Li and Yang *Journal of Inequalities and Applications* (2015) 2015:124
DOI 10.1186/s13660-015-0644-1

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

A nonmonotone hybrid conjugate gradient method for unconstrained optimization

Wenyu Li and Yueting Yang*

*Correspondence: yangyueting@163.com
School of Mathematics and Statistics, Beihua University, Jilin Street No. 15, Jilin, China

Abstract
A nonmonotone hybrid conjugate gradient method is proposed, in which the technique of the nonmonotone Wolfe line search is used. Under mild assumptions, we prove the global convergence and linear convergence rate of the method. Numerical experiments are reported.

Keywords: unconstrained optimization; nonmonotone hybrid conjugate gradient algorithm; global convergence; linear convergence rate

1 Introduction

Let us take the following unconstrained optimization problem:

$$\min_{x\in R^n} f(x), \tag{1}$$

where $f : R^n \to R$ is continuously differentiable. For solving (1), the conjugate gradient method generates a sequence $\{x_k\}$: $x_{k+1} = x_k + \alpha_k d_k$, $d_0 = -g_0$, and $d_k = -g_k + \beta_k d_{k-1}$, where the stepsize $\alpha_k > 0$ is obtained by the line search, d_k is the search direction, $g_k = \nabla f(x_k)$ is the gradient of $f(x)$ at the point x_k, and β_k is known as the conjugate gradient parameter. Different parameters correspond to different conjugate gradient methods. A remarkable survey of conjugate gradient methods is given by Hager and Zhang [1].

Plenty of hybrid conjugate gradient methods were presented in [2–7] after the first hybrid conjugate algorithm was proposed by Touati-Ahmed and Storey [8]. In [5], Lu *et al.* proposed a new hybrid conjugate gradient method (LY) with the conjugate gradient parameter β_k^{LY},

$$\beta_k^{LY} = \begin{cases} \frac{g_k^T(g_k-d_{k-1})}{d_{k-1}^T(g_k-g_{k-1})}, & \text{if } |1-\frac{g_k^T d_{k-1}}{\|g_k\|^2}| \le \mu, \\ \frac{\mu\|g_k\|^2}{d_{k-1}^T g_k-\lambda d_{k-1}^T g_{k-1}}, & \text{otherwise}, \end{cases} \tag{2}$$

where $0 < \mu \le \frac{\lambda-\sigma}{1-\sigma}$, $\sigma < \lambda \le 1$. Numerical experiments show that the LY method is effective.

It is well known that the nonmonotone algorithms are promising methods for solving highly nonlinear large-scale and possibly ill-conditioned problems. The first nonmonotone line search framework was proposed by Grippo *et al.* in [9] for Newton's methods.

At each iteration, the current function value is defined as follows:

$$f_{l(k)} = \max_{0\leq j\leq m(k)} f(x_{k-j}), \tag{3}$$

where $m(0) = 0$, $0 \leq m(k) \leq \min\{m(k-1)+1, M\}$, M is some positive integer. Zhang and Hager [10] proposed another nonmonotone line search technique, they adopted C_k to replace the current function f_k, where

$$C_k = \frac{\zeta_{k-1}Q_{k-1}C_{k-1} + f_k}{Q_k}, \tag{4}$$

$Q_0 = 1$, $C_0 = f(x_0)$, $\zeta_{k-1} \in [0,1]$, and

$$Q_k = \zeta_{k-1}Q_{k-1} + 1. \tag{5}$$

To obtain the global convergence (see [4, 11–14]) and implement the algorithms, the line search in the conjugate gradient is usually chosen by a Wolfe line search; the stepsize α_k satisfies the following two inequalities:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \rho\alpha_k g_k^T d_k, \tag{6}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \tag{7}$$

where $0 < \rho < \sigma < 1$. In particular, a nonmonotone version line search can relax the choice of the stepsize. Therefore the nonmonotone Wolfe line search requires the stepsize α_k to satisfy

$$f(x_k + \alpha_k d_k) \leq f_{l(k)} + \rho\alpha_k g_k^T d_k \tag{8}$$

and (7), or

$$f(x_k + \alpha_k d_k) \leq C_k + \rho\alpha_k g_k^T d_k \tag{9}$$

and (7).

The aim of this paper is to propose a nonmonotone hybrid conjugate gradient method which combines the nonmonotone line search technique with the LY method. It is based on the idea that the larger values of the stepsize α_k may be accepted by the nonmonotone algorithmic framework and improve the behavior of the LY method.

The paper is organized as follows. A new nonmonotone hybrid conjugate gradient algorithm is presented and the global convergence of the algorithm is proved in Section 2. The line convergence rate of the algorithm is shown in Section 3. In Section 4, numerical results are reported.

2 Nonmonotone hybrid conjugate gradient algorithm and global convergence

Now we present a nonmonotone hybrid conjugate gradient algorithm.

Algorithm 1

Step 1. Given $x_0 \in R^n$, $\epsilon > 0$, $d_0 = -g_0$, $C_0 = f_0$, $Q_0, \zeta_0, k := 0$.

Step 2. If $\|g_k\| < \epsilon$, then stop. Otherwise, compute α_k by (9) and (7), set $x_{k+1} = x_k + \alpha_k d_k$.

Step 3. Compute β_{k+1} by (2), set $d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$, $k := k + 1$, and go to Step 2.

Assumption 1 We make the following assumptions:

(i) The level set $\Omega_0 = \{x \in R^n : f(x) \leq f(x_0)\}$ is bounded, where x_0 is the initial point.

