Malaysian Journal of Science 40(3): 64-75 (Oct 2021)

AN EFFICIENT HYBRID DERIVATIVE-FREE PROJECTION ALGORITHM FOR CONSTRAINT NONLINEAR EQUATION

Kanikar Muangchoo[1a*]

[a]Faculty of Science and Technology, Rajamangala University of Technology Phra Nakhon (RMUTP) 1381, Pracharat 1 Road, Wongsawang, Bang Sue, Bangkok 10800, THAILAND. E-mail: kanikar.m@rmutp.ac.th[1]
*Corresponding Author: kanikar.m@rmutp.ac.th
Received: 12th Aug 2020 Accepted: 8th Apr 2021 Published: 31st Oct 2021
DOI: https://doi.org/10.22452/mjs.vol40no3.6

ABSTRACT In this paper, by combining the Solodov and Svaiter projection technique with the conjugate gradient method for unconstrained optimization proposed by Mohamed et al. (2020), we develop a derivative-free conjugate gradient method to solve nonlinear equations with convex constraints. The proposed method involves a spectral parameter that satisfies the sufficient descent condition. The global convergence is proved under the assumption that the underlying mapping is Lipschitz continuous and satisfies a weaker monotonicity condition. Numerical experiment shows that the proposed method is efficient.

Keywords: nonlinear equations, conjugate gradient method, projection method, sufficient descent condition, global convergence

---

1. INTRODUCTION

Iterative methods for solving nonlinear equations with convex constraints are a rapidly developing field that has many spectacular results. Our focus in this article is to solve the following nonlinear equation with convex constraints:

$$T(k) = 0,\ k \in \Lambda, \tag{1}$$

where Λ is a nonempty closed and convex subset of the Euclidean space, and the mapping $T: \mathrm{R}^n \to \mathrm{R}^n$ (where R^n is an n-dimensional Euclidean space) is continuous. The constraints on nonlinear equation (1) appear in various applications such as the economic equilibrium problem (Dirkse & Ferris, 1995), financial forecasting problems (Dai et al., 2020), nonlinear compressed sensing (Blumensath, 2013) and non-negative matrix factorisation (Berry et al., 2007; Lee & Seung, 2001). To this effect, several algorithms for solving (1) have been proposed in the literature in recent years. Newton's method, quasi-Newton methods, the Levenberg–Marquardt method and many of their variations (Dennis & More, 1974; Dennis Jr & Moré, 1977; Qi & Sun, 1993; Yamashita & Fukushima, 2001) are attractive for solving (1) because of their rapid convergence from a sufficiently good initial guess. However, these methods are not suitable for solving large-scale nonlinear equations because they need to solve a linear system of equations at each iteration using or approximating the Jacobian matrix. Consequently, derivative-free methods for solving nonlinear equations have attracted several authors. See, for example, (Huang et al., 2016).

The projection method developed in (Solodov & Svaiter, 1999) has garnered significant consideration. Motivated by the work of Solodov and Saiter (Solodov & Svaiter, 1999), Zhang and Zhou (Zhang & Zhou, 2006) proposed a spectral gradient projection method for solving a nonlinear equation involving monotone mapping. Most recently, (Ibrahim, et al., 2020) proposed a derivative-free projection method for solving nonlinear equation (1). The method combines the projection technique and the Liu-Storey-Fletcher Reeves (LS-FR) conjugate gradient method proposed by Djordjević (Djordjević, 2019). The proposed method does not store a matrix at each iteration. Several other derivative-free methods have been developed with the help of the projection scheme (Abubakar, et al., 2020; Abubakar, et al., 2020; Ibrahim et al., 2019; Ibrahim, 2020; Ibrahim, et al., 2020; Ibrahim, et al., 2020; Mohammad & Bala Abubakar, 2020).

In this paper, by applying the projection technique in (Solodov & Svaiter, 1999) to the new hybrid coefficient proposed by (Mohamed et al., 2020), we propose a derivative-free conjugate gradient method for solving large-scale nonlinear equations with convex constraints. The proposed method generates a sufficient descent direction per iteration and does not require additional computation costs. The global convergence of the method is established under the assumption that the underlying operator is Lipschitz continuous and satisfies a weaker monotonicity assumption.

