Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2026, 14(4), 1587-1607
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Note on the Backward-Douglas-Rachford Splitting Method for Generalized DC Programming

Xinyi Zhu, Zidan Wang*

Key Laboratory of Optimization Theory and Applications, School of Mathematical Sciences, China West Normal University, Nanchong, China
Email: zxy.scnc@qq.com, *1265130810@qq.com

How to cite this paper: Zhu, X.Y. and Wang, Z.D. (2026) A Note on the Backward-Douglas-Rachford Splitting Method for Generalized DC Programming. *Journal of Applied Mathematics and Physics*, **14**, 1587-1607.
https://doi.org/10.4236/jamp.2026.144074

Received: March 31, 2026
Accepted: April 20, 2026
Published: April 23, 2026

Abstract

We revisit the Backward-Douglas-Rachford algorithm proposed by Pham *et al.* for solving generalized DC programming problems composed of differentiable, lower semicontinuous, and convex functions. By employing the convergence analysis framework of Themelis *et al.*, based on the lower bound of smooth functions, and under the same standard assumptions and conditions, we recharacterize the range of the step size. Compared with the result of Pham *et al.*, we obtain a larger step size range and simplify its expression. Moreover, under this step size condition, we prove that the sequence generated by the algorithm possesses subsequential convergence.

Keywords

Backward-Douglas-Rachford Splitting Method, DC Programming, Nonconvex Optimization, Step Size Result

1. Introduction

We consider the following problem

$$\min_{x\in\mathbb{R}^n} F(x) := f(x) + h(x) - g(x), \tag{1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ is a differentiable function with a Lipschitz continuous gradient, $h:\mathbb{R}^n\to(-\infty,+\infty]$ is a lower semicontinuous function, and $g:\mathbb{R}^n\to\mathbb{R}$ is a continuously convex function. When all functions in (1) are convex, the problem is referred to as a generalized difference-of-convex (DC) programming [1]. In particular, when $f=0$, problem (1) reduces to the standard DC programming [2].

The classical algorithm for DC programming is the difference-of-convex algorithm [3] and its variants [4]. The alternating direction method of multipliers has also been employed to address DC programming [5]. The Douglas-Rachford splitting method (DRSM) constitutes another powerful approach for DC programming. In particular, Chuang *et al.* [1] proposed a unified DR splitting framework for solving generalized DC programming problems of the form (1), but its convergence analysis relies on the strict assumption of strong convexity and fails to cover the applicability of the classical DRSM. More recently, Pham *et al.* [6] introduced a Backward Douglas-Rachford splitting method (BDRSM), which studies problem (1) under weaker assumptions than those imposed in earlier works. The algorithm only requires that g in problem (1) be convex, without assuming its differentiability, and does not require the convexity of f or h, thus having a broader range of applicability.

DRSM was originally proposed by Douglas and Rachford [7] in 1956 to compute numerical solutions of the heat differential equation. After its extension to the monotone operator setting by Lions and Mercier [8], the DRSM has been extensively studied in convex optimization.

In recent years, the application of the DRSM to the following nonconvex optimization problem

$$\min_{x} f(x)+g(x), \tag{2}$$

has attracted considerable attention; see, for example [9]-[13]. Here, $f:\mathbb{R}^n \to \mathbb{R}$ is a differentiable function and $g:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ a proper lower semicontinuous function. By introducing the Douglas-Rachford envelope (DRE) function and utilizing the Kurdyka–Łojasiewicz (KL) inequality [14] [15], Li and Pong [9] established the full sequence convergence for the DRSM in a nonconvex setting for (2). Full sequence convergence is achieved as long as the following condition holds

$$\gamma < \frac{-(2.5l+2L)+\sqrt{(2.5l+2L)^2+2L^2}}{2L^2},$$

where L is the Lipschitz constant, $l \in \mathbb{R}$ is such that $f+\frac{l}{2}\|\cdot\|^2$ is convex. More recently, Themelis *et al.* [10] conducted a refined and compact analysis of the sufficient descent property of DRSM using lower bounds for smooth functions, extending the range of the relaxation parameter to (0, 2). Within this parameter range, they obtained a broader step size condition $\gamma < \frac{1}{L}$ compared to that of Li and Pong. Meanwhile, they unified the convergence analysis framework for the alternating direction method of multipliers and DRSM. For the generalized DC programming (1), Pham *et al.* [6] proposed BDRSM. The iteration of BDRSM is as follows

Algorithm 1. Backward-Douglas-Rachford splitting method (BDRSM).

Step 1. Choose initial points $y_0, z_0, w_0 \in \mathbb{R}^n$ and set $n = 0$. Let $\gamma > 0, \tau \geq 0$, and $\nu \in (0,2)$.

Step 2. Compute

$$x_{n+1} \in \arg\min_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{1}{2\gamma} \|x - y_n\|^2 \right\},$$

$$w_{n+1} = \arg\min_{w \in \mathbb{R}^n} \left\{ g^*(w) - \langle w, z_n \rangle + \frac{\tau}{2} \|w - w_n\|^2 \right\},$$

$$z_{n+1} \in \arg\min_{z \in \mathbb{R}^n} \left\{ h(z) + \frac{1}{2\gamma} \left\| z - (2x_{n+1} - y_n + \gamma w_{n+1}) \right\|^2 \right\},$$

$$y_{n+1} = y_n + \nu (z_{n+1} - x_{n+1}).$$

Step 3. If a termination criterion does not hold, set $n = n+1$ and go to Step 2.