(ii) The gradient function $g(x) = \nabla f(x)$ of the objective function f is Lipschitz continuous in a neighborhood $\mathcal{N}$ of level set Ω_0, *i.e.* there exists a constant $L \geq 0$ such that

$$\|g(x) - g(\bar{x})\| \leq L\|x - \bar{x}\|,$$

for any $x, \bar{x} \in \mathcal{N}$.

Lemma 2.1 *Let the sequence* $\{x_k\}$ *be generated by Algorithm* 1. *Then* $d_k^T g_k < 0$ *holds for all* $k \geq 1$.

Proof From Lemma 2 and Lemma 3 in [5], the conclusion holds. □

Lemma 2.2 *Let Assumption* 1 *hold and the sequence* $\{x_k\}$ *be obtained by Algorithm* 1, α_k *satisfies the nonmonotone Wolfe conditions* (9) *and* (7). *Then*

$$\alpha_k \geq \frac{\sigma - 1}{L} \frac{g_k^T d_k}{\|d_k\|^2}. \tag{10}$$

Proof From (7), we have

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

and by (ii) of Assumption 1 it implies that

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

By combining these two inequalities, we obtain

$$\alpha_k \geq \frac{\sigma - 1}{L} \frac{g_k^T d_k}{\|d_k\|^2}.$$ □

Lemma 2.3 *Let the sequence* $\{x_k\}$ *be generated by Algorithm* 1 *and* $d_k^T g_k < 0$ *hold for all* $k \geq 1$. *Then*

$$f_k \leq C_k. \tag{11}$$

Proof See Lemma 1.1 in [10]. □

Lemma 2.4 *Let Assumption* 1 *hold, and the sequence* $\{x_k\}$ *be obtained by Algorithm* 1, *where* d_k *satisfies* $d_k^T g_k < 0$, α_k *is obtained by the nonmonotone Wolfe conditions* (9) *and* (7).

Then

$$\sum_{k\geq 0}\frac{1}{Q_{k+1}}\frac{(d_k^T g_k)^2}{\|d_k\|^2} < +\infty. \tag{12}$$

Proof By (9) and (10), we have

$$f_{k+1} \leq C_k - c_0\frac{(d_k^T g_k)^2}{\|d_k\|^2}, \tag{13}$$

where $c_0 = \rho(1-\sigma)/L$.

From (4), (5), and (13), we have

$$C_{k+1} = \frac{\zeta_k Q_k C_k + f(x_{k+1})}{Q_{k+1}} \leq \frac{\zeta_k Q_k C_k + C_k - c_0\frac{(d_k^T g_k)^2}{\|d_k\|^2}}{Q_{k+1}} \leq C_k - \frac{c_0}{Q_{k+1}}\frac{(d_k^T g_k)^2}{\|d_k\|^2}. \tag{14}$$

Since $f(x)$ is bounded from below in the level set Ω_0 and by (11) for all k, we know that C_k is bounded from below. It follows from (14) that (12) holds. □

Theorem 2.1 *Suppose that Assumptions* 1 *hold and the sequence* $\{x_k\}$ *is generated by the Algorithm* 1. *If* $\zeta_{\max} < 1$, *then either* $g_k = 0$ *for some* k *or*

$$\lim_{k\to\infty}\inf\|g_k\| = 0. \tag{15}$$

Proof We prove by contradiction and assume that there exists a constant $\epsilon > 0$ such that

$$\|g_k\|^2 \geq \epsilon, \quad k = 0,1,2,3,\ldots. \tag{16}$$

By Lemma 4 in [5], we have $|\beta_k|^{LY} \leq \frac{\mu\|g_k\|^2}{d_{k-1}^T g_k - \lambda d_{k-1}^T g_{k-1}}$. Then we have $\|d_k\|^2 = (\beta^{LY})^2 \times \|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2 \leq (\frac{\mu\|g_k\|^2}{d_{k-1}^T g_k - \lambda d_{k-1}^T g_{k-1}})^2\|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2$. The rest of the proof is similar to Theorem 2 and Theorem 1 in [5], and we also conclude

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \frac{\epsilon}{k}.$$

Furthermore, by $\zeta_{\max} < 1$ and (5), we have

$$Q_k = 1 + \sum_{j=0}^{k-1}\prod_{i=0}^{j}\zeta_{k-1-i} \leq \frac{1}{1-\zeta_{\max}}, \tag{17}$$

then

$$\frac{1}{Q_{k+1}}\frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq (1-\zeta_{\max})\frac{\epsilon}{k},$$

which indicates

$$\sum_{i=1}^{\infty}\frac{1}{Q_{k+1}}\frac{(g_k^T d_k)^2}{\|d_k\|^2} = +\infty.$$

This contradicts (12). Therefore (15) holds. □

3 Linear convergence rate of algorithm

We analyze the linearly convergence rate of the nonmonotone hybrid conjugate gradient method under the uniform convex assumption of $f(x)$. The nonmonotone strong Wolfe line search is adopted in this section, given by (9) and

$$\left|g(x_k+\alpha_k d_k)^T d_k\right| \le -\sigma g_k^T d_k. \tag{18}$$

We suppose that the object function $f(x)$ is twice continuously differentiable and uniformly convex on the level set Ω_0. Then the point x^* denotes a unique solution of the problem (1); there exists a positive constant τ such that

$$f(x)-f\left(x^*\right) \le \left\|\nabla f(x)\right\|\left\|x-x^*\right\| \le \tau\left\|\nabla f(x)\right\|^2, \quad \text{for all } x \in R^n. \tag{19}$$

The above conclusion (19) can be found in [10].