The remainder of this manuscript is structured as follows: The proposed derivative-free method and its algorithm are presented in the next section. Section 3 is devoted to the proof of the theoretical/convergence analysis of the proposed method. In Section 4, we present the results of some numerical experiments and analyse the experimental results. Finally, the paper is concluded in Section 5.

2. ALGORITHM

In this section, we present our method of solving nonlinear equation (1) with convex constraints. Our approach is based on the Hybrid-Syarafina-Mustafa-Rivaie (HSMR) (Mohamed et al., 2020) approach for the unconstrained optimisation problem. The HSMR approach solves

$$\min f(k), k \in \mathcal{R}^n, \qquad (2)$$

where $f: \mathcal{R}^n \to \mathcal{R}$ (where $\mathcal{R}$ is the set of real numbers) is smooth and its gradient $\nabla f(k)$ is denoted by $g(k)$. The Mohamed (2020) method generates a sequence of iterates $\{k_t\}$ by the following recursive scheme:

$$k_{t+1} = k_t + \alpha_t j_t, t \geq 0, \qquad (3)$$

where k_t is the current iterative point and $k_0 \in \mathcal{R}^n$ is the starting point of the sequence. In (3), $\alpha_t > 0$ is known as the step length and j_t is the search direction defined by the rule:

$$j_t := \begin{cases} -g_t & \text{if } t = 0, \\ -g_t + \delta_t^{HSMR} j_{t-1} & \text{if } t > 0, \end{cases} \quad (4)$$

where $g_t = g(k_t)$ and the conjugate gradient parameter δ_t^{HSMR} is defined as

$$\delta_t^{HSMR} := \max\{0, \min\{\delta_t^{SMR}, \delta_t^{RMIL}\}\}, \quad (5)$$

where $\delta_t^{SMR} := \max\{0, \frac{\|g_t\|^2 - |g_t^T g_{t-1}|}{\|j_{t-1}\|^2}\}$ and $\delta_t^{RMIL} := \frac{g_t^T y_{t-1}}{\|j_{t-1}\|^2}, y_{t-1} := g_t - g_{t-1}$.

Next, we implement a derivative-free projection method, based on the HSMR method, to solve (1). The method we propose generates a sequence of iterates using the following relation:

$$v_t = k_t + \alpha_t j_t \quad (6)$$

$$j_t := \begin{cases} -T_t & \text{if } t = 0, \\ -\theta_t T_t + \delta_t^{EHSMR} j_{t-1} & \text{if } t > 0, \end{cases} \quad (7)$$

where $T_t = T(k_t)$ and δ_t^{EHSMR} is defined as

$$\delta_t^{EHSMR} := \max\{0, \min\{\delta_t^{ESMR}, \delta_t^{ERMIL}\}\}, \quad (8)$$

where

$$\delta_t^{ESMR} := \max\{0, \frac{\| T_t \|^2 - |T_t^T T_{t-1}|}{\| j_{t-1} \|^2}\}$$

and

$$\delta_t^{ERMIL} := \frac{T_t^T y_{t-1}}{\| j_{t-1} \|^2}, y_{t-1} := T_t - T_{t-1}.$$

It can be observed that, for $t = 0$, the direction (7) obviously satisfies the descent condition; that is, for all $t \geq 0$, if j_t is generated by Algorithm 1, then

$$T_t^T j_t = -c \| T_t \|^2, c > 0. \quad (9)$$

We first note that

$$\delta_t^{EHSMR} \leq \delta_t^{ERMIL}. \quad (10)$$

Thus, for $t > 0$, we have

$$T_t^T j_t = -\left(\theta_t - \frac{\|y_{t-1}\|}{\|j_{t-1}\|}\right) \| T_t \|^2.$$

To satisfy (9), it is only necessary that

$$\theta_t \geq c + \frac{\| y_{t-1} \|}{\| j_{t-1} \|}.$$

In this paper, we choose θ_t as

$$\theta_t = c + \frac{\| y_{t-1} \|}{\| j_{t-1} \|}.$$

Definition 1

Let $\Lambda \subseteq \mathcal{R}^n$ be a nonempty closed convex set. Then, for any $b \in \mathcal{R}^n$, its projection onto Λ, denoted by $P_\Lambda[b]$, is defined by

$$P_\Lambda[b] := \text{argmin}\{\| b - a \| : a \in \Lambda\}.$$

For any $b, a \in \mathcal{R}^n$, the projection operator P_Λ has the following nonexpansive property:

$$\| P_\Lambda[y] - P_\Lambda[x] \| \leq \| y - x \|. \tag{11}$$

In what follows, we state our method's iterative procedures/steps.

