Bulletin of the *Transilvania* University of Braşov
Series III: Mathematics and Computer Science, Vol. 5(67), No. 2 - 2025, 179-192
https://doi.org/10.31926/but.mif.2025.5.67.2.14

A NEW HYBRID CONJUGATE GRADIENT METHOD AS A CONVEX COMBINATION METHODS

Mehamdia ABDELHAMID[*,1], Tahar BECHOUAT[2] and Yacine CHAIB [3]

Abstract

The conjugate gradient (CG) method is a widely employed algorithm for solving large-scale unconstrained optimization problems due to its fast convergence and efficient memory usage. In this paper, we suggest a new hybrid nonlinear conjugate gradient method, which the conjugate gradient coefficient β_k is a convex combination of β_k^{NPRP} and β_k^{DY}. The parameter θ_k is computed in such a way that the conjugacy condition is satisfied. With the strong Wolfe line search, the descent property and global convergence of the new hybrid method are proved. The numerical results also show that our method is robust and efficient.

2000 *Mathematics Subject Classification:* 90C25, 90C30.
Key words: hybrid conjugate gradient method, sufficient descent direction, global convergence, numerical comparisons.

1 Introduction

The conjugate gradient method is a very important and efficient technique for solving large-scale nonlinear optimization due to the simplicity of their iteration, very low memory requirements, and good convergence analysis In this work, we consider the unconstrained optimization problem

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous and differentiable function bounded from below. The conjugate gradient method to solve the problem (1.1) starts from an initial point $x_0 \in \mathbb{R}^n$. It generates a sequence $\{x_k\}_{k\geq 0}$, such that:

[1*] *Corresponding author*, Universite M'Hamed Bougara de Boumerdes, Algeria, e-mail: mehamdiaabdelhamid56@gmail.com
[2]Laboratory Informatics and Mathematics, Mohamed Cherif Messaadia University, Souk Ahras, Algeria, e-mail: t.bechouat@gmail.com
[3]Laboratory Informatics and Mathematics, Mohamed Cherif Messaadia University, Souk Ahras, Algeria, e-mail: y.chaib@univ-soukahras.dz

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

where x_k is the current iteration point, the stepsize α_k is a positive scalar determined by some line search, and d_k is the search direction defined by the following formula

$$d_0 = -g_0;\; d_{k+1} = -g_{k+1} + \beta_k d_k, \tag{1.3}$$

where $g_{k+1} = \nabla f(x_{k+1})$ is the gradient of f at x_{k+1} and the parameter β_k is a known as the conjugate gradient coefficient.

The steplength α_k is very important for global convergence of conjugate gradient methods. The most used line search conditions for the stepsize determination are the so called standard Wolfe line search conditions:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \tag{1.4}$$

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k, \tag{1.5}$$

where $0 < \delta < \sigma < 1$. The first condition (1.4), called the Armijo condition, ensures a sufficient reduction of the objective function value, while the second condition (1.5), called the curvature condition, ensures unacceptable short stepsizes. It is worth mentioning that a stepsize computed by the Wolfe line search conditions (1.4) and (1.5) may not be sufficiently close to a minimizer of $f(x_k + \alpha d_k),\ \alpha > 0$. In these situations, the strong Wolfe line search conditions may be used, which consist of (1.4), and instead of (1.5), the following strengthened version

$$\left|g_{k+1}^T d_k\right| \leq -\sigma g_k^T d_k, \tag{1.6}$$

is used. From (1.6), we see that if $\sigma \to 0$, then the stepsize which satisfies (1.4) and (1.6) tends to be the optimal stepsize. Observe that if a stepsize α_k satisfies the strong Wolfe line search, then it satisfies the standard Wolfe conditions. Now, we denote $y_k = g_{k+1} - g_k$ and $\|.\|$ the Euclidean norm. In Table 1, we will give some famous formulas of the parameter β_k:

Parameter β_k	**References**	**Properties**
$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T d_k}$	Hestenes-Stiefel [14]	In general, may not be convergent, but usually they have better numerical results.
$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\Vert g_k \Vert^2}$	Polak-R-Polyak [19, 20]	
$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T d_k}$	Liu -Storey [18]	
$\beta_k^{FR} = \frac{\Vert g_{k+1} \Vert^2}{\Vert g_k \Vert^2}$	Fletcher-Reeves [12]	They have strong convergent properties, but they may not performwell in practicedue to jamming [3] and [4].
$\beta_k^{CD} = \frac{\Vert g_{k+1} \Vert^2}{-g_k^T d_k}$	Conjugate Descent [13]	
$\beta_k^{DY} = \frac{\Vert g_{k+1} \Vert^2}{y_k^T d_k}$	Dai-Yuan [8]	