Note that, when $g = 0$, BDRSM reduces to the relaxed DRSM; If we further set $\nu = 1$, it coincides with the classical DRSM. Compared with existing works [1] [16], the convergence analysis in [6] relies on weaker assumptions, it only requires g to be convex. Under mild conditions, Pham *et al.* [6] established the global convergence of the full sequence of iterates and derived corresponding convergence rate results, when γ satisfies the following inequality

$$\gamma < \frac{-\nu\rho_f + \sqrt{\nu^2 \rho_f^2 + 8(2-\nu) L_f^2}}{4L_f^2}, \tag{3}$$

where L_f is the gradient Lipschitz constant, ρ_f is the weak convexity constant. In view of this, motivated by the work of [9] [10], under the same original assumptions, we utilize the lower bounds of smooth functions to recharacterize the parameter selection range for the iterative step size, and obtain a step size that has a larger range and a more concise expression compared to (3).

The remainder of this paper is structured as follows. In Section 2, we present the necessary preliminaries required throughout this section. In Section 3, through a more refined analysis utilizing lower bounds for smooth functions, we obtain a step size that has a larger range and a more concise expression. Based on this, by constructing a Lyapunov function, we proved the subsequential convergence. In Section 4, the research work and results of this paper are systematically summarized, and possible future research directions are discussed.

Remark 1 (Existence of solutions to subproblems). According to the setup of Pham *et al.* [6], the three minimization subproblems in the above iteration admit solutions under the following conditions. Since f is L_f-smooth and $\frac{1}{2\gamma}\|x - y_n\|^2$ is strongly convex, the x-subproblem admits a unique solution. Since g is a continuous convex function, its conjugate g^* is convex and lower semicontinuous; together with the term $\frac{\tau}{2}\|w - w_n\|^2$, the w-subproblem is strongly convex and coercive, hence admits a unique solution. For the z-subproblem, h is lower semicontinuous and assumed to be prox-friendly (*i.e.*, its proximal operator can be computed efficiently), so that the subproblem is solvable.

2. Preliminaries

In this section, we introduce the basic concepts and several important lemmas and theorems. First, the meanings of some special symbols commonly used in the text are provided.

Let $\mathbb{R}_+$ denote the set of nonnegative real numbers and $\mathbb{R}_{++}$ the set of positive real numbers. Let $\mathbb{R}^n$ denote the n-dimensional Euclidean space, equipped with the inner product $\langle \cdot,\cdot \rangle$, and the induced Euclidean norm $\|\cdot\|$. Consider a function $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$. The domain of f is defined as $\operatorname{dom}(f) := \{x\in\mathbb{R}^n : f(x) < +\infty\}$. The function f is called proper if $\operatorname{dom}(f) \neq \varnothing$ and it does not take the value $-\infty$. It is said to be coercive if $f(x) \to +\infty$ as $\|x\| \to +\infty$. The epigraph of f is defined by $\operatorname{epi}(f) := \{(x,\rho)\in\mathbb{R}^n\times\mathbb{R} : f(x)\le\rho\}$. The function f is called lower semicontinuous if its epigraph $\operatorname{epi}(f)$ is a closed set.

Let $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be a proper function. Suppose $x\in\operatorname{dom}(f)$. The subdifferential of f at x is defined by

$$\hat{\partial} f(x) := \left\{ x^* \in \mathbb{R}^n : \liminf_{y\to x} \frac{f(y) - f(x) - \langle x^*, y-x \rangle}{\|y-x\|} \ge 0 \right\}$$

and the limiting subdifferential of f at x is defined by

$$\partial f(x) := \left\{ x^* \in \mathbb{R}^n : \exists x_n \xrightarrow{f} x, x_n^* \to x^* \text{ with } x_n^* \in \hat{\partial} f(x_n) \right\},$$

where the notation $y \xrightarrow{f} x$ means that $y \to x$ with $f(y) \to f(x)$. When $x \notin \operatorname{dom}(f)$, both the subdifferential and the limiting subdifferential of f at x are defined to be empty. It follows directly from the definition that the limiting subdifferential satisfies the robustness property

$$\partial f(x) = \left\{ x^* \in \mathbb{R}^n : \exists y_n \xrightarrow{f} x, y_n^* \to x^* \text{ with } y_n^* \in \partial f(y_n) \right\}.$$

The domain of the subdifferential ∂f is defined as

$$\operatorname{dom}\partial f := \{ x\in\mathbb{R}^n : \partial f(x) \neq \varnothing \}.$$

Let a function $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, the *Fenchel conjugate* of f is denoted by $f^*:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, is defined as

$$f^*(v) = \sup_{x\in\mathbb{R}^n} \{ \langle v, x\rangle - f(x) \}.$$

The following proposition presents key properties of the Fenchel conjugate.

Proposition 1. [6] Let $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be a proper function and let $x, v \in \mathbb{R}^n$. Then, the following assertions hold:

(i) f^* is a proper lower semicontinuous and convex function, then it holds:

$$f(x)+f^*(v) \geq \langle x, v\rangle.$$

(ii) If f is a lower semicontinuous and convex, then the following statements are equivalent:

$$v \in \partial f(x) \Leftrightarrow f(x)+f^*(v)=\langle x, v\rangle \Leftrightarrow x \in \partial f^*(v).$$

Next, we recall the definition of the Kurdyka-Łojasiewicz (KL) property, which will play a central role in our subsequent analysis.