Algorithm 1

Input. Set an initial point $k_0 \in \Lambda$ and the positive constants $Tol > 0, r \in (0,1), x \in (0,2), a > 0, \mu > 0$. Set $t = 0$.

Step 0. Compute T_t. If $\| T_t \| \leq Tol$, stop. Otherwise, generate the search direction j_t using (7).

Step 1. Determine the step size $\alpha_t = \max\{ar^m | m \geq 0\}$ such that

$$T(k_t + \alpha_t j_t)^T j_t \geq \mu \alpha_t \| j_t \|^2. \tag{12}$$

Step 2. Compute $v_t = k_t + \alpha_t j_t$, where v_t is a trial point.

Step 3. If $v_t \in \Lambda$ and $\| T(v_t) \| = 0$, stop. Else, compute

$$k_{t+1} = P_\Lambda\left[k_t - x \frac{T(v_t)^T (k_t - v_t)}{\| T(v_t) \|^2} T(v_t)\right]. \tag{13}$$

Step 4. Let $t = t + 1$. Then, return to step 1.

3. GLOBAL CONVERGENCE

In this section, we obtain the global convergence property of Algorithm 1 and list the following assumptions on the mapping T.

Assumption 1

(i) The solution set of constrained nonlinear equation (1), denoted by Λ^*, is nonempty.

(ii) The mapping T is Lipschitz continuous on $\mathcal{R}^n$. That is, there exists a constant $L > 0$ such that

$$\| T(\alpha) - T(\delta) \| \leq L \| \alpha - \delta \| \; \forall \alpha, \delta \in \mathcal{R}^n. \tag{14}$$

(iii) For any $\delta \in \Lambda^*$ and $\alpha \in \mathcal{R}^n$, it holds that

$$T(\alpha)^T(\alpha - \delta) \geq 0. \quad (15)$$

Lemma 1
Let $\{j_t\}$ and $\{k_t\}$ be two sequences generated by Algorithm 1. Then, there exists a step size α_t satisfying the line search (12) for all $t \geq 0$.

Proof. For any $m \geq 0$, suppose (12) does not hold for the iterate t_0 −th. Then, we have

$$-T(k_{t_0} + ar^m j_{t_0})^T j_{t_0} < \mu a r^m \parallel j_{t_0} \parallel^2.$$

Thus, by the continuity of T and with $0 < r < 1$, it follows that by letting $m \to \infty$, we have

$$-T(k_{t_0})^T j_{t_0} \leq 0,$$

which contradicts (9).

Lemma 2
Let the sequences $\{k_t\}$ and $\{v_t\}$ be generated by the Algorithm 1 method under Assumption 1. Then,

$$\alpha_t \geq \max\left\{a, \frac{rc\|T_t\|^2}{(L+\mu)\|j_t\|^2}\right\}. \quad (16)$$

Proof. Let $\hat{\alpha}_t = \alpha_t r^{-1}$. Assume $\alpha_t \neq a$. However, $\hat{\alpha}_t$ does not satisfy (12). That is,

$$-T(k_t + \hat{\alpha}_t j_t)^T j_t < \mu\hat{\alpha}_t \parallel j_t \parallel^2.$$

From (14) and (9), it can be seen that

$$\begin{aligned} c \parallel T_t \parallel^2 &\leq -T_t^T j_t \\ &= (T(k_t + \hat{\alpha}_t j_t) - T_t)^T j_t - T(k_t + \hat{\alpha}_t j_t)^T j_t \\ &\leq L\hat{\alpha}_t \parallel j_t \parallel^2 + \mu\hat{\alpha}_t \parallel j_t \parallel^2 \\ &\leq \hat{\alpha}_t(L+\mu) \parallel j_t \parallel^2. \end{aligned}$$

This gives the desired inequality (16).