Table 1 List of the conjugate gradient coefficients famous

Dai and Yuan [8] proved that the DY method always generates descent directions and converges globally with the standard Wolfe line conditions (1.4) and (1.5). On the other hand, the PRP method has erratic global convergence properties. If function f is strongly convex and the line search is exact, then Polak and Ribière [19] and Polyak [20] established the global convergence of the PRP method. If $g_{k+1}^T g_k > 0$ and $\Vert g_{k+1} \Vert^2 > g_{k+1}^T g_k$ we have $0 \leq \beta_k^{PRP} \leq \beta_k^{FR}$ and the PRP method has good convergence and theoretical properties. In 1992, Gilbert and Nocedal [15] established a convergence result for a modified PRP method with the following parameter

$$\beta_k^{PRP+} = max\left\{0,\ \beta_k^{PRP}\right\}.$$

In recent years, based on the above six formulas and their hybridization, many works have devoted their time and effort to come up with new formulae to increase the efficiency and effectiveness of the PRP method. Wei et al. [22] gave a variant of the PRP method, the WYL method, where the parameter β_k is yielded by

$$\beta_k^{WYL} = \frac{\Vert g_{k+1} \Vert^2 - \frac{\Vert g_{k+1} \Vert}{\Vert g_k \Vert} g_{k+1}^T g_k}{\Vert g_k \Vert^2}.$$

Huang et al. [16] proved that the WLY method satisfies the sufficient descent condition and converges globally under the strong Wolfe line search with the parameter $\sigma < \frac{1}{4.}$. Zhang [23] took a little modification to the WYL method and constructed the NPRP method as follows

$$\beta_k^{NPRP} = \frac{\Vert g_{k+1} \Vert^2 - \frac{\Vert g_{k+1} \Vert}{\Vert g_k \Vert} \left| g_{k+1}^T g_k \right|}{\Vert g_k \Vert^2}.$$

The NPRP method satisfies the sufficient descent condition and converges globally if the strong Wolfe line search is used and the parameter σ is restricted in $\left(0, \frac{1}{2}\right)$

[23]. The hybrid conjugate gradient method considered is to combine the standard conjugate gradient methods in a tow distinct ways. The first class is based on the projection concept. Recently, Touati-Ahmed and Storey [21] introduced the first hybrid conjugate gradient algorithm, where the parameter β_k is computed as

$$\beta_k^{TaS} = \min\left\{\beta_k^{FR},\ \beta_k^{PRP}\right\}.$$

The authors proved that β_k^{TaS} has good convergence properties and numerically outperforms both the β_k^{FR} and β_k^{PRP} algorithms. Soon afterward, Hu and Storey [17] introduced another hybrid conjugate gradient method with the β_k computed as

$$\beta_k^{HuS} = max\left\{0,\ \min\left\{\beta_k^{FR},\ \beta_k^{PRP}\right\}\right\}.$$

Also, Gilbert and Nocedal [15] proposed a combination of PRP and FR methods is computed as

$$\beta_k^{GN} = max\left\{-\beta_k^{FR},\ \min\left\{\beta_k^{FR},\ \beta_k^{PRP}\right\}\right\}.$$

Since β_k^{FR} is always nonnegative, it follows that β_k^{GN} can be negative. The method of Gilbert and Nocedal [15] has the same advantage of avoiding jamming. Dai and Yuan [9] combined the DY method with the HS method, proposing the following two-hybrid methods

$$\beta_k^{hDY} = \max\left\{-c\beta_k^{DY},\ \min\left\{\beta_k^{HS},\ \beta_k^{DY}\right\}\right\},$$

$$\beta_k^{hDYz} = \max\left\{0,\ \min\left\{\beta_k^{HS},\ \beta_k^{DY}\right\}\right\},$$

where $c = \frac{1-\sigma}{1+\sigma}$. For the standard Wolfe conditions (1.4) and (1.5), under the Lipschitz continuity of the gradient, Dai and Yuan [9] established the global convergence of these hybrid computational schemes.