To analyze the convergence of the nonmonotone line search hybrid conjugate gradient method, the main difficulty is that the search directions do not usually satisfy the direction condition:

$$g_k^T d_k \le -c\|g_k\|^2, \tag{20}$$

for some constant $c>0$ and all $k \ge 1$. The following lemma has proven that the direction generated by Algorithm 1 with the strong Wolfe line search (9) and (18) in this paper satisfies the direction condition (20) by the observation for $g_k^T d_{k-1}$.

Lemma 3.1 *Suppose that the sequence* $\{x_k\}$ *is generated by Algorithm* 1 *with the strong Wolfe line search* (9) *and* (18), $0<\sigma<\frac{\lambda}{1+\mu}$. *Then there exists some constant* $c>0$ *such that the direction condition* (20) *holds.*

Proof According to the choice of the conjugate gradient parameter β_k^{LY}, the result is discussed by two cases. In the first case,

$$\left|1-\frac{g_k^T d_{k-1}}{\|g_k\|^2}\right| \le \mu, \quad \textit{i.e. } 1-\mu \le \frac{g_k^T d_{k-1}}{\|g_k\|^2} \le 1+\mu, \tag{21}$$

then $\beta_k = \frac{g_k^T(g_k-d_{k-1})}{d_{k-1}^T(g_k-g_{k-1})}$. If $\beta_k \ge 0$, then, by (21) and $d_{k-1}^T g_{k-1}<0$, $d_{k-1}^T g_k \ge 0$ and $g_k^T(g_k-d_{k-1}) \ge 0$. Furthermore, we have, by (18),

$$\begin{aligned}
g_k^T d_k &= -\|g_k\|^2 + \frac{g_k^T(g_k-d_{k-1})}{d_{k-1}^T(g_k-g_{k-1})} g_k^T d_{k-1} \\
&\le -\|g_k\|^2 - \sigma\frac{g_k^T(g_k-d_{k-1})}{d_{k-1}^T(g_k-g_{k-1})} g_{k-1}^T d_{k-1} \\
&\le -\|g_k\|^2 - \sigma\frac{g_k^T(g_k-d_{k-1})}{-d_{k-1}^T g_{k-1}} g_{k-1}^T d_{k-1} \\
&= -\|g_k\|^2 + \sigma\left(\|g_k\|^2 - g_k^T d_{k-1}\right) \\
&= -(1-\sigma)\|g_k\|^2 - \sigma g_k^T d_{k-1}
\end{aligned}$$

$$\leq -(1-\sigma)\|g_k\|^2 - \sigma(1-\mu)\|g_k\|^2$$
$$= -(1-\sigma\mu)\|g_k\|^2. \tag{22}$$

If $\beta_k < 0$, we have, by (18) and (21),

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \frac{g_k^T(g_k - d_{k-1})}{d_{k-1}^T(g_k - g_{k-1})} g_k^T d_{k-1} \\ &\leq -\|g_k\|^2 + \sigma \frac{g_k^T(g_k - d_{k-1})}{d_{k-1}^T(g_k - g_{k-1})} g_{k-1}^T d_{k-1} \\ &\leq -\|g_k\|^2 + \sigma \frac{g_k^T(g_k - d_{k-1})}{-d_{k-1}^T g_{k-1}} g_{k-1}^T d_{k-1} \\ &\leq -\|g_k\|^2 - \sigma\left(\|g_k\|^2 - g_k^T d_{k-1}\right) \\ &= -(1+\sigma)\|g_k\|^2 + \sigma g_k^T d_{k-1} \\ &= -(1+\sigma)\|g_k\|^2 + \sigma(1+\mu)\|g_k\|^2 \\ &= -(1-\sigma\mu)\|g_k\|^2. \end{aligned} \tag{23}$$

In the second case,

$$\left|1 - \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right| > \mu, \tag{24}$$

then $\beta_k = \frac{\mu\|g_k\|^2}{d_{k-1}^T g_k - \lambda d_{k-1}^T g_{k-1}} > 0$. By (18), we have

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \frac{\mu\|g_k\|^2}{d_{k-1}^T g_k - \lambda d_{k-1}^T g_{k-1}} g_k^T d_{k-1} \\ &\leq -\|g_k\|^2 - \frac{\sigma\mu\|g_k\|^2}{(\sigma-\lambda)d_{k-1}^T g_{k-1}} g_{k-1}^T d_{k-1} \\ &= -\left(1 - \frac{\mu\sigma}{\lambda-\sigma}\right)\|g_k\|^2. \end{aligned} \tag{25}$$

From (23) and (25), we obtain (20), where $c = \min\{1-\sigma\mu, 1-\frac{\mu\sigma}{\lambda-\sigma}\} > 0$. The proof is completed. □

Lemma 3.2 *Suppose the assumptions of Lemma* 2.2 *hold and, for all k*,

$$\|d_k\| \leq c_1\|g_k\|, \tag{26}$$

then there exists a constant $c_2 > 0$ *such that*

$$\alpha_k \geq c_2, \quad \textit{for all } k. \tag{27}$$

Proof By Lemma 2.2 and Lemma 3.2, we have

$$\alpha_k \geq \frac{\sigma-1}{L}\frac{g_k^T d_k}{\|d_k\|^2} \geq -\frac{\sigma-1}{L}\frac{c\|g_k\|^2}{c_1\|g_k\|^2} \geq c_2,$$

where $c_2 = \frac{c(1-\sigma)}{c_1 L}$. □

Table 1 Test problems

No.	Problem name	No.	Problem name
1	Helical valley function	21	Separable cubic function
2	BIGGS6	22	Nearly separable function
3	Gaussian function	23	Allgower function
4	POWELLBS	24	Schittkowski function 302
5	BOX3	25	Discrete integral equation function
6	BEALE	26	BDQRTIC
7	WOODS	27	ARGLINB
8	FREUROTH	28	ARWHEAD
9	Osborne 1 function	29	NONDIA
10	Osborne 2 function	30	NONDQUAR
11	Powell singular function	31	DQDRTIC
12	Meyer function	32	EG2
13	Bard function	33	CURLY20
14	PENALTY2	34	LIARWHD
15	VARDIM	35	ENGVAL1
16	PENALTY1	36	CRAGGLVY
17	EXTROSNB	37	the Edensch function
18	Extended Powell singular function	38	the Explin 1 function
19	CHEBYQAD	39	the Argling function
20	BROYDN3D	40	NONSCOMP