Lemma 3
Suppose that Assumption 1 holds. Let $\{k_t\}$ and $\{v_t\}$ be sequences generated by Algorithm 1. Then, for any solution k^* contained in the solution set Λ^*, the inequality

$$\parallel k_{t+1} - k^* \parallel^2 \leq \parallel k_t - k^* \parallel^2 - \mu^2 \parallel k_t - v_t \parallel^4 \quad (17)$$

holds. In addition, $\{k_t\}$ is bounded and

$$\sum_{t=0}^{\infty} \parallel k_t - v_t \parallel^4 < +\infty. \quad (18)$$

Proof. First, we begin by using the weak monotonicity assumption from Assumption 1 (ii) on the mapping T. Thus, for any solution $k^* \in \Lambda^*$,

$$T(v_t)^T(k_t - k^*) \geq T(v_t)^T(k_t - v_t).$$

The above inequality, together with (12), gives

$$T(k_t + \alpha_t j_t)^T(k_t - v_t) \geq \mu\alpha_t^2 \parallel j_t \parallel^2 \geq 0. \tag{19}$$

From (11) and (19), we have the following:

$$\begin{aligned}
\parallel k_{t+1} - k^* \parallel^2 &= \left\| j_\Lambda \left[k_t - x\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel^2} T(v_t) \right] - k^* \right\|^2 \\
&\leq \left\| \left[k_t - x\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel^2} T(v_t) \right] - k^* \right\|^2 \\
&= \parallel k_t - k^* \parallel^2 - 2x\left(\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel^2}\right) T(v_t)^T(k_t - k^*) \\
&\quad + x^2\left(\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel}\right)^2 \\
&= \parallel k_t - k^* \parallel^2 - 2x\left(\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel^2}\right) T(v_t)^T(k_t - v_t) \\
&\quad + x^2\left(\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel}\right)^2 = \\
&\parallel k_t - k^* \parallel^2 - x(2 - x)\left(\frac{T(v_t)^T(k_t - v_t)}{\parallel T(v_t) \parallel}\right)^2 \leq \parallel k_t - k^* \parallel^2.
\end{aligned}$$

Thus, the sequence $\{\parallel k_t - k^* \parallel\}$ has a nonincreasing and convergent property. Therefore, $\{k_t\}$ is a bounded sequence; that is, $\forall t$,

$$\parallel k_t \parallel \leq b_0, b_0 > 0, \tag{20}$$

and therefore the following holds:

$$\sigma^2 \sum_{t=0}^{\infty} \parallel k_t - v_t \parallel^4 < \parallel k_0 - k^* \parallel^2 < +\infty.$$

Remark 1

Considering the definition of v_t and also (18), it can be deduced that

$$\lim_{t\to\infty} \alpha_t \parallel j_t \parallel = 0. \tag{21}$$

Theorem 1

Suppose Assumption 1 holds. Let $\{k_t\}$ and $\{v_t\}$ be sequences generated by Algorithm 1. Then,

$$\liminf_{t\to\infty} \parallel T_t \parallel = 0. \tag{22}$$

Proof. Suppose (22) is not valid; that is, there exists a constant, say, $s > 0$ such that $s \leq \| T_t \|$, $t \geq 0$. Then this, along with (9), implies that

$$\| j_t \| \geq cs, \quad \forall t \geq 0. \tag{23}$$

It can be seen from Lemma 3 and Remark 1 that the sequences $\{k_t\}$ and $\{v_t\}$ are bounded. Also, the continuity of T further implies that $\{\| T_t \|\}$ is bounded by a constant, say, u. From (7) and (10), it follows that for all $t \geq 1$,

$$\begin{aligned} \| j_t \| &= \|-\theta_t T_t + \delta_t^{EHSMR} j_{t-1}\| \\ &= \left\| -\left(c + \frac{\| y_{t-1} \|}{\| j_{t-1} \|}\right) T_t + \delta_t^{EHSMR} j_{t-1} \right\| \\ &\leq c \| T_t \| + \frac{\| y_{t-1} \|}{\| j_{t-1} \|} \| T_t \| + |\delta_t^{EHSMR}| \| j_{t-1} \| \\ &\leq c \| T_t \| + \frac{\| y_{t-1} \|}{\| j_{t-1} \|} \| T_t \| + \frac{\| y_{t-1} \|}{\| j_{t-1} \|} \| T_t \| \\ &= c \| T_t \| + 2\frac{\| y_{t-1} \|}{\| j_{t-1} \|} \| T_t \| \\ &\leq c \| T_t \| + 2\frac{L \| k_t - k_{t-1} \|}{\| j_{t-1} \|} \\ &\leq cu + 4L\frac{b_0 u}{cs} \triangleq \gamma. \end{aligned}$$

From (16), we have

$$\begin{aligned} \alpha_t \| j_t \| &\geq \max\left\{a, \frac{rc \| T_t \|^2}{(L+\mu) \| j_t \|^2}\right\} \| j_t \| \\ &\geq \max\{acs, \frac{rcs^2}{(L+\mu)\gamma}\} > 0, \end{aligned}$$

which contradicts (21). Hence, (22) is valid.