The second class of hybrid conjugate gradient methods is based on the convex combination of the standard methods. Recently, Andrei [2] introduced the first hybrid conjugate gradient methods based on HS and DY methods (denoted as HYBRID method) for solving unconstrained optimization problems (1.1), calculating the parameter β_k^c as a convex combination of β_k^{HS} and β_k^{DY} i.e:

$$\beta_k^c = (1-\theta_k)\,\beta_k^{HS} + \theta_k\beta_k^{DY},$$

where θ_k is a scalar parameter satisfying $0 \le \theta_k \le 1$. Convergence with the standard Wolfe conditions was established and numerical results show that this hybrid computational scheme outperforms the Hestenes-Stiefel and the Dai-Yuan conjugate gradient algorithms. In 2009, this author also studied the global convergence of the CCOMB method [5] under strong Wolfe line search, such that the parameter β_k as a convex combination of β_k^{PRP} and β_k^{DY} i.e:

$$\beta_k^N = (1-\theta_k)\,\beta_k^{PRP} + \theta_k\beta_k^{DY}.$$

Djordjevic in [10] introduced the hybridization of LS and CD by their convex combination, which calls the HLSCD method, such that

$$\beta_k^{HLSCD} = (1-\theta_k)\,\beta_k^{LS} + \theta_k \beta_k^{CD}.$$

The compilation of parameter θ_k in β_k^{HLSCD} is in such a way that the conjugacy condition is satisfied. The global convergence of this method is proved under the strong Wolfe line search without convexity assumption on the objective function. The aim of this paper is to propose new hybrid conjugate gradient as a convex combination of NPRP and DY conjugate gradient algorithms. We establish, under a strong Wolfe line search, convergence properties of the proposed CGM. Numerical results show that the new method is efficient and robust, and outperforms as six CGMs algorithms famous. Now, we will organize our work as follows. In the next section, we consist of a new hybrid method and determine the parameter θ_k. Also, we present the specific algorithm and we prove the sufficient descent condition. In section 3, we prove the global convergence of the proposed method with a strong Wolfe line search. The numerical results are contained in section 4. Finally, we make a summary of our paper.

1.1 Convex combination method

In this section, we combine NPRP and DY methods to get hNPRPDY method. The parameter β_k in the presented method, denoted as $\beta_k^{hNPRPDY}$, is computed as a convex combination of β_k^{NPRP} and β_k^{DY}, i.e :

$$\beta_k^{hNPRPDY} = (1-\theta_k)\,\beta_k^{NPRP} + \theta_k \beta_k^{DY}, \tag{2.1}$$

where θ_k is a scalar parameter satisfying $0 \le \theta_k \le 1$, which follows to be determined. It is obvious that if $\theta_k = 0$, then $\beta_k^{hNPRPDY} = \beta_k^{NPRP}$ and if $\theta_k = 1$, then $\beta_k^{hNPRPDY} = \beta_k^{DY}$. On the other hand, if $0 < \theta_k < 1$, then $\beta_k^{hNPRPDY}$ is a convex combination appropriate parameters β_k^{NPRP} and β_k^{DY}. The search direction d_k of our algorithm is computed by

$$d_0 = -g_0,\; d_{k+1} = -g_{k+1} + \beta_k^{hNPRPDY} d_k. \tag{2.2}$$

The traditional conjugacy condition $d_{k+1}^T y_k = 0$ plays an important role in the convergence analyses and numerical calculation. The coefficients θ_k is chosen in such a way that the conjugation condition $d_{k+1}^T y_k = 0$ is satisfied. Indeed, we multiply the two sides of the relation (2.2) through y_k^T, we obtain so

$$\theta_k = \frac{(y_k^T g_{k+1} - \beta_k^{NPRP} y_k^T d_k)}{\beta_k^{DY} y_k^T d_k - \beta_k^{NPRP} y_k^T d_k}.$$

After simplification, we find

$$\theta_k = \frac{\varepsilon - \eta}{\mu - \eta}, \tag{2.3}$$

where
$\varepsilon = \|g_{k+1}\|^2\, y_k^T g_{k+1}$, $\mu = \|g_{k+1}\|^2\, \|g_k\|^2$ and $\eta = \left(\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|}\left|g_{k+1}^T g_k\right|\right) y_k^T d_k$.