Theorem 3.1 *Let x^* be the unique solution of problem* (1) *and the sequence $\{x_k\}$ be generated by Algorithm* 1 *with the nonmonotone Wolfe conditions* (9) *and* (18), $0 < \sigma < \frac{\lambda}{1+\mu}$. *If $\alpha_k \leq \nu$ and $\zeta_{\max} < 1$, then there exists a constant $\vartheta \in (0,1)$ such that*

$$f_k - f(x^*) \leq \vartheta^k \bigl(f_0 - f(x^*)\bigr). \tag{28}$$

Proof The proof is similar to that Theorem 3.1 given in [10]. By (9), (20), and (27), we have

$$f_{k+1} \leq C_k + \rho\alpha_k g_k^T d_k \leq C_k - cc_2\rho\|g_k\|^2. \tag{29}$$

By (ii) in Assumption 1, $x_{k+1} = x_k + \alpha_k d_k$ and (27), we have

$$\|g_{k+1}\| \leq \|g_{k+1} - g_k\| + \|g_k\| \leq \alpha_k L\|d_k\| + \|g_k\| \leq (1 + c_1\nu L)\|g_k\|. \tag{30}$$

In the first case, $\|g_k\|^2 \geq \beta(C_k - f(x^*))$, where

$$\beta = 1/\bigl(cc_2\rho + \tau(1 + c_1\nu L)^2\bigr). \tag{31}$$

By (4) and (29), we have

$$\begin{aligned} C_{k+1} - f(x^*) &= \frac{\zeta_k Q_k(C_k - f(x^*)) + (f_{k+1} - f(x^*))}{1 + \zeta_k Q_k} \\ &\leq \frac{\zeta_k Q_k(C_k - f(x^*)) + (C_k - f(x^*)) - cc_2\rho\|g_k\|^2}{1 + \zeta_k Q_k} \\ &= C_k - f(x^*) - \frac{cc_2\rho\|g_k\|^2}{Q_{k+1}}. \end{aligned} \tag{32}$$

Since $Q_{k+1} \leq \frac{1}{1-\zeta_{\max}}$ by (17), we have

$$C_{k+1} - f(x^*) \leq C_k - f(x^*) - cc_2\rho(1 - \zeta_{\max})\|g_k\|^2.$$

Table 2 Numerical comparisons of LY, NMLY1, and NMLY2

No.	Dim	LY				NMLY2				NMLY1			
		it	nf	ng	t	it	nf	ng	t	it	nf	ng	t
1	3	36	69	48	0.0156	59	137	92	0.0156	40	84	60	0.0156
2	6	147	243	200	0.0468	413	853	697	0.1404	146	270	213	0.0468
3	3	2	5	4	0	4	10	8	0	3	9	8	0
4	2	9	26	13	0	11	48	28	0	13	38	17	0.0312
5	3	15	28	19	0	29	75	68	0.0312	14	35	28	0
6	2	37	61	46	0.0156	28	64	45	0	20	60	30	0.0156
7	4	74	130	100	0.0156	313	623	458	0.0936	62	119	90	0.0156
8	2	44	82	61	0.0156	37	98	70	0	5	25	9	0.0156
9	5	128	233	184	0.0936	181	383	278	0.1404	104	199	153	0.078
10	11	3	42	3	0.0156	4	29	4	0	3	13	3	0.0312
11	4	6	49	7	0	169	348	254	0.0468	59	121	77	0.0312
12	3	9	45	14	0	170	456	328	0.1248	13	80	34	0.0156
13	3	25	49	33	0.0156	64	135	101	0.0468	39	75	52	0.0312
14	500	14	70	28	0.0156	10	82	40	0.0156	3	46	10	0.0156
15	500	13	68	14	0	13	118	48	0.0312	17	187	122	0.0156
	1,000	16	70	16	0	13	87	35	0.0156	19	301	211	0.0312
	2,000	14	58	14	0.0156	15	116	50	0.0156	26	362	237	0.0624
16	500	11	35	22	0.0156	40	147	101	0.0156	34	103	87	0.0156
	1,000	16	40	23	0	49	141	99	0.0312	21	82	62	0.0156
	2,000	20	51	28	0.0312	34	122	73	0.0312	57	156	108	0.0624
	1,000	22	75	30	0.078	28	96	67	0.1248	31	134	70	0.156
17	500	42	87	58	0.0156	133	294	213	0.078	3	18	3	0
	1,000	64	120	87	0.0312	187	406	299	0.0936	3	19	3	0
	2,000	53	105	72	0.0312	134	304	221	0.0936	3	19	3	0.0156
	1,000	56	113	79	0.2028	184	393	292	0.624	3	19	3	0.0312
18	500	220	364	295	0.1092	184	353	253	0.1092	110	197	160	0.0624
	1,000	180	288	238	0.1404	163	322	235	0.1404	104	191	150	0.078
	2,000	181	308	248	0.2184	181	361	262	0.2496	112	205	164	0.156
	1,000	217	361	293	1.2792	320	625	458	2.1372	133	244	197	0.8892
19	500	68	105	86	5.694	103	210	146	10.9201	70	112	95	6.1308
	1,000	99	143	125	32.7602	88	185	133	38.891	113	201	144	40.0923
	2,000	72	110	95	90.5742	190	387	287	306.2456	78	141	104	110.3239
	1,000	121	197	142	983.8203	35	156	80	708.6189	24	67	45	322.8129
20	500	51	74	60	0.0468	55	114	73	0.0156	53	83	68	0.0312
	1,000	54	78	62	0.0312	55	117	73	0.0312	93	143	115	0.0468
	2,000	44	74	58	0.0312	58	130	87	0.0624	33	61	39	0.0156
	1,000	43	69	57	0.1872	52	115	71	0.1872	42	75	59	0.1248
21	500	4	9	8	0.0468	10	20	19	0.1092	8	18	17	0.1092
	1,000	5	11	10	0.234	10	20	19	0.468	9	20	19	0.468
	2,000	5	11	10	0.9204	10	20	19	1.8408	9	20	19	1.8252
	1,000	5	11	10	6.63	11	23	22	15.3973	10	21	20	13.8997
22	500	50	91	68	1.1076	86	207	133	2.2308	36	92	65	1.0764
	1,000	33	96	52	3.3852	52	146	100	6.552	43	104	71	4.602
	2,000	27	113	48	12.1681	43	147	91	23.6186	52	167	83	22.9477
	1,000	102	222	142	246.5128	76	196	140	244.2664	42	123	71	124.442
23	500	4	32	8	0.1092	13	69	30	0.4368	3	36	6	0.078
	1,000	6	69	9	0.4836	17	230	61	3.4944	3	57	5	0.234
	2,000	5	61	8	1.794	16	165	41	10.3585	3	37	3	0.5304
	1,000	5	62	11	16.1617	12	125	20	30.6074	3	56	5	6.1932
24	500	59	137	88	0.0156	62	232	133	0.0312	34	192	121	0.0156
	1,000	26	91	34	0.0156	34	174	98	0.0312	56	292	204	0.0624
	2,000	17	68	26	0.0156	16	112	35	0.0312	22	272	172	0.0624
	1,000	37	89	42	0.1092	34	164	59	0.156	25	266	146	0.2496