Next, we will introduce the concepts of L_f-smooth.

Definition 1. [17] Let $f:\mathbb{R}^n \to \mathbb{R}$ be a differentiable function, if its gradient ∇f is L_f-Lipschitz continuous, there exists constant $L_f > 0$,

$$\|\nabla f(x)-\nabla f(y)\| \leq L_f\|x-y\| \text{ for all } x, y \in \mathbb{R}^n,$$

then the function f is said to be L_f-smooth.

The following descent lemma provides a useful tool for convergence analysis.

Definition 2. [17] Let $f:\mathbb{R}^n \to \mathbb{R}$ be a function. If there exists a constant $\rho_f \in [-L_f, L_f]$ such that

$$f+\frac{\rho_f}{2}\|x-y\|^2$$

is convex, then f is said to be ρ_f-hypoconvex convex.

An equivalent characterization of this property for differentiable functions is given below.

Lemma 1. [17] Let $f:\mathbb{R}^n \to \mathbb{R}$ be an L_f-smooth function, where $L_f \geq 0$, for any $x, y \in \mathbb{R}^n$, we have

$$\left|f(y)-f(x)-\langle\nabla f(x), y-x\rangle\right| \leq \frac{L_f}{2}\|y-x\|^2.$$

Theorem 2. [17] A continuously differentiable function $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is ρ_f-weakly convex if and only if for all $x, y \in \mathbb{R}^n$, the following inequality holds:

$$f(y) \geq f(x)+\langle\nabla f(x), y-x\rangle-\frac{\rho_f}{2}\|x-y\|^2.$$

For L_f-smooth and ρ_f-weakly convex function, we have the following property.

Theorem 3. [10] Let $f:\mathbb{R}^n \to \mathbb{R}$ be a L_f-smooth and ρ_f-weakly convex function. Then, for all $x, y \in \mathbb{R}^n$, it holds that

$$f(y) \geq f(x)+\langle\nabla f(x), y-x\rangle+\frac{1}{2(L_f-\rho_f)}\|\nabla f(x)-\nabla f(y)\|^2$$
$$-\frac{\rho_f L_f}{2(L_f-\rho_f)}\|x-y\|^2,$$

where $\rho_f \in [0, L_f)$.

Moreover, the above inequality is also valid if L_f is replaced with any $L \geq L_f$ and ρ_f with any $\rho \in [\rho_f, L)$.

Theorem 4. [17] Let $f: \mathbb{R}^n \to \mathbb{R}$ be a L_f -smooth and convex function. Then, for all $x, y \in \mathbb{R}^n$, it holds that

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2L_f} \|\nabla f(x) - \nabla f(y)\|^2.$$

Remark 2. (Scaling of Smoothness and Hypoconvex Constants). The following scaling properties are crucial for our subsequent analysis.

Smoothness If f is L_f -smooth, then it is also L -smooth for any $L \geq L_f$. Indeed, for any x, y, the inequality

$$\|\nabla f(x) - \nabla f(y)\| \leq L_f \|x - y\| \leq L \|x - y\|,$$

which satisfies the definition of L -smoothness, hence f is L -smooth.

Hypoconvexity If f is ρ_f -weakly convex with $\rho_f > 0$. Then it is also ρ -weakly convex for any $\rho \geq \rho_f$, To see this, note that $-\frac{\rho_f}{2}\|x - y\|^2 \geq -\frac{\rho}{2}\|x - y\|^2$. Substituting this into the inequality below yields

$$\begin{aligned} f(y) &\geq f(x) + \langle \nabla f(x), y - x \rangle - \frac{\rho_f}{2}\|x - y\|^2 \\ &\geq f(x) + \langle \nabla f(x), y - x \rangle - \frac{\rho}{2}\|x - y\|^2, \end{aligned}$$

which satisfies the definition of ρ -weakly convexity. Therefore, f is ρ -weakly convex. If f is ρ_f -strongly convex with $\rho_f < 0$, a similar argument shows it is ρ -strongly convex for any $\rho \leq \rho_f$.

In summary, if f is an L_f -smooth and ρ_f -hypoconvex function, then it is also L -smooth and ρ -hypoconvex. In the Section 3, we will frequently use these properties. Specifically, we will appropriately inflate the constants L and ρ to meet the specific conditions required for applying certain inequalities or to simplify the derivation of step-size rules.

3. Step Size Result and Convergence Analysis

This section focuses on the Backward-Douglas-Rachford splitting method (BDRSM) proposed by Pham *et al.* [6], primarily characterizing the range of the iterative step size parameter and analyzing the convergence of subsequences. Inspired by the work of Themelis *et al.* [10], we utilize the lower bounds of smooth functions to prove the sufficient descent property of the Lyapunov function and conduct a piecewise refined analysis of the descent constant, thereby obtaining a larger step size range. Under this condition, we prove the convergence of the subsequences generated by the algorithm.

Assumption 1. [6]

(i) $f: \mathbb{R}^n \to \mathbb{R}$ is differentiable ρ_f -hypoconvex function with an L_f -Lipschitz continuous gradient, where $\rho_f \in [-L_f, L_f]$.

(ii) $g:\mathbb{R}^n \to \mathbb{R}$ is a continuous and convex function.