Observe that the parameter θ_k given by (2.3) can be outside the interval $[0,1]$. However, in order to have a real convex combination in (2.1) the following rule is considered:

if $\theta_k < 0$, then set $\theta_k = 0$ in (2.1), i.e. $\beta_k^{hNPRPDY} = \beta_k^{NPRP}$; if $\theta_k > 1$, then set $\theta_k = 1$ in (2.1), i.e. $\beta_k^{hNPRPDY} = \beta_k^{DY}$. Therefore, under this rule for θ_k selection, the direction d_{k+1} in (2.2) combines the properties of NPRP and DY algorithms. Using $\beta_k^{hNPRPDY}$ and d_{k+1} as defined in (2.1) and (2.2) respectively, we now present hNPRPDY algorithm.

Step 1: Initialization:

Choose an initial point $x_0 \in \mathbb{R}^n$ and the parameters $0 < \delta < \sigma < \frac{1}{2}$. Compute $f(x_0)$ and g_0. Set $d_0 = -g_0$.

Step 2: Test for a continuation of iterations.

Si $\|g_k\|_\infty \leq 10^{-6}$, then stop. Otherwise, go to the next step.

Step 3: Line search:

Calculate α_k satisfies the linear search conditions of strong Wolfe (1.4) and (1.6) and update the variables

$$x_{k+1} = x_k + \alpha_k d_k.$$

Step 4: Compute θ_k

If $\mu - \eta = 0$, then set $\theta_k = 0$, else find θ_k from (2.3).

Step 5: $\beta_k^{hNPRPDY}$ conjugate gradient parameter computation

If $0 < \theta_k < 1$, then calculate $\beta_k^{hNPRPDY}$ as in (2.1). If $\theta_k \geq 1$, then set $\beta_k^{hNPRPDY} = \beta_k^{DY}$. If $\theta_k \leq 0$, then set $\beta_k^{hNPRPDY} = \beta_k^{NPRP}$.

Step 6. Compute the search direction.

Generate $d_{k+1} = -g_{k+1} + \beta_k^{hNPRPDY} d_k$.

Step 7: Set $k = k+1$ and go to Step 2.

1.2 Sufficient descent condition

Now, we prove that the search direction d_k obtained by the new hybrid conjugate gradient method satisfies in some conditions the sufficient descent condition.

Theorem 1. *Let the sequences $\{d_k\}$ and $\{g_k\}$ be generated by hNPRPDY algorithm. Then, the search direction d_k satisfies the sufficient descent direction.*

$$g_k^T d_k \leq -c \|g_k\|^2, \ \forall\ k \geqslant 0. \tag{2.4}$$

This proof is based on the following intermediate results.

Lemma 1. *Suppose that the conjugate gradient NPRP is implemented with the strong Wolfe line search (1.4) and (1.6) where $0 < \sigma < \frac{1}{2}$. Then, the NPRP method generates descent directions d_k satisfying the following inequalities:*

$$g_k^T d_k \leq -c_1 \|g_k\|^2, \ \forall k \geq 0. \tag{2.5}$$

where $c_1 = \frac{1-2\sigma}{1-\sigma}$.

Proof. The proof of this lemma is similar to the proof of theorem 2.2 of Zhang [23].

□

Lemma 2. *Consider any method* (1.2) *and* (1.3) , *where* $\beta_k = \beta_k^{DY}$ *with the strong Wolfe line search (1.4) and (1.6) where* $0 < \sigma < \frac{1}{2}$. *We get*

$$g_k^T d_k \leq -\frac{1-2\sigma}{1-\sigma} \|g_k\|^2, \text{ for all } k \geq 0. \tag{2.6}$$

Proof. The proof is given by induction as follows. For $k = 0$, $g_0^T d_0 = -\|g_0\|^2$, we conclude that the sufficient descent condition holds for $k = 0$. Now, we assume (2.6) holds for k and prove that for $k+1$. The search direction d_k of method DY (see Dai and Yuan [8]) is computed by

$$d_{k+1}^{DY} = -g_{k+1} + \beta_k^{DY} d_k. \tag{2.7}$$

Multiplying (2.7) by g_{k+1}^T from the left, we get

$$g_{k+1}^T d_{k+1}^{DY} = -\|g_{k+1}\|^2 + \|g_{k+1}\|^2 \frac{g_{k+1}^T d_k}{y_k^T d_k}.$$

From (1.6) , we have

$$d_k^T y_k \geq (1-\sigma)\left(-g_k^T d_k\right). \tag{2.8}$$

Using the line search condition (1.6) and (2.8) , we obtain

$$g_{k+1}^T d_{k+1}^{DY} \leq -c_1 \|g_{k+1}\|^2.$$

The result can be achieved.

