 16]. Several test functions, dimensions, and initial points in 98 problems are stated in Table 1.

Numerical results are obtained by running programs written in Matlab R2019a software and using a personal laptop with specification processor Intel Core i7, 16 GB RAM, and operating system Windows 10 Pro 64 bit. In the program, we consider $\epsilon = 10^{-6}$, so the stopping criteria is $\|\mathrm{g}_k\| < 10^{-6}$ and we use parameters $\sigma = 0.001$ and $\delta = 0.0001$. Numerical results are said to fail if the number of iterations (NOI) exceeds 10,000 or never reaches the optimum value. In this paper, we will compare the performance of the MMSSS2 method with the RMIL method, FR method, CD method, DY method, WYL method, and NPRP method based on NOI and central processing unit (CPU) times. Summary of numerical results written in Table 2.

From the numerical results, we can determine the performance profile curve. The performance profile curve results using a performance profile introduced by Dolan and Moré [17]. In the performance profile figure, the $\rho_s(\tau)$ is the probability for solvers s at τ and $\tau(p,s)$ is computing time (NOI or CPU) needed to solve problem p by solver s. In general, solvers with high values of $\rho_s(\tau)$ or in the upper right of the curves represent the best solvers. The results of performance profiles in Figure 1 and Figure 2.

Table 1: List of the test functions, dimensions, and initial points

Problem	Test Function	Dimension	Initial point
1	Extended White & Holst	1000	(-1.2,1,...,-1.2,1)
2	Extended White & Holst	1000	(10,...,10)
3	Extended White & Holst	10000	(-1.2,1,...,-1.2,1)
4	Extended White & Holst	10000	(5,...,5)
5	Extended Rosenbrock	1000	(-1.2,1,...,-1.2,1)
6	Extended Rosenbrock	1000	(10,...,10)
7	Extended Rosenbrock	10000	(-1.2,1,...,-1.2,1)
8	Extended Rosenbrock	10000	(5,...,5)
9	Extended Freudenstein & Roth	4	(0.5,-2,0.5,-2)
10	Extended Freudenstein & Roth	4	(5,5,5,5)
11	Extended Beale	1000	(1,0.8,...,1,0.8)
12	Extended Beale	1000	(0.5,...,0.5)
13	Extended Beale	10000	(-1,...,-1)
14	Extended Beale	10000	(0.5,...,0.5)
15	Extended wood	4	(-3,-1,-3,-1)
16	Extended wood	4	(5,5,5,5)
17	Raydan 1	10	(1,...,1)
18	Raydan 1	10	(10,...,10)
19	Raydan 1	100	(-1,...,-1)
20	Raydan 1	100	(-10,...,-10)
21	Extended Tridiagonal 1	500	(2,...,2)
22	Extended Tridiagonal 1	500	(10,...,10)
23	Extended Tridiagonal 1	1000	(1,...,1)
24	Extended Tridiagonal 1	1000	(-10,...,-10)
25	Diagonal 4	500	(1,...,1)
26	Diagonal 4	500	(-20,...,-20)

(Continued on next page)

Table 1 – *Continued*

Problem	Test Function	Dimension	Initial point
27	Diagonal 4	1000	(1,...,1)
28	Diagonal 4	1000	(-30,...,-30)
29	Extended Himmelblau	1000	(1,...,1)
30	Extended Himmelblau	1000	(20,...,20)
31	Extended Himmelblau	10000	(-1,...,-1)
32	Extended Himmelblau	10000	(50,...,50)
33	FLETCHCR	10	(0,...,0)
34	FLETCHCR	10	(10,...,10)
35	Extended Powel	100	(3,-1,0,1,...,1)
36	Extended Powel	100	(5,...,5)
37	NONSCOMP	2	(3,3)
38	NONSCOMP	2	(10,10)
39	Extended DENSCHNB	10	(1,...,1)
40	Extended DENSCHNB	10	(10,...,10)
41	Extended DENSCHNB	100	(10,...,10)
42	Extended DENSCHNB	100	(-50,...,-50)
43	Extended Penalty	10	(1,2,3,...,10)
44	Extended Penalty	10	(-10,...,-10)
45	Extended Penalty	100	(5,...,5)
46	Extended Penalty	100	(10,...,10)
47	Hager	10	(1,...,1)
48	Hager	10	(-10,...,-10)
49	Extended Maratos	10	(1.1,0.1)
50	Extended Maratos	10	(-1,...,-1)
51	Six hump camel	2	(-1,2)
52	Six hump camel	2	(-5,10)
53	Three hump camel	2	(-1,2)
54	Three hump camel	2	(2,-1)
55	Booth	2	(5,5)
56	Booth	2	(10,10)
57	Trecanni	2	(-1,0.5)

(Continued on next page)