(iii) $h:\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function.

The following lemma will be utilized in the analysis.

Lemma 5. [6] Suppose that f is differentiable with L_f-Lipschitz continuous gradient. Let $\{(x_n, y_n, z_n, w_n)\}_{n\in\mathbb{N}^*}$ be a sequence generated by **Algorithm 1**. Then, for all $n \in \mathbb{N}^*$, the following hold

(i) $y_n = x_{n+1} + \gamma \nabla f(x_{n+1})$.

(ii) $z_n \in \partial g^*(w_{n+1}) + \tau(w_{n+1} - w_n)$.

(iii) $w_{n+1} \in \partial h(z_{n+1}) - \frac{1}{\gamma}(x_{n+1} - z_{n+1}) - \frac{1}{\gamma}(x_{n+1} - y_n)$.

(iv) $\|y_{n+1} - y_n\| \le (1+\gamma L_f)\|x_{n+2} - x_{n+1}\|$.

(v) $\|z_{n+2} - z_{n+1}\| \le \left(\frac{1+\gamma L_f}{\nu}\right)\|x_{n+3} - x_{n+2}\| + \left(1 + \frac{1+\gamma L_f}{\nu}\right)\|x_{n+2} - x_{n+1}\|$.

The convergence analysis for the problem (1) is based on the following Lyapunov function [6].

$$\begin{aligned}\Phi(x,y,z,w) &:= f(x) + h(z) + g^*(w) - \langle w, z\rangle \\ &\quad + \frac{1}{2\gamma}\|x-y\|^2 - \frac{1}{2\gamma}\|y-z\|^2 + \frac{1-\nu}{\gamma}\|x-z\|^2.\end{aligned} \tag{4}$$

Let $a = x - y, b = x - z$. Using the identity $2\langle a,b\rangle = \|a\|^2 + \|b\|^2 - \|a-b\|^2$, we obtain

$$\begin{aligned}&\frac{1}{2\gamma}\|x-y\|^2 - \frac{1}{2\gamma}\|y-z\|^2 + \frac{1-\nu}{\gamma}\|x-z\|^2 - \langle w, z\rangle \\ &= \frac{1}{2\gamma}\left(\|x-y\|^2 - \|y-z\|^2 + \|x-z\|^2 + \|x-z\|^2\right) \\ &\quad - \frac{\nu}{\gamma}\|x-z\|^2 + \langle x - z - x, w\rangle \\ &= \frac{1}{\gamma}\langle x-z, x-y\rangle + \frac{1}{2\gamma}\|x-z\|^2 - \frac{\nu}{\gamma}\|x-z\|^2 + \langle x-z, w\rangle - \langle x, w\rangle \\ &= \frac{1}{\gamma}\langle x-z, x-y+\gamma w\rangle + \frac{1}{2\gamma}\|x-z\|^2 - \frac{\nu}{\gamma}\|x-z\|^2 - \langle x, w\rangle \\ &= \frac{1}{2\gamma}\|(x-z) + (x-y+\gamma w)\|^2 - \frac{1}{2\gamma}\|x-y+\gamma w\|^2 - \frac{\nu}{\gamma}\|x-z\|^2 - \langle x, w\rangle \\ &= \frac{1}{2\gamma}\|2x-y-z+\gamma w\|^2 - \frac{1}{2\gamma}\|x-y+\gamma w\|^2 - \frac{\nu}{\gamma}\|x-z\|^2 - \langle x, w\rangle.\end{aligned}$$

Therefore, (4) can also be written as

$$\begin{aligned}\Phi(x,y,z,w) &:= f(x) + h(z) + g^*(w) - \langle x, w\rangle \\ &\quad + \frac{1}{2\gamma}\|2x-y-z+\gamma w\|^2 - \frac{1}{2\gamma}\|x-y+\gamma w\|^2 - \frac{\nu}{\gamma}\|x-z\|^2.\end{aligned} \tag{5}$$

Next, we revisit the step size condition γ established by Pham *et al.*, with the aim of obtaining a step size that has a more concise expression and a larger range, while ensuring that subsequent convergence still holds under this step size condition. We first establish the sufficient descent property.

Theorem 6 (sufficient descent). Suppose that Assumption 1 hold, then stepsize γ satisfies

$$\gamma < \min\left\{\frac{1}{L_f}, \frac{2-\nu}{2\left[\rho_f\right]_+}\right\},$$

where, $\left[\rho_f\right]_+ = \max\left\{\rho_f, 0\right\}$.

Let $\left\{\left(x_n, y_n, z_n, w_n\right)\right\}_{n\in\mathbb{N}^*}$ be a sequence generated by Algorithm 1. Then the following hold:

For all $n \in \mathbb{N}^*$,

$$\Phi\left(x_n, y_n, z_n, w_n\right) - \Phi\left(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}\right) \geq \frac{\delta}{2}\left\|x_{n+1} - x_n\right\|^2 + \frac{\tau}{2}\left\|w_{n+1} - w_n\right\|^2.$$

$$\frac{\delta}{2} = \frac{2-\nu}{2\nu\gamma} + \begin{cases} \min\left\{\dfrac{-\rho_f L_f}{2\left(L_f - \rho_f\right)}, \dfrac{L_f}{2} - \dfrac{\gamma L_f^2}{\nu}\right\} & 0 < \nu < 2\left(1 - \dfrac{\rho_f}{L_f}\right), \\ -\dfrac{\rho_f}{\nu} & 2\left(1 - \dfrac{\rho_f}{L_f}\right) \leq \nu < 2. \end{cases}$$

Then the sequence $\left\{\Phi\left(x_n, y_n, z_n, w_n\right)\right\}_{n\in\mathbb{N}^{\mathbb{A}}}$ is nonincreasing.