□

Proof of Theorem 1. From (2.1) and (2.2) , we have

$$d_{k+1} = -g_{k+1} + ((1-\theta_k)\beta_k^{NPRP} + \theta_k \beta_k^{DY}) d_k.$$

Thus, we can obtain

$$d_{k+1} = (1-\theta_k) d_{k+1}^{NPRP} + \theta_k d_{k+1}^{DY}. \tag{2.9}$$

Multiplying (2.9) by g_{k+1}^T from the left, we get

$$g_{k+1}^T d_{k+1} = (1-\theta_k) g_{k+1}^T d_{k+1}^{NPRP} + \theta_k g_{k+1}^T d_{k+1}^{DY}. \tag{2.10}$$

Now, we assume that $0 < \theta_k < 1$, i.e. there are two constants a_1 and a_2 positive such as: $0 < a_1 \leq \theta_k \leq a_2 < 1$. By two relations $g_{k+1}^T d_{k+1}^{DY} \leq -c_1 \|g_{k+1}\|^2 \leq 0$ and $g_{k+1}^T d_{k+1}^{NPRP} \leq -c_1 \|g_{k+1}\|^2 \leq 0$, we obtain

$$a_2 g_{k+1}^T d_{k+1}^{DY} \leq \theta_k g_{k+1}^T d_{k+1}^{DY} \leq a_1 g_{k+1}^T d_{k+1}^{DY},$$

and

$$(1-a_1)\, g_{k+1}^T d_{k+1}^{NPRP} \leq (1-\theta_k)\, g_{k+1}^T d_{k+1}^{NPRP} \leq (1-a_2)\, g_{k+1}^T d_{k+1}^{NPRP}.$$

From (2.10), we conclude:

$$g_{k+1}^T d_{k+1} \leq a_1 g_{k+1}^T d_{k+1}^{DY} + (1-a_2)\, g_{k+1}^T d_{k+1}^{NPRP}. \tag{2.11}$$

Then, lemma 1 and lemma 2 we get

$$g_{k+1}^T d_{k+1} \leq -c \left\| g_{k+1} \right\|^2,$$

where $c = (a_1 + 1 - a_2)\, c_1$.

1.3 Global convergence

In this section, it is assumed that $g_k \neq 0$ for all $k \geq 0$, otherwise a stationary point is found To establish the global convergence of our method, we need the following basic assumptions on the objective function.
Assumption H1. The level set

$$\Lambda = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\},$$

is bounded.
Assumption H2. In some open convex neighborhood $\mathcal{N}$ of Λ, the function f is continuously differentiable and its gradient is Lipschitz continuous, namely, there exists a constant $L > 0$ such that:

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \forall x, y \in \mathcal{N}. \tag{3.1}$$

These assumptions imply that there exists a positive constant $\Gamma \geq 0$, we refer the reader to [4, 10, 13, 15, 22], such that

$$\| \nabla f(x) \| \leq \Gamma, \text{ for all } x \in \mathcal{N}. \tag{3.2}$$

The following lemma is necessary to prove the global convergence of our proposed method

Lemma 3. *Let assumptions H1 and H2 hold. Consider the method* (1.2) *and* (1.3)*, where d_k is a descent direction, and α_k is obtained by the strong Wolfe line search. If*

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} = \infty, \tag{3.3}$$

then

$$\lim_{k \to \infty} \inf \|g_k\| = 0.$$

Proof. This result was demonstrated by Dai et al in [7]. □

We need also this lemma to prove the convergence of our method.

Lemma 4. *Let assumptions H1 and H2 hold and the sequence* $\{x_k\}$ *be obtained by hNPRPDY method,* α_k *satisfies the strong Wolfe conditions* (1.4) *and* (1.6). *Then*

$$\alpha_k \geq \frac{(1-\sigma) \mid g_k^T d_k \mid}{L \parallel d_k \parallel^2}. \tag{3.4}$$

Proof. From (1.6), we have

$$(g_{k+1}^T d_k - g_k^T d_k) \geq (\sigma - 1) g_k^T d_k.$$

Using the Cauchy Schwarz inequality and (3.1), it holds that

$$(\sigma - 1) g_k^T d_k \leq (g_{k+1} - g_k)^T d_k \leq L\alpha_k \parallel d_k \parallel^2 .$$

By combining these two inequalities, the result can be achieved.

