CARPATHIAN J. MATH.
Volume **40** (2024), No. 2,
Pages 515 - 533

Online version at https://www.carpathian.cunbm.utcluj.ro/
Print Edition: ISSN 1584 - 2851; Online Edition: ISSN 1843 - 4401
DOI: https://doi.org/10.37193/CJM.2024.02.18

Efficient nonlinear conjugate gradient techniques for vector optimization problems

JAMILU YAHAYA[1,3] POOM KUMAM [1,2*] and JAMILU ABUBAKAR [4]

ABSTRACT. Conjugate gradient techniques are known for their simplicity and minimal memory usage. However, it is known that in the vector optimization context, the Polak-Ribiére-Polyak (PRP), Liu-Storey (LS), and Hestenes-Stiefel (HS) conjugate gradient (CG) techniques fail to satisfy the sufficient descent property using Wolfe line searches. In this work, we propose a variation of the PRP, LS, and HS CG techniques that we termed YPR, YLS, and YHS, respectively. These techniques exhibit the desirable property of sufficient descent without line search, except for the YHS which uses Wolfe line search for its sufficient descent property. Under certain standard assumptions and employing strong Wolfe conditions, we investigate the global convergence properties of the proposed techniques. The global convergence analysis extends beyond convexity assumption on the objective functions. Additionally, we present numerical experiments and comparisons to demonstrate the implementation, efficiency, and robustness of the proposed techniques.

1. INTRODUCTION

Conjugate gradient (CG) techniques for solving vector optimization problems (VOPs) have gained substantial attention from researchers since their introduction to vector setting in 2018 by Lucambio Pérez and Prudente, [43]. These techniques have captured interest due to their simplicity, minimal memory usage, and suitability for large-scale problems.

Before we extensively explore the main topic of discussion, let us begin by considering an unconstrained single-objective problem of minimizing a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Several β_k parameters were studied in the literature for this problem, these include the Polak–Ribiére–Polyak (PRP) [47], Hestenes–Stiefel (HS) [31] and Liu–Storey (LS) [40], which are defined respectively as follows:

$$\beta_k^{PRP} := \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad \beta_k^{HS} := \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{LS} := -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}},$$

where $y_{k-1} := (g_k - g_{k-1})$, $g_k = \nabla f(z_k)$, and $||\cdot||$ is the Euclidean norm. Other well-known CG techniques are the Fletcher-Reeves (FR) [16], Conjugate Descent (CD) [17], and Dai-Yuan (DY) [11]. For each β_k, a search direction d_k needs to satisfies a descent property given by $g_k^T d_k \leq 0$, for all $k \geq 1$. There are many other choices for the parameter β_k, we briefly listed the most common ones as seen above. For other choices of the β_k parameter, see for example [3, 29, 56, 61], and the references therein.

The CG techniques FR, CD, and DY have one distinguishing characteristic, that is, their search directions satisfy sufficient descent condition (SDC)

$$g_k^T d_k \leq -c\|g_k\|^2, \ \forall\, k \geq 1, \tag{1.1}$$

Received: 05.09.2023. In revised form: 12.02.2024. Accepted: 19.02.2024
2020 *Mathematics Subject Classification.* 90C29, 90C52, 90C30, 90C26, 49M37 .
Key words and phrases. *Conjugate gradient method; Pareto-optimality; sufficient descent condition; vector optimization.*
Corresponding author: Poom Kumam; poom.kum@kmutt.ac.th

with $c > 0$, when a Wolfe line search is utilized. In contrast, the PRP, HS, and LS do not necessarily satisfy (1.1), see, for instance, Powell [48].

Wei et al. [55] introduced new variations of CD and FR techniques for solving unconstrained large-scale optimization. The proposed techniques encompass several crucial properties, including the satisfaction of the SDC without any line search and guaranteeing the Zoutendijk condition is satisfied when a line search technique is used. Moreover, the techniques exhibit specific characteristics of the PRP CG technique. The authors established the global convergence of the proposed techniques by applying the standard Wolfe condition (WWC) and the standard Armijo line search. Since then, numerous other extensions or modifications of these CG techniques have been explored. See, for instance, [34, 59, 60] and the references therein.

In the following, we consider an unconstrained vector optimization problem (VOP) defined in the following form

$$\text{Minimize}_Q\ F(z), \quad z \in \mathbb{R}^n, \tag{1.2}$$

where $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ in C^1 (continuously differentiable functions), $Q \subset \mathbb{R}^m$ is closed, convex and pointed cone with nonempty-interior. We emphasize that VOP, as shown in [36] and [53], can represent several problems across science, engineering, and social sciences.

The VOPs have applications in diverse fields, including bi-level programming, cancer treatment planning, engineering, environmental analysis, location science, management science, and statistics. Notable examples include, [12, 20, 21, 28, 33, 37, 39, 53] and the references therein. The partial order defined in $\mathbb{R}^m$, $\preccurlyeq_Q$, generated by Q (respectively, $\prec_Q$, generated by $int(Q)$) is given by

$$a \preccurlyeq_Q b \Longleftrightarrow b - a \in Q,$$

respectively,

$$a \prec_Q b \Longleftrightarrow b - a \in int(Q).$$

Now, beginning with an initial point $z_1 \in \mathbb{R}^n$, the CG technique in the vector setting, recursively generates a sequence of iterations as

$$z_{k+1} = z_k + t_k d_k,\ k \geq 1, \tag{1.3}$$

where the step size $t_k > 0$, is acquired via a line search technique and the search direction d_k is given as

$$d_k := \begin{cases} u(z_k), & k = 1, \\ u(z_k) + \beta_k d_{k-1}, & k \geq 2, \end{cases} \tag{1.4}$$

where β_k is a scalar parameter.

One of the primary solution strategies for VOPs is scalarization approaches. Here, multiobjective optimization problems are parameterized by reducing to single-objective optimization problems and solved, resulting in a corresponding number of Pareto-optimal points. In this approach, the decision-maker must select the parameters because they are not predetermined. For some problems, making this choice can pose significant challenges or become impossible. Consequently, to overcome these drawbacks, some descent-based algorithms, including the conjugate gradient algorithm, have been suggested for solving VOPs.

Over the last two decades, there has been a growing interest in adapting descent-based algorithms, initially designed for single-objective optimization, extending to VOPs. This trend can be traced back to earlier work in 2005, such as [15] and [6]. Subsequently, numerous other studies have followed this trajectory, exploring similar directions [2, 4, 7–9, 18, 22, 24, 26, 27, 44, 50], and the references therein.

The authors in [43] presented the broad concept of Wolfe and Zoutendijk conditions for VOPs. In particular, they extended and studied certain properties of the following CG parameters

$$\beta_k^{FR} := \frac{\zeta(z_k, u(z_k))}{\zeta(z_{k-1}, u(z_{k-1}))}, \quad \beta_k^{CD} := \frac{\zeta(z_k, u(z_k))}{\zeta(z_{k-1}, d_{k-1})}, \quad \beta_k^{DY} := \frac{-\zeta(z_k, u(z_k))}{\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, d_{k-1})},$$

$$\beta_k^{PRP} := \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-\zeta(z_{k-1}, u(z_{k-1}))}, \quad \beta_k^{HS} := \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, d_{k-1})},$$

where $\zeta(\cdot, \cdot)$ will be defined in the next section. Their study encompassed numerical implementations of these techniques, which were analyzed and discussed. Among these techniques, the nonnegative PRP and HS exhibited superior performance compared to the others, while DY and CD surpassed FR. These are extensions of β_k that were originally proposed for an unconstrained single-objective optimization in [1, 10, 11, 25] to the vector setting.

Furthermore, Goncalves et al. in [26] presented the following LS CG technique and two of its variants

$$\beta_k^{LS} := \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-\zeta(z_{k-1}, d_{k-1})}.$$

The search direction of the LS could not satisfy the SDC. However, one of their variants satisfied this property. They investigated the global convergence of these techniques using both Wolfe and Armijo line searches. Their numerical experiments suggest that the technique with the Wolfe line search is competitive with the modified technique where the Armijo line search is used. Therefore, in general the search directions of the PRP, LS, and HS CG techniques, as defined above, could not establish sufficient descent property. Some few other researches in this direction emerges, see for instance [30, 57, 58].