Table 2 (***Continued***)

No.	Dim	LY				NMLY2				NMLY1			
		it	nf	ng	t	it	nf	ng	t	it	nf	ng	t
25	500	7	67	43	12.8545	13	85	53	15.9589	13	80	26	10.4521
	1,000	7	59	28	26.7698	8	78	51	41.9019	47	229	130	115.8931
	2,000	22	141	90	275.5914	7	55	28	96.081	7	68	28	107.2507
	1,000	7	63	36	716.4346	16	175	117	2148.9606	11	115	63	1265.6829
26	500	54	105	77	1.6224	55	127	88	1.7628	46	97	68	1.3728
	1,000	53	101	75	5.4132	92	195	140	10.4053	46	94	69	5.0856
	2,000	42	89	62	15.1945	71	162	113	27.5654	40	92	65	16.0213
	1,000	29	68	43	44.4603	58	150	105	110.0275	44	97	68	69.7168
27	500	3	52	4	0.2184	11	178	91	1.9968	9	152	12	0.6708
	1,000	13	193	99	8.7049	5	68	13	1.5756	7	106	27	2.9484
	2,000	5	101	46	16.3645	7	141	40	16.5205	7	141	48	18.7045
	1,000	9	112	32	99.279	8	140	33	112.5859	13	224	48	172.6151
28	500	8	23	10	0	22	60	38	0.0468	19	45	27	0.0156
	1,000	9	25	13	0.0156	7	45	31	0.0156	11	59	43	0.0312
	2,000	16	60	26	0.0312	13	78	42	0.0624	16	72	36	0.0468
	1,000	5	22	7	0.0312	10	53	30	0.2496	11	60	40	0.2808
29	500	28	65	41	0.0468	13	69	52	0.0312	13	50	27	0.0312
	1,000	4	17	5	0	22	86	53	0.0468	11	56	34	0.0156
	2,000	7	23	8	0	10	49	26	0.0468	13	66	35	0.0468
	1,000	7	26	9	0.0624	4	40	28	0.156	12	124	101	0.7332
30	500	287	452	387	14.6017	365	691	649	21.0913	249	380	334	10.1089
	1,000	292	456	407	44.5071	343	646	593	69.3736	257	404	359	40.8411
	2,000	343	524	477	154.3786	376	721	655	213.9242	271	457	376	122.1176
	1,000	310	474	424	640.2749	376	729	663	1006.2533	236	412	337	512.9157
31	500	70	110	87	0.3276	73	158	113	0.39	58	115	90	0.2964
	1,000	49	88	66	0.468	50	131	93	0.624	30	68	45	0.312
	2,000	38	69	49	0.7644	64	142	103	1.3416	34	75	49	0.6396
	1,000	27	52	32	2.0904	53	128	90	6.1152	34	85	60	4.0248
32	500	6	50	9	0.0156	53	144	89	0.0624	4	60	6	0
	1,000	30	88	44	0.0468	18	211	35	0.0468	6	93	7	0.0156
	2,000	3	23	6	0.0156	10	107	16	0.0468	4	45	5	0.0156
	1,000	5	39	10	0.0936	8	80	17	0.1872	4	55	8	0.0936
33	500	320	442	406	7.6596	427	822	614	11.8405	4	24	7	0.1248
	1,000	268	381	342	20.6233	338	652	483	29.1566	4	25	7	0.39
	2,000	195	269	248	47.1435	235	461	342	66.7996	4	25	7	1.2792
	1,000	145	198	182	137.4681	223	434	326	247.1524	4	25	7	4.5396
34	500	105	187	146	0.1404	101	234	162	0.156	98	191	143	0.1404
	1,000	33	76	49	0.1248	73	197	136	0.2496	97	203	156	0.2028
	2,000	123	226	172	0.39	173	416	306	0.6708	89	196	141	0.312
	1,000	175	311	243	2.6676	236	537	395	4.6176	151	319	239	2.6676
35	500	12	25	14	0.0468	13	35	23	0.078	5	18	5	0.0156
	1,000	12	25	14	0.078	11	30	19	0.1248	5	18	5	0.0312
	2,000	11	23	13	0.1716	11	34	23	0.2808	5	18	5	0.0624
	1,000	10	23	14	0.9984	10	31	22	1.56	5	18	5	0.2808
36	500	37	228	122	0.6864	10	78	17	0.1404	9	34	16	0.078
	1,000	17	75	25	0.2652	11	78	20	0.234	9	51	13	0.156
	2,000	37	137	57	1.1076	11	101	21	0.5148	15	90	34	0.6864
	1,000	16	82	23	2.496	10	82	18	2.1684	15	76	21	2.106
37	500	11	25	16	0.0156	10	33	19	0	3	17	6	0
	1,000	12	26	15	0.0156	11	29	19	0.0156	3	17	6	0
	2,000	10	22	12	0.0312	10	27	18	0.0312	3	18	6	0.0156
	5,000	8	18	10	0.0936	10	27	18	0.156	3	18	6	0.0624
38	500	6	75	50	0.0156	10	123	72	0.0624	8	105	51	0.0312
	1,000	4	63	22	0.0156	7	107	43	0.0468	4	56	24	0.0156
	2,000	7	121	65	0.0624	13	211	130	0.1248	8	132	68	0.078
	5,000	5	95	49	0.2184	5	84	34	0.156	6	115	51	0.2496