□

Remark 1. *Assuming the beginning of this section and* (2.4), *it is easy to obtain that* $g_k^T d_k \neq 0$ *for all* $k \geq 0$. *Suppose* $\alpha_k = 0$, *by* (1.6) *and the sufficient descent direction of hNPRPDY method, we get*

$$-g_k^T d_k \leq -\sigma g_k^T d_k,$$

hence

$$\sigma \geq 1.$$

This is a contradiction with $0 < \sigma < 1$. *This indicates that* α_k *obtained in the hNPRPDY method is not equal to zero, i.e., there exists a constant* $\lambda > 0$ *such that*

$$\alpha_k \geq \lambda \quad \textit{for all } k \geq 0. \tag{3.5}$$

The following theorem establishes to global convergence of hNPRPDY method with the strong Wolfe line search.

Theorem 2. *Suppose that assumptions H1 and H2 hold. Consider any conjugate gradient method in the form* (1.2) *and* (1.3), *with the conjugate gradient parameter* β_k *defined by* (2.1), *in which the steplength* α_k *is determined to satisfy the strong Wolfe conditions* (1.4) *and* (1.6). *If the search directions satisfy the decent condition* (2.4). *Then this method converges in the sense that*

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{3.6}$$

Proof. *We prove by contradiction and assume that there exists a positive constant* γ *such that*

$$\|g_k\| \geq \gamma \quad \text{for } k \quad \text{enough large.} \tag{3.7}$$

We have since the definition of β_k^{NPRP} *and Cauchy Schwarz inequality, that*

$$0 = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|\|g_{k+1}\|\|g_k\|}{\|g_k\|}}{\|g_k\|^2} \leq \beta_k^{NPRP} \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2}.$$

Using (3.2) *and* (3.7) , *we get*

$$0 \leq \beta_k^{NPRP} \leq \frac{\Gamma^2}{\gamma^2}. \tag{3.8}$$

On the other hand, by using (1.6) , (2.6) *and* (3.7)

$$d_k^T y_k = d_k^T (g_{k+1} - g_k) \geq (1-\sigma)\left(-d_k^T g_k\right) \geq c_1(1-\sigma)\gamma^2. \tag{3.9}$$

Using the definition of β_k^{DY}, (3.2) *and* (3.9), *we have*

$$0 \leq \beta_k^{DY} \leq \frac{\Gamma^2}{c_1(1-\sigma)\gamma^2}. \tag{3.10}$$

Since $0 \leq \theta_k \leq 1$, *from* (2.1), (3.8) *and* (3.10) *we have*

$$0 \leq \beta_k^{hNPRPDY} \leq \beta_k^{NPRP} + \beta_k^{DY} \leq E,$$

where

$$E = \frac{\Gamma^2}{c(1-\sigma)\gamma^2} + \frac{\Gamma^2}{\gamma^2}.$$

Thus, it follows from (1.2), (2.2), (3.5) *and* (3.11) *that*

$$\|d_{k+1}\| \leq \| g_{k+1} \| + \beta_k^{hNPRPDY} \frac{\| x_{k+1} - x_k \|}{\alpha_k} \leq M, \tag{3.12}$$

where

$$M = \Gamma + E\frac{D}{\lambda},$$

and

$$D = \max\left\{\|y - z\| : y, z \in \Lambda\right\}.$$

By taking the summation $k \geq 0$

$$\sum_{k\geq 0} \frac{1}{\|d_{k+1}\|^2} = \infty.$$

So, applying lemma 3, we conclude that (3.6) *is true. This is a contradiction with* (3.7), *so we have proved* (3.6).

1.4 Numerical experiments

In this section, we present some numerical experiments obtained with the new proposed conjugate gradient method with the hybridization parameter β_k given by (2.1). The test problems have been taken to the CUTE library [1, 6]. All the algorithms have been coded in MATLAB 2013 and compiler settings on the PC machine (2.5 GHz, 3.8 GB RAM) with Windows XP operating system. We compare the computational results of our method (hNPRPDY method) against the DY [8], PRP [19, 20], hDYz [9], CCOMB [5], HLSCD [10] and HuS [17] methods. In this numerical result, all algorithms implement the strong Wolfe line search conditions with $\delta = 10^{-3}$ and $\sigma = 10^{-1}$. The iteration is terminated if one of the following conditions is satisfied (i) $\|g_k\|_\infty < 10^{-6}$, where $\|.\|_\infty$ is the maximum absolute component of a vector, (ii) The number of iterations exceeded 2000, (iii) The computing time is more than 500 s. We show the performance difference clearly between our hNPRPDY method and six conjugate gradient algorithms. We chose the performance profile introduced by Dolan and Morè [11] to compare the performance according to the number of iterations and CPU time to rule as follows. Let S be the set of methods and P is the set of the test problems with n_p, n_s is the number of the test problems and the number of the methods, respectively. For each problem $p \in P$ and solver $s \in S$, denote $\tau_{p,s}$ be the computing time (the number of iterations or CPU time) required to solve problems $p \in P$ by solver $s \in S$. Then comparison between different solvers based on the performance ratio is given by