4. NUMERICAL EXPERIMENT

This section assesses the computational efficiency of the proposed algorithm using the Dolan and Moré (2002) performance profile, whose metric takes into consideration the number of iterations, the number of function evaluations and the running time of the CPU. The performance of Algorithm 1, which we will refer to as Extended HSMR (EHSMR), is compared with the derivative-free iterative method proposed in Ibrahim et al. (2019) and Lieu and Feng (2019). We refer to these two methods as ERMIL and PDY. We note that all codes were coded and implemented in the MATLAB environment. The control parameters for EHSMR were chosen as $a = 1, r = 0.75, \mu = 10^{-4}, x = 1.2$ and $Tol = 10^{-6}$. The parameters for ERMIL and PDY were set as reported in the numerical section of their respective papers. We made use of various dimensions, including $1000, 5000, 10{,}000, 50{,}000, 100{,}000$ and different initial points: $k_1 = (0.1,0.1,\cdots,0.1)^T$, $k_2 = (0.2,0.2,\cdots,0.2)^T$, $k_3 = (0.5,0.5,\ldots,0.5)^T$, $k_4 = (1.2,1.2,\cdots,1.2)^T$, $k_5 = (1.5,1.5,\cdots,1.5)^T$, $k_6 = (2,2,\ldots,2)^T$ and $k_7 = \text{rand}(0,1)$. In what follows, we give the list of test problems used for the experiment. We note that the $T = (T_1, T_2, \cdots, T_n)$ are given below:

Problem 1 (Ding et al., 2017)

$$t_i = \sum_{i=1}^{n} k_i^2, d = 10^{-5}.$$
$$T_i(k) = 2d(k_i - 1) + 4(t_i - 0.25)k_i, i = 1,2,3, \ldots, n \text{ and } \quad \Lambda = \mathcal{R}_+^n.$$

Problem 2 The Trig exp function, (La Cruz et al., 2006)

$$T_1(k) = 3k_1^3 + 2k_2 - 5 + \sin(k_1 - k_2)\sin(k_1 + k_2)$$
$$T_i(k) = 3k_i^3 + 2k_{i+1} - 5 + \sin(k_i - k_{i+1})\sin(k_i + k_{i+1}) + 4k_i - k_{i-1}e^{k_{i-1}-k_i} - 3$$
$$T_n(k) = k_{n-1}e^{k_{n-1}-k_n} - 4k_n - 3, \quad \text{where h} = \frac{1}{m+1} \text{and } \Lambda = \mathcal{R}_+^n.$$

Problem 3 Nonsmooth Function, **3** (Yu et al., 2009)

$$T_i(k) \quad = k_i - \sin|k_i - 1|, i = 1,2,3, \ldots, n,$$
$$\text{and} \Lambda \quad = \left\{ k \in \mathcal{R}^n : \sum_{i=1}^{n} k_i \leq n, k_i \geq -1, i = 1,2, \cdots, n \right\}.$$

Problem 4 Tridiagonal Exponential Function, (Bing & Lin, 1991)

$$T_1(k) \quad = k_1 - e^{\cos(h(k_1+k_2))},$$
$$T_i(k) \quad = k_i - e^{\cos(h(k_{i-1}+k_i+k_{i+1}))}, \text{for} i = 2, \ldots, n-1,$$
$$T_n(k) \quad = k_n - e^{\cos(h(k_{n-1}+k_n))},$$
$$h \quad = \frac{1}{n+1}.$$

Problem 5 Strictly Convex Function II, (La Cruz et al., 2006)

$$T_i(k) \quad = \frac{i}{n}e^{k_i} - 1, \text{for} i = 1,2, \ldots, n,$$
$$\text{and} \quad \Lambda = \mathcal{R}_+^n.$$