Proof. (i) When $\nu \in (0,2)$, we choose $L \geq L_f$, $\rho \in \mathbb{R}_+$. By Theorem 3, the following inequality holds

$$\begin{aligned} & f\left(x_n\right) - f\left(x_{n+1}\right) \\ & \geq \left\langle \nabla f\left(x_{n+1}\right), x_n - x_{n+1}\right\rangle + \frac{1}{2(L-\rho)}\left\|\nabla f\left(x_{n+1}\right) - \nabla f\left(x_n\right)\right\|^2 \\ & \quad - \frac{\rho L}{2(L-\rho)}\left\|x_{n+1} - x_n\right\|^2. \end{aligned}$$

Combining the above inequality with (4), we obtain

$$\begin{aligned} & \Phi\left(x_{n+1}, y_n, z_n, w_{n+1}\right) - \Phi\left(x_n, y_n, z_n, w_{n+1}\right) \\ & = f\left(x_{n+1}\right) - f\left(x_n\right) + \frac{1}{2\gamma}\left\|x_{n+1} - y_n\right\|^2 - \frac{1}{2\gamma}\left\|x_n - y_n\right\|^2 \\ & \quad + \frac{1-\nu}{\gamma}\left\|x_{n+1} - z_n\right\|^2 - \frac{1-\nu}{\gamma}\left\|x_n - z_n\right\|^2 \\ & \leq \left\langle \nabla f\left(x_{n+1}\right), x_{n+1} - x_n\right\rangle - \frac{1}{2(L-\rho)}\left\|\nabla f\left(x_n\right) - \nabla f\left(x_{n+1}\right)\right\|^2 \\ & \quad + \frac{\rho L}{2(L-\rho)}\left\|x_n - x_{n+1}\right\|^2 + \frac{1}{2\gamma}\left\|x_{n+1} - y_n\right\|^2 - \frac{1}{2\gamma}\left\|x_n - y_n\right\|^2 \\ & \quad + \frac{1-\nu}{\gamma}\left\|x_{n+1} - z_n\right\|^2 - \frac{1-\nu}{\gamma}\left\|x_n - z_n\right\|^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\gamma}\|x_{n+1}-x_n\|^2 + \frac{1}{\gamma}\langle x_{n+1}-x_n, x_n-y_n\rangle + \langle \nabla f(x_{n+1}), x_{n+1}-x_n\rangle \\
&\quad - \frac{1}{2(L-\rho)}\|\nabla f(x_n)-\nabla f(x_{n+1})\|^2 + \frac{\rho L}{2(L-\rho)}\|x_n-x_{n+1}\|^2 \\
&\quad + \frac{1-\nu}{\gamma}\|x_{n+1}-z_n\|^2 - \frac{1-\nu}{\gamma}\|x_n-z_n\|^2 \\
&= \frac{1}{\gamma}\langle x_{n+1}-x_n, x_{n+1}-y_n\rangle - \frac{1}{2\gamma}\|x_{n+1}-x_n\|^2 + \langle \nabla f(x_{n+1}), x_{n+1}-x_n\rangle \\
&\quad - \frac{1}{2(L-\rho)}\|\nabla f(x_n)-\nabla f(x_{n+1})\|^2 + \frac{\rho L}{2(L-\rho)}\|x_n-x_{n+1}\|^2 \\
&\quad + \frac{1-\nu}{\gamma}\|x_{n+1}-z_n\|^2 - \frac{1-\nu}{\gamma}\|x_n-z_n\|^2.
\end{aligned}
$$

From Lemma 5(i), we derive the following equality

$$
0 = \nabla f(x_{n+1}) + \frac{1}{\gamma}(x_{n+1}-y_n). \tag{6}
$$

Using equality (6), we obtain that

$$
\begin{aligned}
&\Phi(x_{n+1}, y_n, z_n, w_{n+1}) - \Phi(x_n, y_n, z_n, w_{n+1}) \\
&\le -\frac{1}{2\gamma}\|x_{n+1}-x_n\|^2 - \frac{1}{2(L-\rho)}\|\nabla f(x_n)-\nabla f(x_{n+1})\|^2 \\
&\quad + \frac{\rho L}{2(L-\rho)}\|x_n-x_{n+1}\|^2 + \frac{1-\nu}{\gamma}\|x_{n+1}-z_n\|^2 - \frac{1-\nu}{\gamma}\|x_n-z_n\|^2.
\end{aligned} \tag{7}
$$

Now, from the Lyapunov function (4), we have

$$
\begin{aligned}
&\Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}) - \Phi(x_{n+1}, y_n, z_{n+1}, w_{n+1}) \\
&= \frac{1}{2\gamma}\|x_{n+1}-y_{n+1}\|^2 - \frac{1}{2\gamma}\|y_{n+1}-z_{n+1}\|^2 - \frac{1}{2\gamma}\|x_{n+1}-y_n\|^2 + \frac{1}{2\gamma}\|y_n-z_{n+1}\|^2.
\end{aligned}
$$