Table 2 (*Continued*)

No.	Dim	LY				NMLY2				NMLY1			
		it	nf	ng	t	it	nf	ng	t	it	nf	ng	t
39	500	8	127	51	2.7924	10	168	76	3.7596	9	127	62	3.042
	1,000	7	107	40	8.5021	11	188	85	15.7405	9	144	61	11.9341
	2,000	10	142	67	46.7691	16	248	104	81.2453	113	573	189	182.7708
	5,000	9	165	72	698.6037	12	162	78	675.7495	13	181	83	756.8077
40	500	170	257	218	0.1092	305	578	430	0.1716	105	171	140	0.1404
	1,000	153	231	196	0.0624	265	505	367	0.1404	114	208	155	0.0624
	2,000	147	222	190	0.0936	272	522	390	0.1872	117	188	155	0.078
	5,000	144	233	184	0.3588	305	579	425	0.8424	125	205	172	0.3276

By $\|g_k\|^2 \geq \beta(C_k - f(x^*))$, we have

$$C_{k+1} - f(x^*) \leq \vartheta\bigl(C_k - f(x^*)\bigr), \tag{33}$$

where $\vartheta = 1 - cc_2\rho\beta(1-\zeta_{\max}) \in (0,1)$.

In the second case, $\|g_k\|^2 < \beta(C_k - f(x^*))$. By (19) and (30), we have

$$f_{k+1} - f(x^*) \leq \tau(1+c_1\nu L)^2\|g_k\|^2 \leq \tau\beta(1+c_1\nu L)^2\bigl(C_k - f(x^*)\bigr).$$

By combining the equality, the first equation of (32), and $Q_{k+1} \leq \frac{1}{1-\zeta_{\max}}$, $\zeta_{\max} < 1$ and (31), we obtain

$$\begin{aligned} C_{k+1} - f(x^*) &\leq \frac{\zeta_k Q_k(C_k - f(x^*)) + \tau\beta(1+c_1\nu L)^2(C_k - f(x^*))}{1+\zeta_k Q_k} \\ &= \left(1 - \frac{1-\tau\beta(1+c_1\nu L)^2}{Q_{k+1}}\right)\bigl(C_k - f(x^*)\bigr) \\ &= \bigl(1 - (1-\tau\beta(1+c_1\nu L)^2)(1-\zeta_{\max})\bigr)\bigl(C_k - f(x^*)\bigr) \\ &= \bigl(1 - cc_2\rho\beta(1-\zeta_{\max})\bigr)\bigl(C_k - f(x^*)\bigr) \\ &\leq \vartheta\bigl(C_k - f(x^*)\bigr). \end{aligned} \tag{34}$$

By (11), (33), and (34), we have

$$f_k - f(x^*) \leq C_k - f(x^*) \leq \vartheta\bigl(C_{k-1} - f(x^*)\bigr) \leq \cdots \leq \vartheta^k\bigl(C_0 - f(x^*)\bigr).$$

The proof is completed. □

4 Numerical experiments

In this section, we report numerical results to illustrate the performance of hybrid conjugate gradient (LY) in [5], Algorithm 1 (NHLYCG1) and Algorithm 2 (NGLYCG2), in which (8) replaces only (9) in Step 2 of Algorithm 1. All codes are written with Matlab R2012a and are implemented on a PC with CPU 2.40 GHz and 2.00GB RAM. We select 12 small-scale and 28 large-scale unconstrained optimization test functions from [15] and the CUTEr collection [16, 17] (see Table 1). All algorithms implement the stronger version of the Wolfe condition with $\rho = 0.45$ and $\sigma = 0.39$, and $\mu = 0.5$, $\lambda = 0.6$, $C_0 = f_0$, $Q_0 = 1$,