Problem 6 Strictly Convex Function I, (La Cruz et al., 2006)

$$T_i(k) \quad = e^{k_i} - 1, \text{for} i = 1,2, \ldots, n,$$
$$\text{and} \quad \Lambda = \mathcal{R}_+^n.$$

Problem 7 (Ding et al., 2017)

$$T_i(k) = \min(\min(|k_i|, k_i^2), \max(|k_i|, k_i^3)) \text{for} i = 2,3, \ldots, n,$$
$$\text{and } \Lambda = \mathcal{R}_+^n.$$

Problem 8 Modified Logarithmic Function, (La Cruz et al., 2006)

$$T_i(k) = \ln(k_i + 1) - \frac{k_i}{n}, \text{for} i = 1,2,3, \ldots, n,$$
$$\text{and} \quad \Lambda = \left\{k \in \mathcal{R}^n : \sum_{i=1}^{n} k_i \leq n, k_i > -1, i = 1,2, \cdots, n\right\}.$$

Problem 9 Exponential Function, (La Cruz et al., 2006)

$$T_1(k) = e^{k_1} - 1,$$
$$T_i(k) = e^{k_i} + k_i - 1, \text{for} i = 2,3, \ldots, n,$$
$$\text{and} \quad \Lambda = \mathcal{R}^n_+.$$

The performance results of the methods for problems 1–5 are presented in Tables 1–5 of the appendix section which can be found in the following link; https://documentcloud.adobe.com/link/review?uri=urn:aaid:scds:US:738a210f-a99-4260-9612-c7ab7d34dcaa.

In Tables 1–5, 'dm' denotes the dimension, 'inp' denotes the initial points, 'it' denotes the iteration number, 'nf' denotes the function evaluation number and 'tm' denotes the CPU running time. Figure 1 shows the iteration performance profiles of the three methods. The EHSMR algorithm's results can be seen in the top curve. EHSMR outperformed ERMIL and PDY, with EHSMR solving 69% of the test problems with few iterations and ERMIL and PDY solving about 41% and 21%, respectively. The profile of the number of function evaluations is reported in Figure 2. We note that EHSMR performed better than the other two methods. EHSMR was able to solve about 63% of the test problems with few iterations, while ERMIL and PDY were able to solve about 40% and 20%, respectively. Figure 3 shows the CPU time performance profiles. EHSMR required the shortest CPU time.