This research is motivated by the works [55, 59]. We introduced three new CG techniques (which we termed YPR, YLS, and YHS) for VOPs that exhibit the SDC without line search, except for the YHS which uses Wolfe line search for its sufficient descent property. Additionally, we establish the global convergence of these techniques using strong Wolfe line search. We show that the sequence generated by our proposed techniques identifies a point that satisfies the first-order necessary condition for Pareto-optimality. These are obtained under appropriate assumptions. Importantly, our comprehensive analysis does not rely on convex assumption on the objective functions. We extensively discuss the results of the numerical experiments, which aim to demonstrate the effectiveness, efficiency, and robustness of the proposed techniques. A comprehensive comparison of these results with nonnegative HS technique, is provided. To the best of our knowledge, the proposed β_k parameters are the first nonnegative PRP and HS variants to satisfy the SDC in the vector optimization setting.

The paper is structured as follows: Section 2 provides the basics and preliminary results, such as lemmas and definitions associated with VOPs. Section 3 presents the convergence analysis: the sufficient descent conditions and the global convergence of the proposed CG techniques. Section 4 presents the numerical results and discusses the comparison with other existing technique. Finally, in Section 5, we conclude and provide closing remarks.

2. Preliminaries

In this section, we present some definitions, basic notions, and lemmas related to VOPs that will subsequently be used in this paper. For some notable preliminaries, the reader is referred to [14, 15, 19, 42, 43].

The concept of optimality is replaced by *Pareto-optimal or Pareto-efficient* in VOP. Thus, in this context, we define Pareto-optimal or Pareto-efficient:

Definition 2.1. [23] A point $\bar{z} \in \mathbb{R}^n$ is Pareto-optimal or efficient if and only if there does not exists a point $z \in \mathbb{R}^n$ such that $F(z) \preccurlyeq_Q F(\bar{z})$ and $F(z) \neq F(\bar{z})$,

and weak Pareto-optimal or weak Pareto-efficient

Definition 2.2. [23] A point $\bar{z} \in \mathbb{R}^n$ is weak Pareto-optimal or weak Pareto-efficient if and only if there does not exists a point $z \in \mathbb{R}^n$ such that $F(z) \prec_Q F(\bar{z})$.

Remark 2.1. If $\bar{z} \in \mathbb{R}^n$ represents a Pareto-optimal point, it also qualifies as a weak Pareto point. However, the reverse statement is often not true.

The positive polar cone of Q is

$$Q^* := \{p \in \mathbb{R}^m \mid \langle p, z\rangle \geq 0,\ \forall\, z \in Q\}.$$

Now, we state some properties of Q and Q^*. Note that since Q is closed and convex, then, $Q = Q^{**}$,

$$-Q = \{z \in \mathbb{R}^m \mid \langle z, p\rangle \leq 0,\ \forall\, p \in Q^*\}\ and\ -int(Q) = \{z \in \mathbb{R}^m \mid \langle z, p\rangle < 0,\ \forall\, p \in Q^* \setminus \{0\}\}.$$

A cone generated by $S \subseteq R^m$ is denoted by $cone(S)$ and a convex hull of S is denoted by $conv(S)$. Now, suppose $C \subseteq Q^*$ and $0 \notin C$ is compact, and define Q^* as:

$$Q^* = cone(conv(C)). \tag{2.5}$$

For instance, in multiobjective optimization setting, $Q = \mathbb{R}^m_+$, implies $Q^* = Q$ and C is taken to be the canonical basis in $\mathbb{R}^m$. If Q is a polyhedral cone, then Q^* is also a polyhedral cone. Additionally, C can be considered as the finite set of extremal rays of polyhedral cone Q^*. Now, for a generic Q (closed, convex and pointed cone with nonempty-interior), we have

$$C = \{p \in Q^* \mid \|p\| = 1\}, \tag{2.6}$$

satisfying (2.5). Throughout this paper we consider C to be defined as (2.6). The first order derivative and the Jacobian of F at z is represented as $JF(z)$. While the Image($JF(z)$) represents the image on $\mathbb{R}^m$ by $JF(z)$. A necessary condition for $Q-$optimality of $\bar{z} \in \mathbb{R}^n$ is given as

$$-int(Q) \cap Image(JF(\bar{z})) = \emptyset.$$

If this condition holds, then the point $z \in \mathbb{R}^n$ is said to be a *stationary* or $Q-$*critical Pareto*. However, if $z \in \mathbb{R}^n$ is not $Q-$critical, then there exists $h \in \mathbb{R}^n$ such that $JF(z)h \in -int(Q)$, this indicates that h is a $Q-$descent direction for F at z, that is, there exists $\epsilon > 0$ such that $F(z + rh) \prec_Q F(z)$ for all $0 < r < \epsilon$, see e.g., [42] for a full discussion on this.

Define $\theta : \mathbb{R}^m \to \mathbb{R}$ as

$$\theta(z) := \sup\{\langle z, p\rangle \mid p \in C\}.$$

By the compactness of C, we have that θ is well-defined. Notice that θ also provides some features of $-Q$ and $-int(Q)$ as follows: $-Q = \{z \in \mathbb{R}^m \mid \theta(z) \leq 0\}$ and $-int(Q) = \{z \in \mathbb{R}^m \mid \theta(z) < 0\}$, respectively.

Now, define $\zeta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$\zeta(z, d) := \theta(JF(z)d) = \sup\{\langle JF(z)d,\ p\rangle \mid p \in C\}. \tag{2.7}$$

Definition 2.3. A $d \in \mathbb{R}^n$ is $Q-$descent direction for F at z when $\zeta(z,d) < 0$ and z is $Q-$critical point for F when $\zeta(z,d) \geq 0$ for all d.

Lemma 2.1. [15] *Suppose $F : \mathbb{R}^n \to \mathbb{R}^m$ is in C^1. Then, the statements below hold:*

(a) $\zeta(z, z' + td) \leq \zeta(z, z') + t\zeta(z,d)$, for $z, z', d \in \mathbb{R}^n$ and $t \geq 0$;

(b) the mapping $(z,d) \longmapsto \zeta(z,d)$ is continuous;

(c) $|\zeta(z,d) - \zeta(z',d)| \leq \|JF(z) - JF(z')\|\|d\|$, for $z, z', d \in \mathbb{R}^n$;

(d) if $\|JF(z) - JF(z')\| \leq L\|z - z'\|$, then $|\zeta(z,d) - \zeta(z',d)| \leq L\|d\|\|z - z'\|$.

Define $u : \mathbb{R}^n \to \mathbb{R}^n$ and $v : \mathbb{R}^n \to \mathbb{R}$ by

$$u(z) := \text{arg min} \left\{ \zeta(z,d) + \frac{\|d\|^2}{2} \mid d \in \mathbb{R}^n \right\} \tag{2.8}$$

and

$$v(z) := \zeta(z, u(z)) + \frac{\|u(z)\|^2}{2}, \tag{2.9}$$

respectively. Considering that the real-valued $\zeta(z,\cdot)$ is a closed and convex function and $d \longmapsto \frac{\|d\|^2}{2}$ is strictly convex, then $u(z)$ exists and is unique.

Now, consider a convex quadratic problem:

$$\begin{aligned} &\text{Minimize } t + \frac{1}{2}\|d\|^2, \\ &\text{subject to } [JF(z)d]_i \leq t,\ i = 1, 2, \cdots, m, \end{aligned} \tag{2.10}$$

with linear inequality constraints, see for instance, [19].

In vector setting, the search direction $d \in \mathbb{R}^n$ is said to satisfies the *sufficient descent condition* (SDC) if

$$\zeta(z,d) \leq c\zeta(z, u(z)), \tag{2.11}$$

for some $c > 0$. In addition, we say that the step size, $t > 0$ can be obtained through an exact line search if

$$\zeta(z + td, d) = 0. \tag{2.12}$$

We now give the vector Wolfe conditions that was introduced by Lucambio Pérez and Prudente [43].