It is well known that

$$
\begin{aligned}
&\langle a-b, c-d\rangle \\
&= \langle a,c\rangle - \langle a,d\rangle - \langle b,c\rangle + \langle b,d\rangle \\
&= \frac{1}{2}\left(\|a\|^2+\|c\|^2-\|a-c\|^2\right) - \frac{1}{2}\left(\|a\|^2+\|d\|^2-\|a-d\|^2\right) \\
&\quad - \frac{1}{2}\left(\|b\|^2+\|c\|^2-\|b-c\|^2\right) + \frac{1}{2}\left(\|b\|^2+\|d\|^2-\|b-d\|^2\right) \\
&= \frac{1}{2}\|a-d\|^2 - \frac{1}{2}\|a-c\|^2 - \frac{1}{2}\|b-d\|^2 + \frac{1}{2}\|b-c\|^2,
\end{aligned}
$$

let $a = x_{n+1}, b = z_{n+1}, c = y_n, d = y_{n+1}$, then we have

$$
\begin{aligned}
&\Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}) - \Phi(x_{n+1}, y_n, z_{n+1}, w_{n+1}) \\
&= \frac{1}{\gamma}\langle x_{n+1}-z_{n+1}, y_n-y_{n+1}\rangle \\
&= \frac{\nu}{\gamma}\|x_{n+1}-z_{n+1}\|^2,
\end{aligned} \tag{8}
$$

where the second equality is got from the updating step of y_{n+1} in **Algorithm 1**. Next, from the definition of w_{n+1} in **Algorithm 1**, we know that

$$g^*(w_{n+1}) - \langle w_{n+1}, z_n \rangle + \frac{\tau}{2}\|w_{n+1} - w_n\|^2$$
$$\le g^*(w_n) - \langle w_n, z_n \rangle + \frac{\tau}{2}\|w_n - w_n\|^2.$$

Rearranging the above inequality, we get

$$g^*(w_{n+1}) - g^*(w_n) + \langle w_n - w_{n+1}, z_n \rangle \le -\frac{\tau}{2}\|w_{n+1} - w_n\|^2.$$

It means that

$$\begin{aligned} &\Phi(x_n, y_n, z_n, w_{n+1}) - \Phi(x_n, y_n, z_n, w_n) \\ &= g^*(w_{n+1}) - g^*(w_n) - \langle w_{n+1}, z_n \rangle + \langle w_n, z_n \rangle \\ &\le -\frac{\tau}{2}\|w_{n+1} - w_n\|^2. \end{aligned} \tag{9}$$

Subsequently, we construct the relationship of Φ regarding to z_n and z_{n+1} from (5), that is

$$\begin{aligned} &\Phi(x_{n+1}, y_n, z_{n+1}, w_{n+1}) - \Phi(x_{n+1}, y_n, z_n, w_{n+1}) \\ &= h(z_{n+1}) + \frac{1}{2\gamma}\|2x_{n+1} - y_n + \gamma w_{n+1} - z_{n+1}\|^2 - \frac{\nu}{\gamma}\|x_{n+1} - z_{n+1}\|^2 \\ &\quad - h(z_n) - \frac{1}{2\gamma}\|2x_{n+1} - y_n + \gamma w_{n+1} - z_n\|^2 + \frac{\nu}{\gamma}\|x_{n+1} - z_n\|^2. \end{aligned}$$

However, by the definition of z_{n+1} in **Algorithm 1**, it yields that

$$\begin{aligned} &\Phi(x_{n+1}, y_n, z_{n+1}, w_{n+1}) - \Phi(x_{n+1}, y_n, z_n, w_{n+1}) \\ &\le -\frac{\nu}{\gamma}\|x_{n+1} - z_{n+1}\|^2 + \frac{\nu}{\gamma}\|x_{n+1} - z_n\|^2. \end{aligned} \tag{10}$$

Combining (7), (8), (9) and (10), we obtain

$$\begin{aligned} &\Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}) - \Phi(x_n, y_n, z_n, w_n) \\ &\le -\frac{1}{2\gamma}\|x_{n+1} - x_n\|^2 - \frac{1}{2(L-\rho)}\|\nabla f(x_n) - \nabla f(x_{n+1})\|^2 \\ &\quad + \frac{\rho L}{2(L-\rho)}\|x_n - x_{n+1}\|^2 + \frac{1}{\gamma}\|x_{n+1} - z_n\|^2 \\ &\quad - \frac{1-\nu}{\gamma}\|x_n - z_n\|^2 - \frac{\tau}{2}\|w_{n+1} - w_n\|^2 \\ &= \frac{1}{2\gamma}\|x_{n+1} - x_n\|^2 - \frac{1}{2(L-\rho)}\|\nabla f(x_n) - \nabla f(x_{n+1})\|^2 \\ &\quad + \frac{\rho L}{2(L-\rho)}\|x_n - x_{n+1}\|^2 + \frac{2}{\gamma}\langle x_{n+1} - x_n, x_n - z_n \rangle \\ &\quad + \frac{\nu}{\gamma}\|x_n - z_n\|^2 - \frac{\tau}{2}\|w_{n+1} - w_n\|^2. \end{aligned} \tag{11}$$

From the iteration of y_{n+1} in **Algorithm 1** and Lemma 5(i), it is not hard to know that

$$\nu(x_n - z_n) = y_{n-1} - y_n = x_n - x_{n+1} + \gamma(\nabla f(x_n) - \nabla f(x_{n+1})).$$