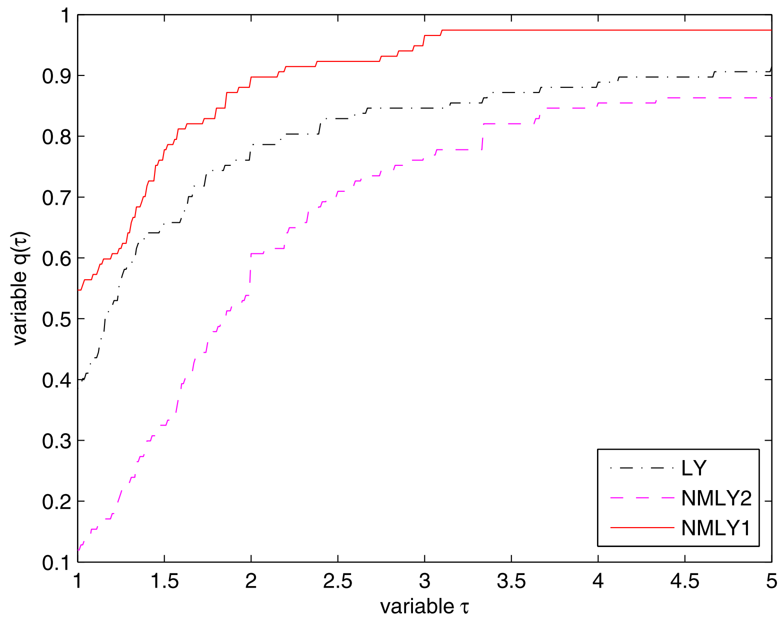

Figure 1 **Performance profile comparing the number of iterations.**

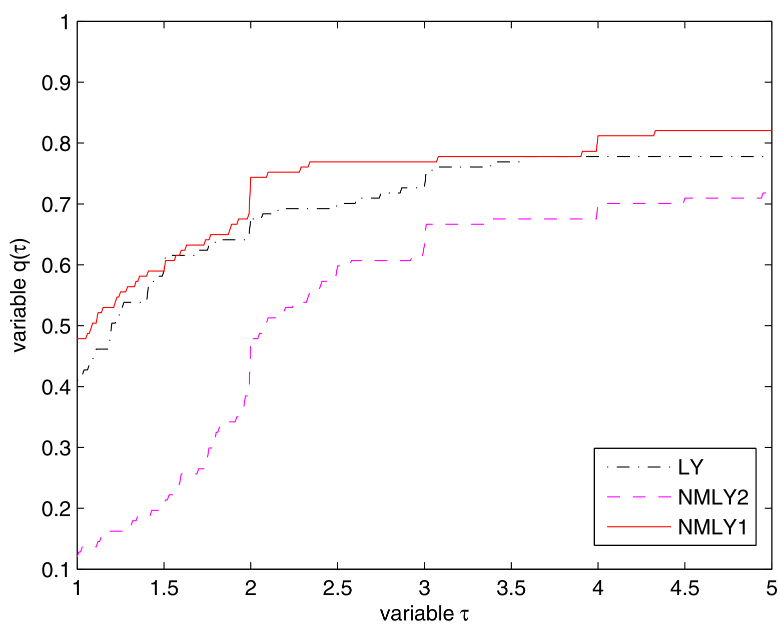

Figure 2 **Performance profile comparing the CPU time.**

$\zeta_0 = 0.08$, $\zeta_1 = 0.04$, $\zeta_{k+1} = \frac{\zeta_k+\zeta_{k-1}}{2}$, and the terminated condition

$$\|g_k\|_2 \le 10^{-6} \quad \text{or} \quad |f_{k+1} - f_k| \le 10^{-6} \max\{1.0, |f_k|\}.$$

Table 2 lists all the numerical results, which include the order numbers and dimensions of the tested problems, the number of iterations (it), the function evaluations (nf), the