Definition 2.4. [43] Suppose $d \in \mathbb{R}^n$ is a $Q-$descent and $e \in Q$, we have

$$0 < \langle p,\ e\rangle \leq 1, \tag{2.13}$$

for all $p \in C$.

Now, $t > 0$ satisfies the *standard Wolfe condition* (WWC) if

$$F(z + td) \preccurlyeq_Q F(z) + \rho t \zeta(z,d) e$$

$$\zeta(z + td, d) \geq \sigma \zeta(z,d), \tag{2.14}$$

where $0 < \rho < \sigma < 1$. Furthermore, $t > 0$ satisfies the *strong Wolfe condition* (SWC) if

$$F(z + td) \preccurlyeq_Q F(z) + \rho t \zeta(z,d) e$$

$$|\zeta(z + td, d)| \leq \sigma |\zeta(z,d)|. \tag{2.15}$$

It is interesting to know that the vector $e \in Q$ given in (2.13), invariably exists. Specifically, for multiobjective optimization $e \in Q$ is considered as $[1, \cdots, 1]^T$. The sets Q and C are considered as $\mathbb{R}^m_+$ and canonical basis of $\mathbb{R}^m$, respectively.

Let us end this section with the following important Lemmas:

Lemma 2.2. *[15]. Let $u(z)$ and $v(z)$ be defined as in* (2.8) *and* (2.9) *respectively:*

(a) *let z be a $Q-$critical for F, then $u(z) = 0$ and $v(z) = 0$,*
(b) *suppose z is not $Q-$critical for F, then $u(z) \neq 0$, $v(z) < 0$, $\zeta(z, u(z)) < -\frac{\|u(z)\|^2}{2} < 0$ and $u(z)$ $Q-$descent direction for F at z,*
(c) *The mappings u and v are continuous.*

3. Main results

In this section, we present the proposed CG techniques YPR, YLS and YHS and investigate their convergence properties.

Assumption 3.1. *Suppose that the cone Q is finitely generated and there exists an open set Δ for which $\mathcal{L} := \{z \mid F(z) \preccurlyeq_Q F(z_1)\} \subset \Delta$, where $z_1 \in \mathbb{R}^n$ and there exists $L > 0$ such that JF satisfies $\|JF(z) - JF(z^{'})\| \leq L\|z - z^{'}\|$ for all $z, z^{'} \in \Delta$.*

Assumption 3.2. *The level set $\mathcal{L} := \{z \mid F(z) \preccurlyeq_Q F(z_1)\}$ is bounded.*

Note that throughout this section, we assume that $0 < \rho < \sigma < 1$, and $e \in Q$ as defined in (2.13).

We propose the following variants of the PRP and LS techniques:

$$\beta_k^{YPR} := \frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\mu_2|\zeta(z_k, d_{k-1})| - \zeta(z_{k-1}, u(z_{k-1}))}, \tag{3.16}$$

$$\beta_k^{YLS} := \frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\mu_2|\zeta(z_k, d_{k-1})| - \zeta(z_{k-1}, d_{k-1})}, \tag{3.17}$$

and

$$\beta_k^{YHS} := \frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, d_{k-1}) + \mu_2|\zeta(z_k, d_{k-1})|}, \tag{3.18}$$

where $\mu_1 \in (0, 1)$ and $\mu_2 \in (\mu_1, \infty)$

Remark 3.2. The proposed techniques (3.16) and (3.18) are considered nonnegative, that is $\max\{\beta_k^{YPR}, 0\}$, $\max\{\beta_k^{YLS}, 0\}$, and $\max\{\beta_k^{YHS}, 0\}$. Additionally, the choice of μ_1 and μ_2 are such that $\frac{\mu_1}{\mu_2} \in (0, 1)$.

Let us consider the following general algorithm for VOPs.

Algorithm 1: Conjugate **G**radient **A**lgorithm (CG Algorithm)

Step 0: Given $z_1 \in \mathbb{R}^n$ and Initialization $k \longleftarrow 1$.
Step 1: Compute $u(z_k)$ and $v(z_k)$ using (2.8) and (2.9), respectively.
Step 2: If $v(z_k) = 0$, then stop. Otherwise, compute

$$d_k = \begin{cases} u(z_k), & k = 1, \\ u(z_k) + \beta_k d_{k-1}, & k \geq 2, \end{cases} \tag{3.19}$$

where β_k is a nonnegative parameter.
Step 3: Compute $t_k > 0$ by using the line search (2.15).
Step 4: Set $z_{k+1} = z_k + t_k d_k$, for $k \longleftarrow k + 1$ and go back to **Step 1.**

Let us now proposed the well-known property $(*)$ as follows:
Property $(*)$ [43] Consider Algorithm 1 and suppose that

$$0 < \bar{\delta} \leq \|u(z_k)\|, \tag{3.20}$$

for all $k \geq 1$. Under this assumption, we get a property $(*)$ if there are some constants $q > 1$ and $\lambda > 0$ for all k such that $|\beta_k| \leq q$, and $\|s_{k-1}\| \leq \lambda$ implies that $|\beta_k| \leq \frac{1}{2q}$, where $s_{k-1} = z_k - z_{k-1}$. The well-known property $(*)$ was originally introduced by Gilbert and Nocedal [25] to analyze the global convergence of PRP and HS in scalar optimization, its vector extension was subsequently provided by Lucambio Pérez and Prudente [43].

The theorem below suggests that using standard assumptions, a CG technique in vector setting which satisfies property $(*)$, converges.

Theorem 3.3. [43] *Consider Algorithm 1 and let Assumptions 3.1 and 3.2 hold, for all k, where:*
(a) β_k is nonnegative;
(b) d_k is a $Q-$descent direction of F at z_k;
(c) t_k satisfies condition (2.15)*;*
(d) property $(*)$ holds. Then,

$$\liminf_{k\to\infty} \|u(z_k)\| = 0.$$

Note that, by Assumption 3.2 we have that $\{z_k\} \subset \mathcal{L}$, there exists $\bar{M} > 0$ such that

$$\|z_k\| \leq \bar{M}, \tag{3.21}$$

for all k. Thus, by the continuity arguments, there are constants $\delta > 0$ and $\gamma > 0$ such that

$$\|u(z_k)\| \leq \delta \text{ and } \|JF(z_k)\| \leq \gamma \tag{3.22}$$

hold. In addition, for all $\bar{p} \in C$, we get

$$0 < -\zeta(z_k, u(z_k)) \leq \langle JF(z_k)u(z_k), \bar{p}\rangle \leq \|JF(z_k)\|\|u(z_k)\| \leq \delta\gamma, \tag{3.23}$$

with $\|\bar{p}\| = 1$.

The following result shows that d_k defined as (3.19) satisfies the SDC (2.11) without any line search.

Lemma 3.3. *Consider Algorithm 1 with β_k defined as* (3.16)*. Then, d_k defined by* (3.19)*, satisfies the SDC* (2.11) *with $c = \left(1 - \frac{\mu_1}{\mu_2}\right)$, for all $k \geq 1$, where μ_1 and μ_2 are defined in* (3.18)*.*

Proof. The proof utilizes an induction technique. We initiate the process by considering the case when $k = 1$, we have $d_1 = u(z_1)$. Since $\zeta(z_1, u(z_1)) < 0$, from Lemma 2.2 (b), we now have

$$\zeta(z_1, d_1) \leq \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_1, u(z_1)).$$

Now, assume that up to some $k \geq 2$, we have

$$\zeta(z_{k-1}, d_{k-1}) \leq \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_{k-1}, u(z_{k-1})) < 0. \tag{3.24}$$

Observe from Remark 3.2, we have $\beta_k^{YPR} \geq 0$ and so it is well-defined.

Additionally, using Lemma 2.1 (a), (3.19) and the fact that $\beta_k^{YPR} \geq 0$, we have

$$\zeta(z_k, d_k) \leq \zeta(z_k, u(z_k)) + \beta_k^{YPR}\zeta(z_k, d_{k-1}). \tag{3.25}$$

If $\zeta(z_k, d_{k-1}) \leq 0$, the result follows trivially with $c = \left(1 - \frac{\mu_1}{\mu_2}\right)$.