Table 1 – *Continued*

Problem	Test Function	Dimension	Initial point
58	Trecanni	2	(-5,10)
59	Zettl	2	(-1,2)
60	Zettl	2	(10,10)
61	Shallow	1000	(0,...,0)
62	Shallow	1000	(10,...,10)
63	Shallow	10000	(-1,...,-1)
64	Shallow	10000	(-10,...,-10)
65	Generalized Quartic	1000	(1,...,1)
66	Generalized Quartic	1000	(20,...,20)
67	Quadratic QF2	50	(0.5,...,0.5)
68	Quadratic QF2	50	(30,...,30)
69	Leon	2	(2,2)
70	Leon	2	(8,8)
71	Gen. Tridiagonal 1	10	(2,...,2)
72	Gen. Tridiagonal 1	10	(10,...,10)
73	Gen. Tridiagonal 2	4	(1,1,1,1)
74	Gen. Tridiagonal 2	4	(10,10,10,10)
75	POWER	10	(1,1,1,1)
76	POWER	10	(10,10,10,10)
77	Quadratic QF1	50	(1,...,1)
78	Quadratic QF1	50	(10,...,10)
79	Quadratic QF1	500	(1,...,1)
80	Quadratic QF1	500	(-5,...,-5)
81	Extended quadratic penalty QP2	100	(1,...,1)
82	Extended quadratic penalty QP2	100	(10,...,10)
83	Extended quadratic penalty QP2	500	(10,...,10)
84	Extended quadratic penalty QP2	500	(50,...,50)
85	Extended quadratic penalty QP1	4	(1,1,1,1)
86	Extended quadratic penalty QP1	4	(10,10,10,10)
87	Quartic	4	(10,10,10,10)
88	Quartic	4	(15,15,15,15)

(Continued on next page)

Table 1 – *Continued*

Problem	Test Function	Dimension	Initial point
89	Matyas	2	(1,1)
90	Matyas	2	(20,20)
91	Colville	4	(2,2,2,2)
92	Colville	4	(10,10,10,10)
93	Dixon and Price	3	(1,1,1)
94	Dixon and Price	3	(10,10,10)
95	Sphere	5000	(1,...,1)
96	Sphere	5000	(10,...,10)
97	Sum squares	50	(0,1,...,0,1)
98	Sum squares	50	(10,...,10)

TABLE 2. Summary of numerical results

Methods	Total of NOI	Total of CPU times	Successful
MMSSS2	4,675	4.8846	100%
RMIL	8,419	5.568807	89%
FR	35,402	28.7177298	93%
CD	37,031	26.830327	93%
DY	32,135	26.5542355	91%
WYL	69,374	157.0359284	97%
NPRP	9,625	7.1383309	96%

Figure 1 and Figure 2 below illustrate the performance profiles respect to NOI and CPU times, respectively. Both figures show that the MMSSS2 method curve is at the top of all curves; this indicates that the MMSSS2 method is the best solver. As well as in Table 2, we can see that the MMSSS2 method has a total of NOI and a total of CPU times a fewest compared to other methods, and successfully 100% solved 98 problems. Hence, the proposed methods very efficient as compared with the other methods.