$$r_{p,s} = \frac{\tau_{p,s}}{\min\{\tau_{p,i},\ 1 \leq i \leq n_s\}}.$$

Suppose that a parameter $r_M \geq r_{p,s}$ for all problems and solvers chosen, and $r_M = r_{p,s}$ if and only if solver s does not solve problem p. The overall evaluation of the performance of the solvers is then given by the performance profile function given by

$$F_s(t) = \frac{size\{p:\ 1 \leq p \leq n_p,\ r_{p,s} \leq t\}}{n_p},$$

where $t \geq 1$ and $size\ A$ is the number of elements in the set A.

The performance profile $F_s : [1, \infty[\to [0, 1]$ for a solver s is a nondecreasing, piecewise constant function, continuous from the right at each breakpoint. $F_s(t)$ is the probability for solver $s \in S$ so that the performance ratio $r_{p,s}$ is within a factor $t \geq 1$ of the best possible ratio. The function F_s is the cumulative distribution function for the performance ratio. Observe that $1 - F_s(t)$ is the fraction of problems that the solver cannot solve within a factor t of the best solver. The value of $F_s(1)$ is the probability that the solver will win the rest of the solvers.

Figure 1 and Figure 2 give a performance comparison of the hNPRPDY method with those methods for the CPU time and the number of iterations, respectively. From these Figures, we can conclude that the hNPRPDY method performs better than DY [8], PRP [19, 20], hDYz [9], CCOMB [5], HLSCD [10] and HuS [17] methods, for the given test problems. These obtained preliminary results are indeed encouraging.

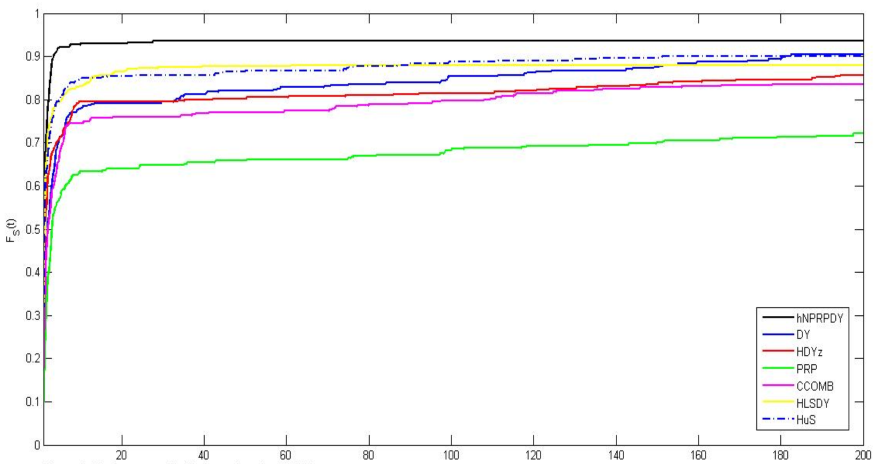

Figure 1. Performance Profile based on the CPU time.