Otherwise, $\zeta(z_k, d_{k-1}) > 0$, and so by using (3.16) in (3.25), we have

$$\zeta(z_k, d_k) \le \zeta(z_k, u(z_k)) + \left(\frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\mu_2\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, u(z_{k-1}))}\right)\zeta(z_k, d_{k-1}). \tag{3.26}$$

Observe from Lemma 2.2(b) that $-\zeta(z_{k-1}, u(z_{k-1})) > 0$. This implies that

$$\mu_2\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, u(z_{k-1})) > 0, \tag{3.27}$$

for all k.

Thus,

$$\zeta(z_k, d_k) \le \zeta(z_k, u(z_k)) + \frac{-\mu_1\zeta(z_k, u(z_k))\zeta(z_k, d_{k-1})}{\mu_2\zeta(z_k, d_{k-1})}. \tag{3.28}$$

Inequality (3.28) arises from omitting the last term in (3.26). Additionally, the positive term $-\zeta(z_{k-1}, u(z_{k-1})) > 0$ in the denominator of the second term in (3.26) is dropped. This further yields

$$\zeta(z_k, d_k) \le \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_k, u(z_k)). \tag{3.29}$$

The proof is complete. □

Theorem 3.4. *Consider Algorithm 1 with* β_k *defined as* (3.16) *such that Assumptions* (3.1) *and* (3.2) *hold. If* t_k *satisfies condition* (2.15)*. Then,*

$$\liminf_{k\to\infty} \|u(z_k)\| = 0. \tag{3.30}$$

Proof. To that end, using Theorem 3.3, it suffices to show that Algorithm 1 with β_k in (3.16) has property$(*)$. Now, assume that (3.20) holds. Then, by (3.22) and (3.20), we have

$$0 < \bar{\delta} \le \|u(z_k)\| \le \delta, \ \forall\, k \ge 1. \tag{3.31}$$

Additionally, by (3.23), Lemma 2.2 (b), and (3.31), we have

$$\frac{\bar{\delta}^2}{2} \le -\zeta(z_k, u(z_k)) \le \delta\gamma. \tag{3.32}$$

We also see from (3.22) and (3.23) that

$$|\zeta(z_{k-1}, u(z_k))| = |\langle JF(z_{k-1})u(z_k), \bar{p}\rangle| \le ||JF(z_{k-1})||||u(z_k)|| \le \delta\gamma, \tag{3.33}$$

then by Assumption 3.1, Lemma 2.1(d), and (3.31), we get

$$|-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))| \le L\|z_k - z_{k-1}\|\|u(z_k)\| \le L\lambda\delta, \tag{3.34}$$

where $\lambda \ge \|s_{k-1}\| = \|z_k - z_{k-1}\|$.

Now, for (3.16), we define $q := \frac{4\delta\gamma}{\bar{\delta}^2}$ and $\lambda := \frac{\bar{\delta}^2}{4L\delta q}$. Observe that for $\beta_k = 0$, nothing to show, therefore, our concern is only for the case when $\beta_k = \beta_k^{YPR} > 0$. Since $\mu_1 \in (0, 1)$ and
$-\zeta(z_k, u(z_k)) > 0$, from Lemma 2.2 (b), we have

$$\beta_k = \beta_k^{YPR} = \frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\mu_2|\zeta(z_k, d_{k-1})| - \zeta(z_{k-1}, u(z_{k-1}))} \le \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{\mu_2|\zeta(z_k, d_{k-1})| - \zeta(z_{k-1}, u(z_{k-1}))}.$$

Again, since $\mu_2|\zeta(z_k, d_{k-1})| > 0$ for all $\mu_2 > 0$, this implies that

$$\beta_k \le \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, u(z_{k-1}))}. \tag{3.35}$$

By (3.32) and (3.33), we get

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) + |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{4\delta\gamma}{\bar{\delta}^2} = q.$$

Observe that

$$-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))| \leq -\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k)). \tag{3.36}$$

Therefore, from (3.35), we have

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-\zeta(z_{k-1}, u(z_{k-1}))}.$$

Again, by (3.32) and (3.34), we get

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{|-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{2L\lambda\delta}{\bar{\delta}^2} = \frac{1}{2q}, \tag{3.37}$$

where $\lambda \geq \|s_{k-1}\| = \|z_k - z_{k-1}\|$. Hence, we conclude that Algorithm 1 with β_k in (3.16) has Property $(*)$. □

The following result shows that d_k defined as (3.19) satisfies the SDC (2.11) without any line search.

Lemma 3.4. *Consider Algorithm 1 with β_k in* (3.17). *Then d_k defined by* (3.19),*satisfies the SDC* (2.11) *with $c = \left(1 - \frac{\mu_1}{\mu_2}\right)$, for all $k \geq 1$, where μ_1 and μ_2 are defined in* (3.18).

Proof. The proof utilizes an induction technique. We initiate the process by considering the case when $k = 1$, we have $d_1 = u(z_1)$. Since $\zeta(z_1, u(z_1)) < 0$, from Lemma 2.2 (b), we now have

$$\zeta(z_1, d_1) \leq \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_1, u(z_1)).$$

Now, assume up to some $k \geq 2$ that, we have

$$\zeta(z_{k-1}, d_{k-1}) \leq \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_{k-1}, u(z_{k-1})) < 0. \tag{3.38}$$

Additionally, using Lemma 2.1 (a), (3.19) and the fact that $\beta_k = \beta_k^{YLS} \geq 0$, we have

$$\zeta(z_k, d_k) \leq \zeta(z_k, u(z_k)) + \beta_k^{YLS}\zeta(z_k, d_{k-1}). \tag{3.39}$$

If $\zeta(z_k, d_{k-1}) \leq 0$, the result follows trivially with $c = \left(1 - \frac{\mu_1}{\mu_2}\right)$.

Otherwise, $\zeta(z_k, d_{k-1}) > 0$, and by using (3.17) in (3.39), we have

$$\zeta(z_k, d_k) \leq \zeta(z_k, u(z_k)) + \left(\frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1}) + \mu_2|\zeta(z_k, d_{k-1})|}\right)\zeta(z_k, d_{k-1}). \tag{3.40}$$

Observe that from (3.38), we have $-\zeta(z_{k-1}, d_{k-1}) + \mu_2|\zeta(z_k, d_{k-1})| > 0$, for all μ_2. Therefore, we drop the last term in inequality (3.40) and get

$$\zeta(z_k, d_k) \leq \zeta(z_k, u(z_k)) + \left(\frac{-\mu_1\zeta(z_k, u(z_k))}{-\zeta(z_{k-1}, d_{k-1}) + \mu_2|\zeta(z_k, d_{k-1})|}\right)\zeta(z_k, d_{k-1}). \tag{3.41}$$

Again, since the term $-\zeta(z_{k-1}, d_{k-1}) > 0$, we get

$$\zeta(z_k, d_k) \leq \left(1 - \frac{\mu_1}{\mu_2}\right)\zeta(z_k, u(z_k)). \tag{3.42}$$

This complete the proof. □

Theorem 3.5. *Consider Algorithm 1 with β_k defined as* (3.17) *such that Assumptions* (3.1) *and* (3.2) *hold. If t_k satisfies condition* (2.15). *Then,*

$$\liminf_{k\to\infty} \|u(z_k)\| = 0. \tag{3.43}$$

Proof. The proof follows the same pattern with that of Theorem 3.4. We utilize Algorithm 1 with β_k in (3.17), we want to show that this method has property$(*)$. Now, assume that (3.20) holds and define $q := \frac{4\delta\gamma}{\bar{\delta}^2 c}$ and $\lambda := \frac{\bar{\delta}^2 c}{4L\delta q}$. Observe that for $\beta_k = 0$, nothing to show, therefore, our concern is only for the case when $\beta_k = \beta_k^{YLS} > 0$. Since $\mu_1 \in (0,1)$ and $-\zeta(z_k, u(z_k)) > 0$, from Lemma 2.2 (b), we have

$$\beta_k = \beta_k^{YLS} = \frac{-\mu_1\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1}) + |\mu_2\zeta(z_k, d_{k-1})|} \leq \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1}) + |\mu_2\zeta(z_k, d_{k-1})|}.$$