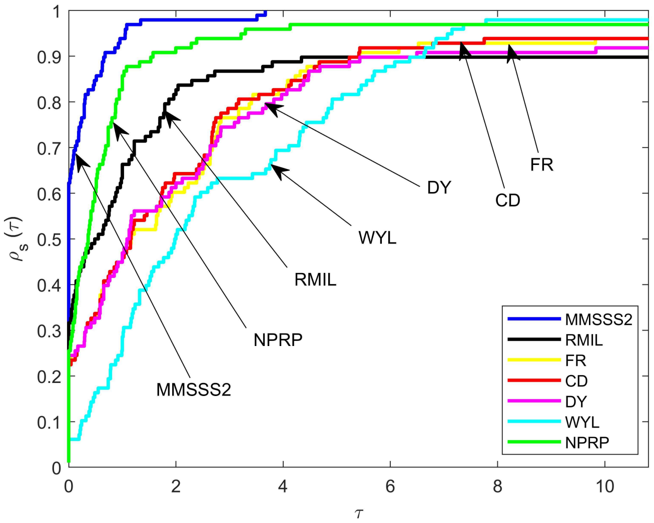

FIGURE 1. Performance profiles respect to NOI

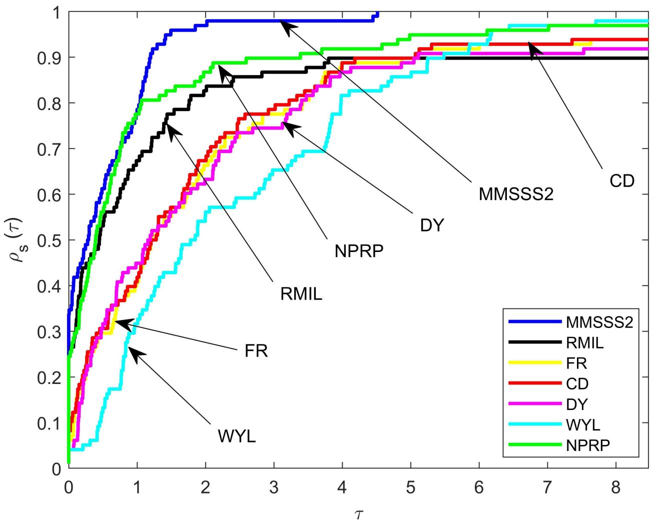

FIGURE 2. Performance profiles respect to CPU times

5. Conclusion

The main objective of this paper is to propose the new formula of the conjugate gradient method. The proposed method is the modification of the NPRP method, which we denote as the MMSSS2 method. The MMSSS2 method satisfies the sufficient descent condition and global convergence properties under the strong Wolfe line search with parameter $\sigma \in (0, \frac{1}{80})$. Based on 98 test problems, the numerical experiments have shown that the MMSSS2 method very efficient as compared with the RMIL method, FR method, CD method, DY method, WYL method, and NPRP method.

References

[1] J. Nocedal, S. J. Wright: *Numerical Optimization*, Springer Science and Business Media, New York, 2006.

[2] R. Pytlak: *Conjugate gradient algorithms in nonconvex optimization*, Springer Science and Business Media, 2008.

[3] R. Fletcher, C. M. Reeves: *Function minimization by conjugate gradients*, The Computer Journal, **7**(2) (1964), 149–154.

[4] R. Fletcher: *Practical methods of optimization*, John wiley and sons, 2013.

[5] Y. H. Dai, Y. Yuan: *A Nonlinear Conjugate Gradient Method with A Strong Global Convergence Property*, SIAM Journal on optimization, **10**(1) (1999), 177–182.

[6] E. Polak, G. Ribiére: *Note sur la convergence de méthodes de directions conjuguées*, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, **3**(R1) (1969), 35–43.

[7] Z. Wei, S. Yao, L. Liu: *The convergence properties of some new conjugate gradient methods*, Applied Mathematics and Computation, **183**(2) (2006), 1341–1350.

[8] M. Rivaie, M. Mamat, L. W. June, I. Mohd: *A new class of nonlinear conjugate gradient coefficients with global convergence properties*, Appl. Math. Comput., **218**(22) (2012), 11323–11332.

[9] G. Zoutendijk: *Nonlinear programming, computational methods*, Integer and nonlinear programming, (1970), 37–86.

[10] M. Al-Baali: *Descent property and global convergence of the Fletcher–Reeves method with inexact line search, Journal of Numerical Analysis*, **5**(1) (1985), 121–124.

[11] Y. X. Yuan, J. Stoer: *A subspace study on conjugate gradient algorithms*, ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, **75**(1) (1995), 69–77.

[12] L. Zhang: *An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation*, Applied Mathematics and Computation, **215**(6) (2009), 2269–2274.

[13] Y. H. DAI: *Nonlinear conjugate gradient methods*, Wiley Encyclopedia of Operations Research and Management Science, 2010.

[14] N. ANDREI: *An unconstrained optimization test functions collection*, Adv. Model. Optim, **10**(1) (2008), 147–161.

[15] M. MALIK, M. MAMAT, S. S. ABAS, SUKONO: *Convergence analysis of a new coefficient conjugate gradient method under exact line search*, International Journal of Advanced Science and Technology, **29**(5) (2020), 187–198.

[16] M. MALIK, S. S. ABAS, M. MAMAT, SUKONO, I. S. MOHAMMED: *A new hybrid conjugate gradient method with global convergence properties*, International Journal of Advanced Science and Technology, **29**(5) (2020), 199–210.

[17] E. D. DOLAN, J. J. MORÉ: *Benchmarking optimization software with performance profiles*, Mathematical Programming, **91**(2) (2002), 201–213.

FACULTY OF INFORMATICS AND COMPUTING
UNIVERSITI SULTAN ZAINAL ABIDIN (UNISZA)
TERENGGANU 22200, MALAYSIA
DEPARTMENT OF MATHEMATICS
UNIVERSITAS INDONESIA (UI)
DEPOK 16424, INDONESIA
E-mail address: m.malik@sci.ui.ac.id

FACULTY OF INFORMATICS AND COMPUTING
UNIVERSITI SULTAN ZAINAL ABIDIN (UNISZA)
TERENGGANU 22200, MALAYSIA
E-mail address: must@unisza.edu.my

FACULTY OF INFORMATICS AND COMPUTING
UNIVERSITI SULTAN ZAINAL ABIDIN (UNISZA)
TERENGGANU 22200, MALAYSIA
E-mail address: sabariahabas@unisza.edu.my

FACULTY OF INFORMATICS AND COMPUTING
UNIVERSITI SULTAN ZAINAL ABIDIN (UNISZA)
TERENGGANU 22200, MALAYSIA
E-mail address: sulaimanib@unisza.edu.my

DEPARTMENT OF MATHEMATICS
UNIVERSITAS PADJADJARAN (UNPAD)
JATINANGOR 45361, INDONESIA
E-mail address: sukono@unpad.ac.id