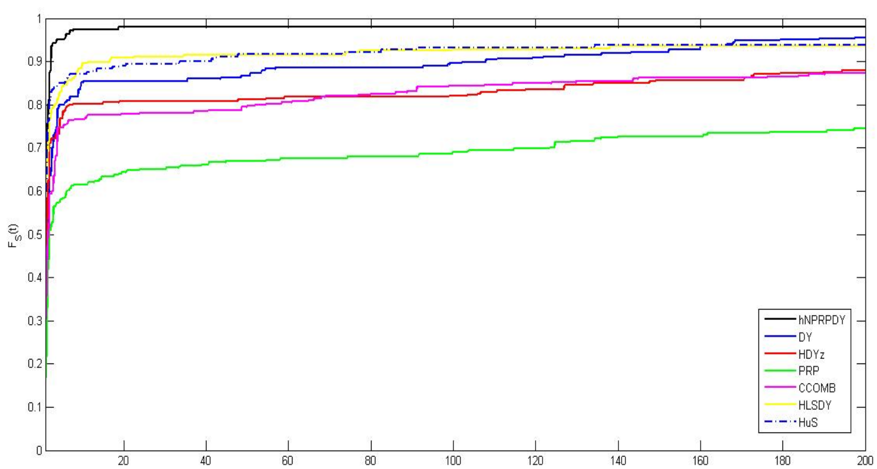

Figure 2. Performance Profle based on the number of iterations

Conclusion

In this paper, we propose hNPRPDY hybrid conjugate gradient algorithm by using a convex combination of NPRP and DY methods. The global convergence properties and the sufficient descent condition of the proposed method have been established, under the strong Wolfe line search conditions. Numerical results also show that our method is very robust and effective.

Acknowledgments

The authors wish to thank the Editors and the Reviewers for their constructive and pertinent suggestions for improving the presentation of the work.

References

[1] Andrei, N., *An unconstrained optimization test functions*, Adv. Model. Optim. **10** (2008), 147-161.

[2] Andrei, N., *Another hybrid conjugate gradient algorithm for unconstrained optimization*, Numer. Algorithms **47** (2008), 143-156.

[3] Andrei, N., *A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan*, Stud. Inform. Control. **17** (2008), 55-70.

[4] Andrei, N., *New hybrid conjugate gradient algorithms for unconstrained optimization*, Ency. Optim. (2009), 2560-2571.

[5] Andrei, N., *New hybrid conjugate gradient algorithms for unconstrained optimization*, J. Optim. Theory Appl. **141** (2009), 249-264.

[6] Bongartz, I., Conn, A. R., Gould, N. and Toint, P. L., *Constrained and unconstrained testing environment*, ACM Trans. Math. Soft. **21** (1995), 123-160.

[7] Dai, Y. H., Han, J. Y., Liu, G.H., Sun, D. F., Yin, X. and Yuan, Y., *Convergence properties of nonlinear conjugate gradient methods*, SIAM J. Optim. **10** (1999), 348-358.

[8] Dai, Y. H. and Yuan, Y. *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., **10** (1999), 177-182.

[9] Dai, Y. H. and Yuan, Y., *An efficient hybrid conjugate gradient method for unconstrained optimization*, Ann. Oper. Res. **103** (2001), 33-47.

[10] Djordjevic, S. S., *New hybrid conjugate gradient mathod as a convex combination of LS and CD methods*, Filomat **31** (2017), 1813-1825.

[11] Dolan, E. D., Morè, J. J., *Benchmarking optimization software with performance profiles*, Math. Prog., **91** (2002), 201-213.

[12] Fletcher, R., Reeves, C., *Function minimization by conjugate gradients*, Comput. J. **7** (1964), 149-154.

[13] Fletcher, R., *Practical methods of optimization*, second ed., John Wiley & Sons, New York, 1987.

[14] Hestenes, M. R. and Stiefel, E. L., *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand. **49** (1952), 409-436.

[15] Gilbert, J. C. and Nocedal, J., *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim. **2** (1992), 21-42.

[16] Huang, H., Wei, Z. and Yao, S., *The proof of the sufficient descent condition of the Wei-Yao-Liu conjugate gradient method under the strong Wolfe-Powell line search*, Appl. Math. Comput. **189** (2007), 1241-1245.

[17] Hu, Y. F. and Storey, C., *Global convergence result for conjugate gradient methods*, J. Optim. Theory Appl. **71** (1991), 399-405.

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

[19] Polak, E. and Ribière, G., *Note sur la convergence de directions conjuguée*, Rev. fr. inform. rech. opér. **16** (1969), 35-43.

[20] Polyak, B. T.,*The conjugate gradient method in extreme problems*, USSR Comput. Math. Math. Phys. **9** (1969), 94-112.

[21] Touati-Ahmed, D. and Storey, C., *Efficient hybrid conjugate gradient techniques*, J. Optim. Theory Appl., **64** (1990), 379-397.

[22] Wei, Z., Yao, S. and Liu, L.,*The convergence properties of some new conjugate gradient methods*, Appl. Math. Comput. **183** (2006), 1341-1350.

[23] Zhang, L., *An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation*, Appl. Math. Comput. **6** (2009), 2269-2274.