Again, since $\mu_2\zeta(z_k, d_{k-1}) > 0$ for all $\mu_2 > 0$, this implies that

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1})}. \tag{3.44}$$

This is further given as

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) + |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1})}.$$

By Lemma 3.4, we have

$$-\zeta(z_{k-1}, d_{k-1}) \geq -c\zeta(z_{k-1}, u(z_{k-1})) > 0. \tag{3.45}$$

Applying (3.45), (3.32), and (3.33), we get

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) + |\zeta(z_{k-1}, u(z_k))|}{-c\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{4\delta\gamma}{\bar{\delta}^2 c} = q.$$

Thus, from (3.44), we get

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) - |\zeta(z_{k-1}, u(z_k))|}{-\zeta(z_{k-1}, d_{k-1})} \leq \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-c\zeta(z_{k-1}, u(z_{k-1}))}.$$

We further have

$$\beta_k \leq \frac{-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))}{-c\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{|-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))|}{-c\zeta(z_{k-1}, u(z_{k-1}))} \tag{3.46}$$

Again, we see from (3.32) and (3.34) that

$$\beta_k \leq \frac{|-\zeta(z_k, u(z_k)) + \zeta(z_{k-1}, u(z_k))|}{-c\zeta(z_{k-1}, u(z_{k-1}))} \leq \frac{2L\lambda\delta}{\bar{\delta}^2 c} = \frac{1}{2q}.$$

Hence, we conclude that Algorithm 1 with β_k in (3.17) has Property $(*)$. □

The following result shows that d_k defined as (3.19) satisfies the SDC (2.11) using WWC.

Lemma 3.5. *Consider Algorithm 1 with β_k in* (3.18) *and suppose* (2.14) *holds. Then d_k defined by* (3.19)*,satisfies the SDC* (2.11) *with $c = \left(1 - \frac{\mu_1}{\mu_2}\right)$, for all $k \geq 1$, where μ_1 and μ_2 are defined in* (3.17).

Proof. Using the WWC (2.14), we have $\zeta(z_k, d_{k-1}) - \zeta(z_{k-1}, d_{k-1}) > 0$ for all $k \geq 1$. Then, the proof follows a similar pattern to that of Lemma 3.4. □

Theorem 3.6. *Consider Algorithm 1 with β_k defined as* (3.18) *such that Assumptions* (3.1) *and* (3.2) *hold. If t_k satisfies condition* (2.15)*. Then,*

$$\liminf_{k\to\infty} \|u(z_k)\| = 0. \tag{3.47}$$

Proof. The proof follows a similar pattern to that of Theorem 3.5. □

4. Numerical experiments and discussions

In this section, we report the performance of the proposed techniques, namely: YPR, YLS, and YHS. The purpose is to assess their effectiveness and robustness in solving benchmark test problems derived from a wide range of multiobjective optimization research articles in the literature. All the Algorithms were implemented using double-precision Fortran 90, and the experiments were conducted on a PC with the following specifications: Intel Core i5-1135G7 CPU running at 2.4GHz, and 16 GB of RAM.

For multiobjective optimization, we take e to be $[1, \cdots, 1]^T \in \mathbb{R}^m$, Q and C are considered as $\mathbb{R}^m_+$, and canonical basis of $\mathbb{R}^m$, respectively.

Below, we present a summary of the CG techniques under consideration. This encompasses both our proposed techniques and those employed for comparison purposes:

- YPR+: a nonnegative variant of the PRP CG technique defined by (3.16);
- YLS+: a nonnegative variant of the LS CG technique defined by (3.17);
- YHS+: a nonnegative variant of the HS CG technique defined by (3.18);
- HS+: a nonnegative Hestenes-Stiefel (HS) CG technique proposed in [43].

An essential part of these techniques is the computation of the steepest descent direction, denoted as $u(z)$. To achieve this, we compute problem (2.10) by using Algencan, a versatile augmented Lagrangian code designed for solving nonlinear problems [5]. In addition, the selection of the step size was performed using a line search strategy that satisfies condition (2.15). Specifically, we employed the line search used for the HS+ technique in [43] for all the proposed techniques. Below are the initial parameters utilized in the implementation of the proposed techniques:

- $\rho = 10^{-4}, \quad \sigma = 0.1 \quad c = 0.4.$

Moreover, it was conveyed in Lemma 2.2 that $z \in \mathbb{R}^n$ is a $Q-$critical point of F if and only if $v(z) = 0$. Consequently, the experimentation was conducted by running all the implemented techniques up to the point of convergence, which is assumed to be $v(z) \geq -5 \times eps^{\frac{1}{2}}$, in which the $v(z)$ is defined by (2.9) and the machine precision, $eps \approx 2.22 \times 10^{-16}$ or whenever, the maximum number of iterations, $\#maxIt = 5000$ is exceeded. Additionally, we consider $\mu_1 = 0.01$ and $\mu_2 = 0.1$ in our numerical computation.

Table 1 presents essential information regarding the selected test problems. In the first column, we have the names of the problems; for instance, "MGH" corresponds to the problem introduced by Moré, Garbow, and Hillstrom in [46], and "SLC2" aligns with the second problem proposed by Schütze, Lara, and Coello in [51]. All other problems follow the same pattern with their corresponding sources. The second and third columns, labeled as "n" and "m," respectively, indicate the numbers of variables under consideration and the objective functions of the problem, respectively. To generate the starting points, a box constraint was utilized, defined as$\{z \in \mathbb{R}^n \mid \bar{l} \leq z \leq \bar{u}\}$, with the lower and upper bounds denoted in the fourth and fifth columns, respectively. While the last column indicates the corresponding references of the problems.

Tables 2 and 3 present the results of our new CG techniques in comparison with HS+ and are organized as follow: "%", "It", "Fe", and "Ge". In this case, "%" denotes the percentage of runs that has reached critical point and for the successful runs, while "It", "Fe", and "Ge" indicate the medians number of iterations, functions and gradient evaluations, respectively. It is important to emphasize that each evaluation of an objective function (respectively, objective gradient) in the corresponding computation is accounted for in the Fe column (respectively, Ge column).

The following table presents the collection of test problems considered in this paper.

TABLE 1. List of Test Problems

Problems	n	m	$\bar{l}^T$	$\bar{u}^T$	Source
JOS1	1000	2	$(-10000, \cdots, -10000)$	$(10000, \cdots, 10000)$	[35]
SLC2	10	2	$(-100, \cdots, -100)$	$(100, \cdots, 100)$	[51]
	1000	2	$(-100, \cdots, -100)$	$(100, \cdots, 100)$	[51]
	2000	2	$(-100, \cdots, -100)$	$(100, \cdots, 100)$	[51]
SLCDT1	2	2	$(-5, -5)$	$(5, 5)$	[52]
AP1	2	3	$(-100, -100)$	$(100, 100)$	[2]
AP2	2	2	$(-100, -100)$	$(100, 100)$	[2]
AP4	2	2	$(-100, -100)$	$(100, 100)$	[2]
Lov1	2	2	$(-100, -100)$	$(100, 100)$	[41]
Lov4	2	2	$(-100, -100)$	$(100, 100)$	[41]
FF1	2	2	$(-1, -1)$	$(1, 1)$	[35]
	2000	2	$(-1, \cdots, -1)$	$(1, \cdots, 1)$	[35]
	10000	2	$(-1, \cdots, -1)$	$(1, \cdots, 1)$	[35]
FDS	100	3	$(-2, \cdots, -2)$	$(2, \cdots, 2)$	[18]
MMR1	2	2	$(0, 0)$	$(1, 1)$	[45]
MMR5	2	2	$(-5, -5)$	$(5, 5)$	[45]
	1000	2	$(-5, \cdots, -5)$	$(5, \cdots, 5)$	[45]
MOP1	2	2	$(-100000, -100000)$	$(100000, 100000)$	[35]
MOP5	2	3	$(-1, -1)$	$(1, 1)$	[35]
DGO1	2	2	$(-10, -10)$	$(13, 13)$	[35]
Far1	2	2	$(-1, -1)$	$(1, 1)$	[35]
MLF2	2	2	$(-100, -100)$	$(100, 100)$	[35]
SSFYY2	2	2	$(-100, -100)$	$(100, 100)$	[35]
SK1	2	2	$(-100, -100)$	$(100, 100)$	[35]
SK2	4	2	$(-10, -10, -10, -10)$	$(10, 10, 10, 10)$	[35]
Hil1	2	2	$(0, 0)$	$(1, 1)$	[32]
DD1	5	2	$(-20, -20, -20, -20, -20)$	$(20, 20, 20, 20, 20)$	[35]
KW2	2	2	$(-3, -3)$	$(3, 3)$	[38]
Toi4	4	2	$(-100, -100, -100, -100)$	$(100, 100, 100, 100)$	[54]
Toi8	2	2	$(-1, -1, -1, -1)$	$(1,1,1,1)$	[54]
MGH26	4	4	$(-1, -1, -1, -1)$	$(1,1,1,1)$	[46]
MGH33	10	10	$(-1, \cdots, -1)$	$(1, \cdots, 1)$	[46]
PNR	2	2	$(-1, -1)$	$(1, 1)$	[49]
SLCDT2	10	3	$(-100, \cdots, -100)$	$(100, \cdots, 100)$	[52]