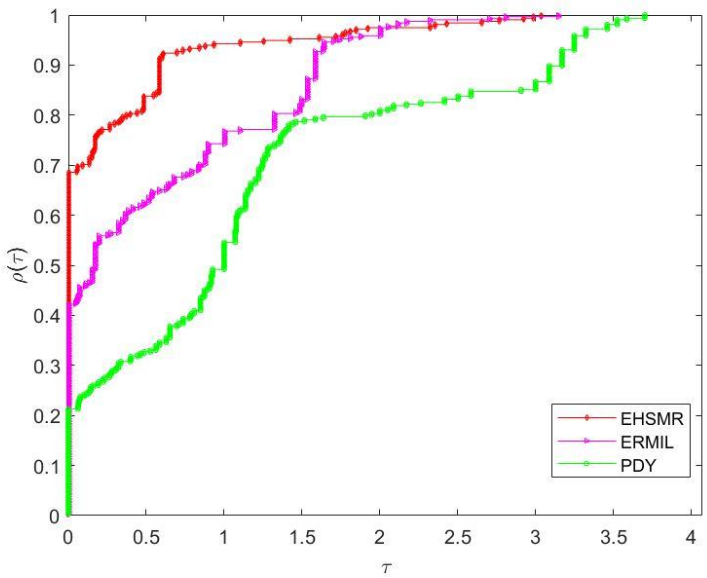

Figure 1. Performance profiles based on number of iterations

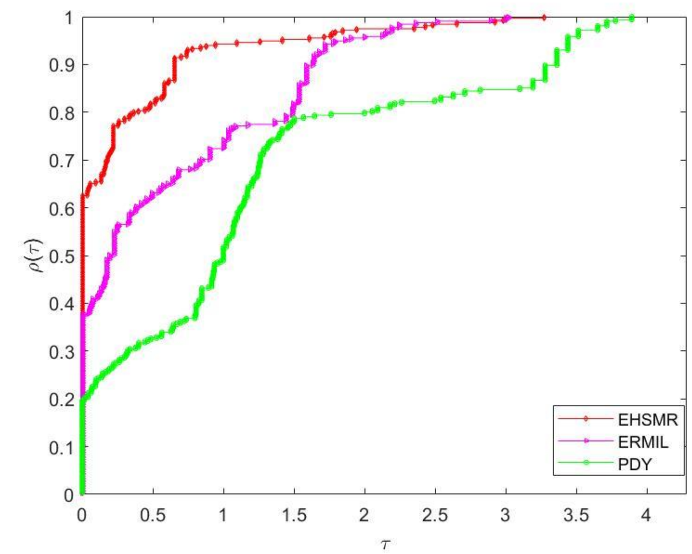

Figure 2. Performance profiles based on the number of function evaluations.

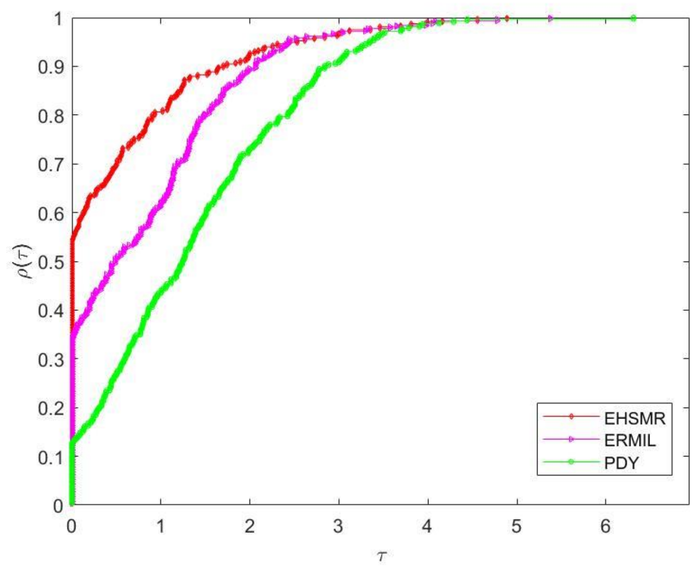

Figure 3. Performance profiles based on the CPU time (in seconds).

5. CONCLUSION

In this paper, we have extended the HSMR conjugate gradient method for unconstrained optimisation problems to solve a nonlinear equation with convex constraints. The proposed method is derivative-free and satisfies the sufficient descent condition. Global convergence is proved under the assumption that the underlying mapping is Lipschitz continuous and satisfies a weaker monotonicity condition. Numerical experiments show that the proposed method is efficient.

6. ACKNOWLEDGEMENT

The author was financially supported by the Rajamangala University of Technology Phra Nakhon (RMUTP) Research Scholarship.

7. REFERENCES

Abubakar, A. B., Kumam, P., Ibrahim, A. H., & Rilwan, J. (2020). Derivative-free HS-DY-type method for solving nonlinear equations and image restoration. Heliyon, 6(11), e05400.

Abubakar, A. B., Rilwan, J., Yimer, S. E., Ibrahim, A. H., & Ahmed, I. (2020). Spectral three-term conjugatedescent method for solving nonlinear monotone equations with convex constraints. Thai Journal of Mathematics, 18(1), 501–517.

Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., & Plemmons, R. J. (2007). Algorithms and applications for approximate nonnegative matrix factorization. Computational Statistics & Data Analysis, 52(1), 155–173.

Bing, Y., & Lin, G. (1991). An efficient implementation of Merrill's method for sparse or partially separable systems of nonlinear equations. SIAM Journal on Optimization, 1(2), 206–221.

Blumensath, T. (2013). Compressed sensing with nonlinear observations and related nonlinear optimization problems. IEEE Transactions on Information Theory, 59(6), 3466–3474.

Dai, Z., Dong, X., Kang, J., & Hong, L. (2020). Forecasting stock market returns: New technical indicators and two-step economic constraint method. The North American Journal of Economics and Finance, 101216.

Dennis, J. E., & More, J. J. (1974). A Characterization of Superlinear Convergence and Its Application to Quasi-Newton Methods. Mathematics of Computation, 28(126), 549. https://doi.org/10.2307/2005926

Dennis Jr, J. E., & Moré, J. J. (1977). Quasi-Newton methods, motivation and theory. SIAM Review, 19(1), 46–89.