Now, we can rewrite $\frac{\nu}{\gamma}\|x_n - z_n\|^2$ as follows

$$\begin{aligned}\frac{\nu}{\gamma}\|x_n-z_n\|^2&=\frac{1}{\gamma\nu}\cdot\nu^2\|x_n-z_n\|^2\\&=\frac{1}{\gamma\nu}\left\|(x_n-x_{n+1})+\gamma\left(\nabla f(x_n)-\nabla f(x_{n+1})\right)\right\|^2\\&=\frac{1}{\gamma\nu}\|x_{n+1}-x_n\|^2+\frac{\gamma}{\nu}\|\nabla f(x_{n+1})-\nabla f(x_n)\|^2\\&\quad+\frac{2}{\nu}\left\langle x_{n+1}-x_n,\nabla f(x_{n+1})-\nabla f(x_n)\right\rangle.\end{aligned}\tag{12}$$

Additionally, we aslo obtain

$$\begin{aligned}&\frac{2}{\gamma}\langle x_{n+1}-x_n,x_n-z_n\rangle\\&=\frac{2}{\gamma}\left\langle x_{n+1}-x_n,\frac{1}{\nu}(x_n-x_{n+1})+\frac{\gamma}{\nu}\left(\nabla f(x_n)-\nabla f(x_{n+1})\right)\right\rangle\\&=-\frac{2}{\nu\gamma}\|x_{n+1}-x_n\|^2+\frac{2}{\nu}\left\langle x_{n+1}-x_n,\nabla f(x_n)-\nabla f(x_{n+1})\right\rangle.\end{aligned}\tag{13}$$

Substituting (12) and (13) into (11), we have

$$\begin{aligned}&\Phi(x_{n+1},y_{n+1},z_{n+1},w_{n+1})-\Phi(x_n,y_n,z_n,w_n)\\&\le-\frac{1}{2(L-\rho)}\|\nabla f(x_n)-\nabla f(x_{n+1})\|^2+\frac{\rho L}{2(L-\rho)}\|x_n-x_{n+1}\|^2\\&\quad-\frac{\tau}{2}\|w_{n+1}-w_n\|^2+\left(\frac{1}{2\gamma}-\frac{1}{\nu\gamma}\right)\|x_{n+1}-x_n\|^2\\&\quad+\frac{\gamma}{\nu}\|\nabla f(x_{n+1})-\nabla f(x_n)\|^2\end{aligned}\tag{14}$$

$$\begin{aligned}&\le\left(\frac{1}{2\gamma}-\frac{1}{\nu\gamma}+\frac{\rho L}{2(L-\rho)}\right)\|x_{n+1}-x_n\|^2\\&\quad+\left(\frac{\gamma}{\nu}-\frac{1}{2(L-\rho)}\right)\|\nabla f(x_{n+1})-\nabla f(x_n)\|^2-\frac{\tau}{2}\|w_{n+1}-w_n\|^2,\end{aligned}\tag{15}$$

If we multiply both sides of inequality (14) by $\frac{1}{L}$, we obtain

$$\begin{aligned}&\frac{1}{L}\left(\Phi(x_n,y_n,z_n,w_n)-\Phi(x_{n+1},y_{n+1},z_{n+1},w_{n+1})\right)\\&\ge\left(-\frac{1}{2\gamma L}+\frac{1}{\nu\gamma L}-\frac{\rho}{2(L-\rho)}\right)\|x_{n+1}-x_n\|^2\\&\quad-\left(\frac{\gamma}{\nu L}-\frac{1}{2L(L-\rho)}\right)\|\nabla f(x_{n+1})-\nabla f(x_n)\|^2+\frac{\tau}{2L}\|w_{n+1}-w_n\|^2,\end{aligned}\tag{16}$$

where $\rho\in[0,L)$. Since $f(\cdot)$ is L_f-smooth, we have $\|\nabla f(x_{n+1})-\nabla f(x_n)\|\le L_f\|x_{n+1}-x_n\|$, (16) can be expressed as follows

$$\begin{aligned}&\frac{1}{L}\left(\Phi(x_n,y_n,z_n,w_n)-\Phi(x_{n+1},y_{n+1},z_{n+1},w_{n+1})\right)\\&\ge\frac{\delta}{2L}\|x_{n+1}-x_n\|^2+\frac{\tau}{2L}\|w_{n+1}-w_n\|^2,\end{aligned}$$

where δ is chosen as

$$\frac{\delta}{2L}=\begin{cases}-\frac{1}{2\gamma L}+\frac{1}{\nu\gamma L}-\frac{\frac{\rho}{L}}{2\left(1-\frac{\rho}{L}\right)} & \frac{\gamma}{\nu L}-\frac{1}{2L(L-\rho)}<0,\\ -\frac{1}{2\gamma L}+\frac{1}{\nu\gamma L}-\frac{\frac{\rho}{L}}{2\left(1-\frac{\rho}{L}\right)}-\left(\frac{\gamma}{\nu L}-\frac{1}{2L(L-\rho)}\right)L_f^2 & \text{otherwise.}\end{cases} \quad (17)$$

Now we want to select a suitable L to ensure that the constant δ is strictly positive. Therefore, we further consider two possibilities based on the size of ν:

Case 1: If $0<\nu<2\left(1-\frac{\rho}{L_f}\right)$.