In the following table, we present the performance of the proposed CG techniques in comparison with HS+ CG technique on a collection of some selected convex and nonconvex multiobjective problems.

TABLE 2. Performance of the proposed techniques in comparison with HS+

	HS+				YPR+			
Problem	%	It	Fe	Ge	%	It	Fe	Ge
JOS1	100	1	2	4	100	1	2	4
SLC2(n=10)	100	14.5	124.5	114	100	21.5	141	121
SLC2 (n=1000)	100	31	208.5	186	100	29	196	172.5
SLC2 (n=2000)	100	31	221	203	100	32.5	220	189.5
SLCDT1	100	2	22	24	100	2	22.5	22
AP1	100	11	104	85	100	11	104	85
AP2	100	1	2	4	100	1	2	4
AP4	100	18	140	131	100	21	163	147.5
Lov1	100	3	6	8	100	3	6	8
Lov4	100	1	6	10	100	2	6	8
FF1 (n=2)	100	12.5	75	64.5	100	12	76.5	66
FF1(n=2,000)	13	25	256	246	100	12	168	162.5
FF1(n=10,000)	15	18	218	209	100	13	138.5	132
FDS	100	45	326.5	294.5	100	45	326.5	294.5
MMR1	100	7	54	43	100	7	54	41
MMR5 (n=2)	100	84	525.5	458.5	100	46.5	230	214.5
MMR5 (n=1000)	100	103.5	584	545	100	53	411.5	392.5
MOP1	100	1	2	4	100	1	2	4
MOP5	100	2	18	19	100	3	24	25
DGO1	100	1	10	11	100	1	10	11
Far1	100	43.5	284	255	100	34.5	199	174
MLF2	100	41	229.5	206	100	34	192	166
SSFYY2	100	1	9	10.5	100	1	9	10.5
SK1	100	2	22	23.5	100	2	21	20
SK2	100	37.5	116.5	119	100	37.5	117.5	119
Hil1	100	11	80	69	100	10	66	56.5
DD1	100	74.5	230	232	100	74.5	225.5	227.5
KW2	100	13.5	108.5	90.5	100	11	81	70
Toi4	100	3	27	29	100	4	22	21
Toi8	100	1	6	7	100	1	6	7
MGH26	100	2	39	41	100	5	64.5	60
MGH33	100	1	22	31	100	1	22	31
PNR	100	14	70	58	100	11	52	44.5
SLCDT2	100	21	187	168	100	21	191	168

TABLE 3. Continuation of Table 2

	YLS+				YHS+			
Problem	%	It	Fe	Ge	%	It	Fe	Ge
JOS1	100	1	2	4	100	1	2	4
SLC2(n=10)	100	23	141	123	100	22	138.5	121
SLC2 (n=1000)	100	28.5	192.5	167.5	100	28	189	163
SLC2 (n=2000)	100	33	218.5	187.5	100	32	226	193.5
SLCDT1	100	2	22.5	22	100	2	22.5	22
AP1	100	11	104	85	100	11	104	85
AP2	100	1	2	4	100	1	2	4
AP4	100	21	163	147.5	100	21	163	147.5
Lov1	100	3	6	8	100	3	6	8
Lov4	100	2	6	8	100	2	6	8
FF1 (n=2)	100	12	76.5	66	100	12	76.5	66
FF1(n=2,000)	100	12	168	162	100	12.5	168	162
FF1(n=10,000)	100	13	138.5	132	100	13	138.5	132
FDS	100	45	326.5	294.5	100	45	326.5	294.5
MMR1	100	7	54	41	100	7	54	41
MMR5 (n=2)	100	46	230	214	100	46	230	214
MMR5 (n=1000)	100	53	411.5	392.5	100	53	411.5	392.5
MOP1	100	1	2	4	100	1	2	4
MOP5	100	3	24	25	100	3	24	25
DGO1	100	1	10	11	100	1	10	11
Far1	100	34.5	199	174	100	34.5	199	174
MLF2	100	34.5	192	166.5	100	34	192	166.5
SSFYY2	100	1.5	9.5	10.5	100	1.5	9.5	10.5
SK1	100	2	21	20	100	2	21	20
SK2	100	38	117.5	119	100	38	117	119
Hil1	100	10	66	56.5	100	10	66	56.5
DD1	100	74.5	225.5	227.5	100	74.5	225.5	227.5
KW2	100	11	81.5	70	100	11	81	70.5
Toi4	100	4	22	21	100	4	22	21
Toi8	100	1	6	7	100	1	6	7
MGH26	100	5	64.5	60	100	5	64.5	60
MGH33	100	1	22	31	100	1	22	31
PNR	100	11	52	44.5	100	11	52	44.5
SLCDT2	100	21	193	172	100	15	174	162

Table 4 provides a summary of the information presented in Tables 2 and 3. In Table 4, the term ”successful” indicates the number of test problems in which the CG techniques achieved a 100% success rate. On the other hand, ”Not successful” refers to the number of test problems in which the CG techniques did not reach a 100% success rate and ”Failure” represents the number of test problems that the CG techniques were unable to solve. Finally, ”Total” indicates the total number of test problems considered for these CG techniques. Notice that each problem here was solved 200 times using a uniform random distribution within a defined box. This process involved exploring the solution space by starting from various initial points. In this case, one significant advantage of our proposed techniques is that they successfully solved all the 34 selected test problems considered here, while HS+ was able to solved 32 successfully out of the 34.

TABLE 4. An overview of the results in Tables 2 and 3.

Run	HS+	YPR+	YLS+	YHS+
Successful	32	34	34	34
Not Successful	2	0	0	0
Failure	0	0	0	0
Total	34	34	34	34

In the context of multiobjective optimization, the primary focus is on approximating the Pareto frontier of the problem being considered. To obtain this approximate Pareto frontier, we adopted an approach where each of the implemented techniques was run 200 times for each problem. The techniques were initialized from uniformly distributed random points within the problem's bounds, which are specified in Table 1. The comparison metrics used include the number of iterations (It), number of function evaluations (Fe), and number of gradient evaluations (Ge). To ensure a fair and appropriate algorithmic comparison, we employed the well-known Dolan and Moré performance profile [13]. This tool allows us to summarize the experimental data presented in Tables 2 and 3.

Now, based on the reported Figures 1-3, the newly developed techniques, YHS+ and YPR+, demonstrate the best performance in terms of efficiency and robustness among all compared techniques, as indicated by the number of iterations. Following closely is the YLS+ technique. On the number of function and gradient evaluations, YHS+, YPR+, and YLS+ techniques evaluate fewer functions and gradients than HS+. These results make our proposed techniques promising, considering that we used the same line search for both HS+ and our proposed techniques, specifically the line search used for HS+ techniques in [43]. These impressive results suggest that the established sufficient descent condition plays a significant role in achieving such performance.