Ding, Y., Xiao, Y., & Li, J. (2017). A class of conjugate gradient methods for convex constrained monotone equations. Optimization, 66(12), 2309–2328. https://doi.org/10.1080/02331934.2017.1372438

Dirkse, S. P., & Ferris, M. C. (1995). Mcplib: A collection of nonlinear mixed complementarity problems. Optimization Methods and Software, 5(4), 319–345. https://doi.org/10.1080/10556789508805619

Djordjević, S. S. (2019). New hybrid conjugate gradient method as a convex combination of LS and FR methods. Acta Mathematica Scientia, 39(1), 214–228. https://doi.org/10.1007/s10473-019-0117-6

Huang, N., Ma, C., & Xie, Y. (2016). The Derivative-Free Double Newton Step Methods for Solving System of Nonlinear Equations. Mediterranean Journal of Mathematics, 13(4), 2253–2270.

Ibrahim, A. H., Garba, A. I., Usman, H., Abubakar, J., & Abubakar, A. B. (2019). Derivative-free RMIL conjugate gradient method for convex constrained equations. Thai Journal of Mathematics, 18(1), 212–232.

Ibrahim, A. H., Kumam, P., Abubakar, A. B., Abubakar, J., & Muhammad, A. B. (2020). Least-Square-Based Three-Term Conjugate Gradient Projection Method for ell_1-Norm Problems with Application to Compressed Sensing. Mathematics, 8(4), 602. https://doi.org/https://doi.org/10.3390/math8040602

Ibrahim, A. H., Kumam, P., Abubakar, A. B., Jirakitpuwapat, W., & Abubakar, J. (2020). A hybrid conjugate gradient algorithm for constrained monotone equations with application in compressive sensing. Heliyon, 6(3), e03466. https://doi.org/10.1016/j.heliyon.2020.e03466

Ibrahim, A. H., Kumam, P., Abubakar, A. B., Yusuf, U. B., Yimer, S. E., & Aremu, K. O. (2020). An efficient gradient-free projection algorithm for constrained nonlinear equations and image restoration. AIMS Mathematics, 2021(6), 1.

Ibrahim, A. H., Kumam, P., & Kumam, W. (2020). A family of derivative-free conjugate gradient methods for constrained nonlinear equations and image restoration. IEEE Access, 8, 162714–162729.

La Cruz, W., Martínez, J. M., & Raydan, M. (2006). Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. Mathematics of Computation, 75(255), 1429–1449. https://doi.org/10.1090/s0025-5718-06-01840-0

Lee, D. D., & Seung, H. S. (2001). Algorithms for non-negative matrix factorization. Advances in Neural Information Processing Systems, 556–562.

Mohamed, N. S., Mamat, M., Rivaie, M., & Shaharudin, S. M. (2020). A new hyhbrid coefficient of conjugate gradient method. Indonesian Journal of Electrical Engineering and Computer Science, 18(3), 1454–1463.

Mohammad, H., & Bala Abubakar, A. (2020). A descent derivative-free algorithm for nonlinear monotone equations with convex constraints. RAIRO - Operations Research, 54(2), 489–505. https://doi.org/10.1051/ro/2020008

Qi, L., & Sun, J. (1993). A nonsmooth version of Newton’s method. Mathematical Programming, 58(1–3), 353–367. https://doi.org/10.1007/BF01581275

Solodov, M. V., & Svaiter, B. F. (1999). A new projection method for variational inequality problems. SIAM Journal on Control and Optimization, 37(3), 765–776. https://doi.org/10.1137/S0363012997317475

Yamashita, N., & Fukushima, M. (2001). On the rate of convergence of the Levenberg-Marquardt method. In Topics in numerical analysis (pp. 239–249). Springer.

Yu, Z., Lin, J., Sun, J., Xiao, Y., Liu, L., & Li, Z. (2009). Spectral gradient projection method for monotone nonlinear equations with convex constraints. Applied Numerical Mathematics, 59(10), 2416–2423.

Zhang, L., & Zhou, W. (2006). Spectral gradient projection method for solving nonlinear monotone equations. Journal of Computational and Applied Mathematics, 196(2), 478–484.