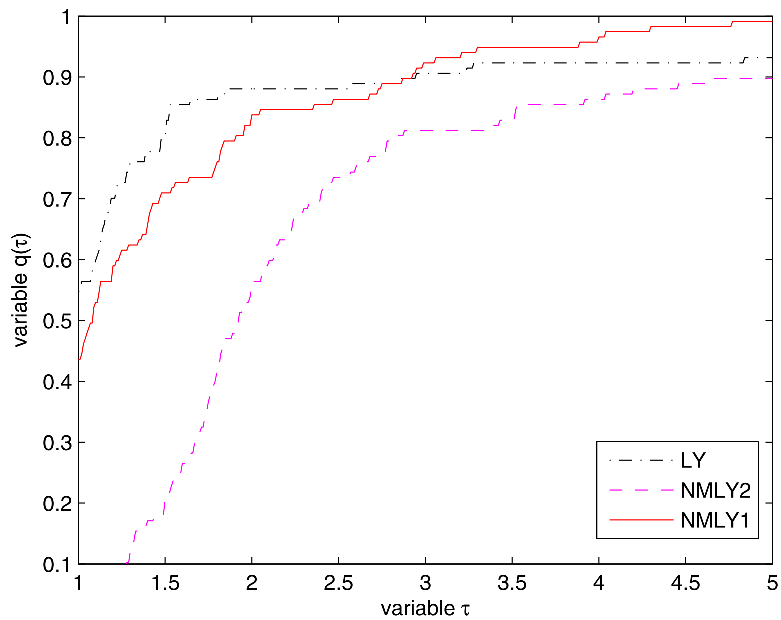

Figure 3 **Performance profile comparing the number of function evaluations.**

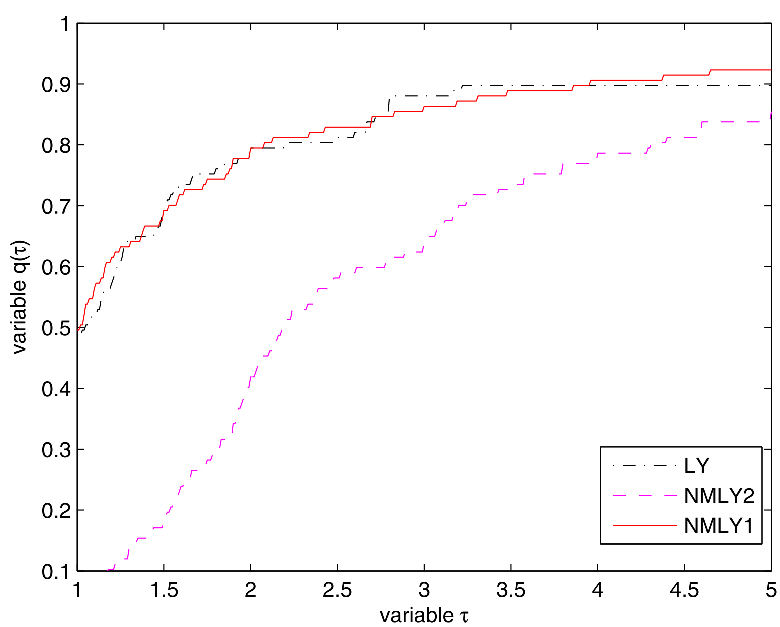

Figure 4 **Performance profile comparing the number of gradient evaluations.**

gradient evaluations (ng), and the CPU time (t) in seconds, respectively. We presented the Dolan and Moré [18] performance profiles for the LY, NHLYCG1, and NGLYCG2. Note that the performance ratio $q(\tau)$ is the probability for a solver s for the tested problems with the factor τ of the smallest cost. As we can see from Figure 1 and Figure 2, NHLYCG1 is superior to LY and NGLYCG2 for the number of iterations and CPU time. Figure 3 shows that NGLYCG1 is slightly better than LY and NGLYCG2 for the number of function value evaluations. Figure 4 shows the performance of NGLYCG1 is very much like that of LY for the number of gradient evaluations. However, the performance of NGLYCG2 with the nonmonotone framework (3) is less than satisfactory.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Acknowledgements
This work is supported in part by the NNSF (11171003) of China.

Received: 16 January 2015 Accepted: 27 March 2015 Published online: 08 April 2015

References

1. Hager, WW, Zhang, H: A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. **16**, 170-192 (2005)
2. Andrei, N: A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan. Stud. Inform. Control **17**(4), 55-70 (2008)
3. Babaie-Kafaki, S, Mahdavi-Amiri, N: Two modified hybrid conjugate gradient methods based on a hybrid secant equation. Math. Model. Anal. **18**(1), 32-52 (2013)
4. Dai, YH, Yuan, Y: An efficient hybrid conjugate gradient method for unconstrained optimization. Ann. Oper. Res. **103**, 33-47 (2001)
5. Lu, YL, Li, WY, Zhang, CM, Yang, YT: A class new conjugate hybrid gradient method for unconstrained optimization. J. Inf. Comput. Sci. **12**(5), 1941-1949 (2015)
6. Yang, YT, Cao, MY: The global convergence of a new mixed conjugate gradient method for unconstrained optimization. J. Appl. Math. **2012**, 93298 (2012)
7. Zheng, XF, Tian, ZY, Song, LW: The global convergence of a mixed conjugate gradient method with the Wolfe line search. Oper. Res. Trans. **13**(2), 18-24 (2009)
8. Touati-Ahmed, D, Storey, C: Efficient hybrid conjugate gradient techniques. J. Optim. Theory Appl. **64**(2), 379-397 (1990)
9. Grippo, L, Lampariello, F, Lucidi, S: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707-716 (1986)
10. Zhang, H, Hager, WW: A nonmonotone line search technique and its application to unconstrained optimization. SIAM J. Optim. **14**, 1043-1056 (2004)
11. Zoutendijk, G: Nonlinear programming, computational methods. In: Abadie, J (ed.) Integer and Nonlinear Programming. North-Holland, Amsterdam (1970)
12. Al-Baali, M: Descent property and global convergence of the Fletcher-Reeves method with inexact line search. IMA J. Numer. Anal. **5**, 121-124 (1985)
13. Gilbert, JC, Nocedal, J: Global convergence properties of conjugate gradient methods for optimization. SIAM J. Optim. **2**(1), 21-42 (1992)
14. Yu, GH, Zhao, YL, Wei, ZX: A descent nonlinear conjugate gradient method for large-scale unconstrained optimization. Appl. Math. Comput. **187**(2), 636-643 (2007)
15. Moré, JJ, Garbow, BS, Hillstrom, KE: Testing unconstrained optimization software. ACM Trans. Math. Softw. **7**, 17-41 (1981)
16. Andrei, N: An unconstrained optimization test functions collection. Adv. Model. Optim. **10**, 147-161 (2008)
17. Gould, NIM, Orban, D, Toint, PL: CUTEr and SifDec: a constrained and unconstrained testing environment, revisited. ACM Trans. Math. Softw. **29**, 373-394 (2003)
18. Dolan, ED, Moré, JJ: Benchmarking optimization software with performance profiles. Math. Program. **9**, 201-213 (2002)