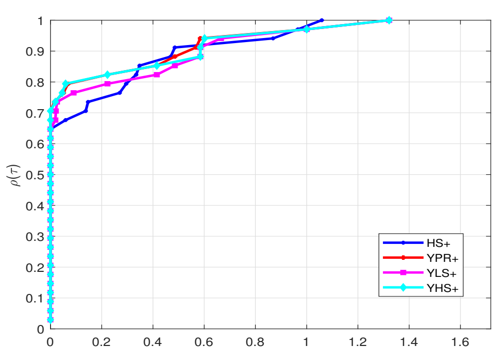

FIGURE 1. Performance on It

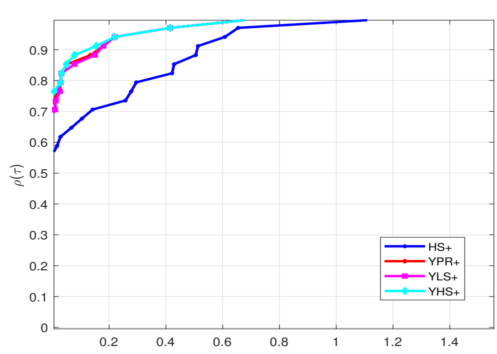

FIGURE 2. Performance on Fe

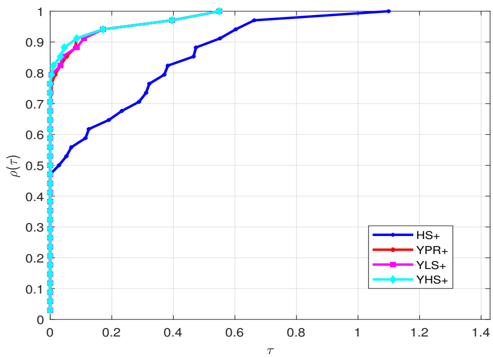

FIGURE 3. Performance on Ge

5. Concluding Remarks

In this paper, we present three new modified CG techniques for VOPs. The assumptions made in this work are built naturally upon those established for the scalar minimization case. Specifically, we proposed three CG techniques, denoted as YPR+, YLS+, and YHS+. We established the SDC of YPR+ and YLS+ techniques without line search. While YHS+ technique with line search. Moreover, without assuming convexity on the objective functions but under certain standard assumptions, we have proven the global convergence of the YPR+, YLS+ and YHS+ CG techniques. We show that the sequence generated by our proposed techniques identifies a point that satisfies the first-order necessary condition for Pareto-optimality. To our knowledge, the proposed techniques are the first PRP and HS variants shown to have SDC in the vector optimization literature. Additionally, we have conducted numerical experiments to demonstrate the practical robustness and efficiency of the proposed techniques. These experiments also include a comparison with the HS+ CG technique. The results indicate that our proposed CG techniques are competitive and promising.

Acknowledgments. The authors express their gratitude for the financial assistance received from the Center of Excellence in Theoretical and Computational Science (TaCS-CoE), King Mongkut's University of Technology Thonburi (KMUTT). Additionally, the authors acknowledge the financial support provided by the Mid-Career Research Grant (N41A640089). The first author acknowledges the support provided by the Petchra Pra Jom Klao PhD scholarship of KMUTT (Contract No. 23/2565).

Statements and Declarations

Conflict of interest: the authors declare that they have no competing interest.
Ethics approval: not applicable.
Availability of data and materials: not applicable.
Code availability: the codes are freely provided in [43]

References

[1] Al-Baali, M. Descent property and global convergence of the fletcher-reeves method with inexact line search. *IMA J. Numer. Anal.* **5** (1985), no. 1, 121–124.

[2] Ansary, M. A. T.; Panda, G. A modified quasi-newton method for vector optimization problem. *Optim.* **64** (2015) no. 11, 2289–2306.

[3] Arzuka, I.; Abu Bakar, M. R.; Leong, W. J. A scaled three-term conjugate gradient method for unconstrained optimization. *J. Inequal. Appl.* **2016** (2016), no. 1, 1–16.
[4] Bello Cruz, J. Y. A subgradient method for vector optimization problems. *SIAM J. Optim.* **23** (2013), no. 4, 2169–2182.
[5] Birgin, E.G. and Martínez, J.M. *Practical augmented Lagrangian methods for constrained optimization*. SIAM, 2014.
[6] Bonnel, H.; Iusem, A. N.; Svaiter, B. F. Proximal methods in vector optimization. *SIAM J. Optim.* **15** (2005), no. 4, 953–970.
[7] Ceng, L. C.; Yao, J. C. Approximate proximal methods in vector optimization. *Eur. J. Oper. Res.* **183** (2007), no. 1, 1–19.
[8] Chuong, T. D. Newton-like methods for efficient solutions in vector optimization. *Comput. Optim. Appl.* **54** (2013), no. 3, 495–516.
[9] Chuong, T. D.; Yao, J. C. Steepest descent methods for critical points in vector optimization problems. *Appl. Anal.*, **91** (2012), no. 10, 1811–1829.
[10] Dai, Y. H.; Yuan, Y. X. Convergence properties of the fletcher-reeves method. *IMA J. Numer. Anal.* **16** (1996), no. 2, 155–164.
[11] Dai, Y. H.; Yuan, Y. X. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.* **10** (1999), no. 1, 177–182.
[12] De, P. Ghosh, J. B.; Wells, C. E. On the minimization of completion time variance with a bicriteria extension. *Oper. Res.* **40** (1992), no. 6, 1148–1155.
[13] Dolan, E. D.; Moré, J. J. Benchmarking optimization software with performance profiles. *Math. Program.* **91** (2002), 201–213.
[14] Grana Drummond, L. M.; Iusem, A. N. A projected gradient method for vector optimization problems. *Comput. Optim. Appl.* **28** (2004), no. 1, 5–29.
[15] Grana Drummond, L. M.; Svaiter, B. F. A steepest descent method for vector optimization. *J. Comput. Appl. Math.* **175** (2005), no. 2, 395–414.
[16] Fletcher, R.; Reeves, C. M. Function minimization by conjugate gradients. *Comput. J.*, **7** (1964), no. 2, 149–154.
[17] Fletcher, R. *Practical methods of optimization: Unconstrained optimization*, **1** John Wiley and Sons, New York, 1980.
[18] Fliege, J.; Grana-Drummond, L. M.; Svaiter, B. F. Newton's method for multiobjective optimization. *SIAM J. Optim.* **20** (2009), no. 2, 602–626.
[19] Fliege, J.; Svaiter, B. F. Steepest descent methods for multicriteria optimization. *Math. Methods Oper. Res.* **51** (2000), no. 3, 479–494.
[20] Fliege, J.; Vicente, L. N. Multicriteria approach to bilevel optimization. *J. Optim. Theory Appl.*, **131** (2006), 209–225.
[21] Fliege, J.; Werner, R. Robust multiobjective optimization & applications in portfolio optimization. *Eur. J. Oper. Res.* **234** (2014), no. 2, 422–433.
[22] Fukuda, E. H.; Graña Drummond, L. M. On the convergence of the projected gradient method for vector optimization. *Optim.* **60** (2011), no. 8-9, 1009–1021.
[23] Fukuda, E. H.; Graña Drummond, L. M. A survey on multiobjective descent methods. *Pesqui. Oper.* **34** (2014), 585–620.
[24] Fukuda, E. H.; Graña Drummond, L. M. Inexact projected gradient method for vector optimization. *Comput. Optim. Appl.* **54** (2013) no. 3, 473–493.
[25] Gilbert, J.C. and Nocedal, J. Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.*, **2** (1992), no. 1, 21–42.
[26] Gonçalves, M.L.N. Lima, F.S. and Prudente, L.F. A study of liu-storey conjugate gradient methods for vector optimization. *Appl. Math. Comput.* **425** (2022), 127099.
[27] Gonçalves, M.L.N. and Prudente, L.F. On the extension of the hager–zhang conjugate gradient method for vector optimization. *Comput. Optim. Appl.* **7** (2020), no. 3, 889–916.
[28] Gravel, M. Marc Martel, J. Nadeau, R. Price, W. and Tremblay, R. A multicriterion view of optimal resource allocation in job-shop production. *Eur. J. Oper. Res.* **61** (1992), no. 1-2, 230–244.
[29] Dai, Y.H. and Liao, L.Z. New conjugacy conditions and related nonlinear conjugate gradient methods. *Appl. Math. Optim.* **43** (2001), 87–101.
[30] He, Q.R. Chen, C.R. and Li, S.J. Spectral conjugate gradient methods for vector optimization problems. *Comput. Optim. Appl.* (2023), 1–33.
[31] Hestenes, M.R. and Stiefel, E. Methods of conjugate gradients for solving. *J. Res. Natl. Bur. Stand.* **49** (1952), no. 6, 409.
[32] Hillermeier, C. Generalized homotopy approach to multiobjective optimization. *J. Optim. Theory Appl.* **110** (2001), no. 3, 557–583.

[33] Hong, T.S. Craft, D.L. Carlsson, F. and Bortfeld, T.R. Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head. *International Journal of Radiation Oncology* Biology* Physics*, **72** (2008), no. 4, 1208–1214.

[34] Huang, H. and Lin, S. A modified wei–yao–liu conjugate gradient method for unconstrained optimization. *Appl. Math. Comput.* **231** (2014), 179–186.

[35] Huband, S. Hingston, P. Barone, L. and While, L. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Trans. Evol. Comput.* **10** (2006), no. 5,(2006), 477–06.

[36] Jahn, J. Vector optimization, theory, application, and extensions, Springer, Berlin, 2004.

[37] Jahn, J. Kirsch, A. and Wagner, C. Optimization of rod antennas of mobile phones. *Math. Methods Oper. Res.* **59** (2004), 37–51.

[38] Kim, I.Y. and De Weck, O.L. Adaptive weighted-sum method for bi-objective optimization: Pareto front generation. *Struct. Multidiscipl. Optim.* **29** (2005), 149–158.

[39] Leschine, T.M. Wallenius, H. and Verdini, W.A. Interactive multiobjective analysis and assimilative capacity-based ocean disposal decisions. *Eur. J. Oper. Res.* **56** (1992), no. 2, 278–289.

[40] Liu, Y. and Storey, C. Efficient generalized conjugate gradient algorithms, part 1: theory. *J. Optim. Theory Appl.* **69** (1991), no. 1, 129–137.

[41] Lovison, A. Singular continuation: Generating piecewise linear approximations to pareto sets via global analysis. *SIAM J. Optim.* **21** (2011), no. 2, 463–490.

[42] Luc, D.T. *Theory of vector optimization*. Springer, 1989.

[43] Lucambio Pérez, L.R. and Prudente, L.F. Nonlinear conjugate gradient methods for vector optimization. *SIAM J. Optim.* **28** (2018), no. 3, 2690–2720.

[44] Lucambio Pérez, L.R. and Prudente, L.F. A wolfe line search algorithm for vector optimization. *ACM Trans. Math. Softw.* **45** (2019), no. 4, 1–23.

[45] Miglierina, E. Molho, E. and Recchioni, M.C. Box-constrained multi-objective optimization: a gradient-like method without "a priori" scalarization. *Eur. J. Oper. Res.* **188** (2008), no. 3, 662–682.

[46] Moré, J.J. Garbow, B.S. and Hillstrom, K.E. Testing unconstrained optimization software. *ACM Trans. Math. Softw.* **7** (1981), no. 1, 17–41.

[47] Polak, E. and Ribiere, G. Note sur la convergence de méthodes de directions conjuguées. *Revue française d'informatique et de recherche opérationnelle. Série rouge*, **3** (1969), no. 16, 35–43.

[48] Powell, M.J.D. Nonconvex minimization calculations and the conjugate gradient method. *Numer. Anal.* (1984), 122–141.

[49] Preuss, M. Naujoks, B. and Rudolph, G. Pareto set and emoa behavior for simple multimodal multiobjective functions. *PPSN*, (2006), 513–522.

[50] Qu, S. Goh, M. and Chan, F.T.S. Quasi-newton methods for solving multiobjective optimization. *Oper. Res. Lett.* **39** (2011), no. 5, 397–399.

[51] Schütze, O. Lara, A. and Coello, C.C. The directed search method for unconstrained multi-objective optimization problems. *Proceedings of the EVOLVE–A Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation*, (2011), 1–4.

[52] Schütze, O. Laumanns, M. Coello, C.A.C. Dellnitz, M. and Talbi, E. Convergence of stochastic search algorithms to finite size pareto set approximations. *J. Glob. Optim.* **41** (2008), 559–577.

[53] Stewart, T. Bandte, O. Braun, H. Chakraborti, N. Ehrgott, M. Göbelt, M. Jin, Y. Nakayama, H. Poles, S. and Stefano, D. D. Real-world applications of multiobjective optimization. *Multiobjective optimization: interactive and evolutionary approaches*, pages 285–327, 2008.

[54] Toint, P.L. Test problems for partially separable optimization and results for the routine pspmin. the university of namur, department of mathematics. *Technical report, Belgium, Tech. Rep.*, 1983.

[55] Wei, Z. Li, G. and Qi, L. New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems. *Appl. Math. Comput.* **179** (2006), no. 2, 407–430.

[56] Wei, Z. Yao, S. and Liu, L. The convergence properties of some new conjugate gradient methods. *Appl. Math. Comput.* **183** (2006), no. 2, 1341–1350.

[57] Yahaya, J. Arzuka, i. and Isyaku, M. Descent modified conjugate gradient methods for vector optimization problems. *Bangmod Int. J. Math. & Comp. Sci.* **9** (2023), 72–91.

[58] Yahaya, J. and Kumam, P. Efficient hybrid conjugate gradient techniques for vector optimization. *Results in Control and Optimization*, **14** (2024), 100348.

[59] Yuan G., Zhao, Y. and Wei, Z. A descent nonlinear conjugate gradient method for large-scale unconstrained optimization. *Appl. Math. Comput.* **187** (2007), no. 2, 636–643.

[60] Yuan G., Lu, X. and Wei, Z. A conjugate gradient method with descent direction for unconstrained optimization. *J. Comput. Appl. Math.* **233** (2009), no. 2, 519–530.

[61] Zhang, L. An improved wei–yao–liu nonlinear conjugate gradient method for optimization computation. *Appl. Math. Comput.* **215** (2009), no. 6, 2269–2274.

[1]CENTER OF EXCELLENCE IN THEORETICAL AND COMPUTATIONAL SCIENCE (TACS-COE)
AND KMUTTFIXED POINT RESEARCH LABORATORY
ROOM SCL 802 FIXED POINT LABORATORY SCIENCE LABORATORY BUILDING
DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY THONBURI (KMUTT)
126 PRACHA-UTHIT ROAD, BANG MOD, THUNG KHRU, BANGKOK 10140, THAILAND.
Email address: poom.kum@kmutt.ac.th

[2]NCAO RESEARCH CENTER, FIXED POINT THEORY AND APPLICATIONS RESEARCH GROUP
CENTER OF EXCELLENCE IN THEORETICAL AND COMPUTATIONAL SCIENCE (TACSCOE)
FACULTY OF SCIENCE, KING MONGKUT'S UNIVERSITY OF TECHNOLOGY THONBURI (KMUTT)
126 PRACHA-UTHIT ROAD, BANG MOD, THUNG KHRU, BANGKOK 10140, THAILAND.

[3]DEPARTMENT OF MATHEMATICS
FACULTY OF PHYSICAL SCIENCES
AHMADU BELLO UNIVERSITY ZARIA
KADUNA STATE, NIGERIA.
Email address: yahayaj@abu.edu.ng

[4]DEPARTMENT OF MATHEMATICS
FACULTY OF PHYSICAL AND COMPUTING SCIENCES
USMANU DANFODIO UNIVERSITY, SOKOTO 840004, NIGERIA
Email address: abubakar.jamilu@udusok.edu.ng