The condition is equivalent to $\frac{\nu L_f}{2(L_f-\rho)}<1$, then $\rho<\frac{(2-\nu)L_f}{2}<L_f$. Since the parameter L in Theorem 3 can be any number satisfying $L\geq L_f$. To keep the analysis as simple as possible and avoid unnecessarily conservative bounds, we take $L=L_f$. Then (17) becomes

$$\frac{\delta}{2L_f}=\frac{2-\nu}{2\nu\gamma L_f}+\begin{cases}-\frac{\frac{\rho}{L_f}}{2\left(1-\frac{\rho}{L_f}\right)} & \gamma L_f<\frac{\nu L_f}{2(L_f-\rho)}\\ \frac{1}{2}-\frac{\gamma L_f}{\nu} & \text{otherwise.}\end{cases}$$

In this case, we will verify that for any γ satisfying $\gamma<\frac{1}{L_f}$, the constant δ is strictly positive.

a) If $0<\gamma L_f<\frac{\nu L_f}{2(L_f-\rho)}<1$,

$$\frac{\delta}{2L_f}=\frac{2-\nu}{2\nu\gamma L_f}-\frac{\frac{\rho}{L_f}}{2\left(1-\frac{\rho}{L_f}\right)}>\frac{2-\nu}{2\nu}-\frac{\frac{\rho}{L_f}}{2\left(1-\frac{\rho}{L_f}\right)}$$

$$>\frac{2-\nu}{2\nu}-\frac{\frac{\rho}{L_f}}{\nu}>\frac{1-\frac{\rho}{L_f}}{\nu}-\frac{1}{2}=\frac{2\left(1-\frac{\rho}{L_f}\right)-\nu}{2\nu}>0.$$

b) If $\frac{\nu L_f}{2(L_f-\rho)}\leq\gamma L_f<1$,

$$\frac{\delta}{2L_f}=\frac{2-\nu}{2\nu\gamma L_f}+\frac{1}{2}-\frac{\gamma L_f}{\nu}>\frac{2-\nu}{2\nu}-\frac{1}{\nu}+\frac{1}{2}=0.$$

According to the analysis of (a) and (b), it implies that whenever $\gamma<\frac{1}{L_f}$, δ

is strictly positive, therefore, we also obtain $\frac{\delta}{2L} > 0$.

Case 2: If $2\left(1-\frac{\rho}{L_f}\right) \le \nu < 2$.

This inequality requires ρ must satisfy $\rho > 0$, otherwise ν will an empty. To obtain a simple and unified expression for the step-size condition, we set $\rho = \rho_f$. The original condition is equivalent to $\frac{1}{L_f} \ge \frac{2-\nu}{2\rho_f}$. Since $L \ge L_f$, we can set $\frac{1}{L} = \frac{2-\nu}{2\rho_f} \le \frac{1}{L_f}$, substituting $\frac{1}{L} = \frac{2-\nu}{2\rho_f}$ into (17) and simplifying yields

$$\frac{\delta}{2} = \frac{2-\nu}{2\nu\gamma} + \begin{cases} -\frac{\rho_f}{\nu} & \gamma < \frac{2-\nu}{2\rho_f} \\ -\frac{\rho_f}{\nu} - \left(\frac{\gamma}{\nu} - \frac{2-\nu}{2\rho_f \nu}\right) \cdot L_f^2 & \text{otherwise.} \end{cases}$$

When $\gamma < \frac{2-\nu}{2\rho_f} = \frac{1}{L}$, it is easy to verify that

$$\frac{\delta}{2L} = \frac{2-\nu}{2\nu\gamma L} - \frac{\rho_f}{\nu L} > \frac{2-\nu}{2\nu} - \frac{\rho_f}{\nu L} = \frac{2-\nu}{2\nu} - \frac{\rho_f}{\nu} \cdot \frac{2-\nu}{2\rho_f} = 0,$$

we also obtain that δ is strictly positive for $\gamma < \frac{2-\nu}{2\rho_f}$. But if $\gamma \ge \frac{2-\nu}{2\rho_f}$, we can only obtain

$$\frac{\delta}{2} = \frac{2-\nu}{2\nu\gamma} - \frac{\rho_f}{\nu} + \left(-\frac{\gamma}{\nu} + \frac{2-\nu}{2\rho_f \nu}\right) \cdot L_f^2 < \frac{2-\nu}{2\nu} \cdot \frac{2\rho_f}{2-\nu} - \frac{\rho_f}{\nu} = 0.$$

Consequently, we obtain that

$$\Phi(x_n, y_n, z_n, w_n) - \Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}) \ge \frac{\delta}{2}\|x_{n+1} - x_n\|^2 + \frac{\tau}{2}\|w_{n+1} - w_n\|^2,$$

where $\frac{\tau}{2}\|w_{n+1} - w_n\|^2 > 0$ and δ is strictly positive with $\gamma < \min\left\{\frac{1}{L_f}, \frac{2-\nu}{2\rho_f}\right\}$. This shows that the sequence $\{\Phi(x_n, y_n, z_n, w_n)\}_{n\in\mathbb{N}^{\mathbb{A}}}$ is nonincreasing.

(ii) When $\nu \in (0,2)$, $\rho < 0$. The focus in the following will be on the strongly convex case. Utilizing Theorems 4, we have

$$f(x_n) - f(x_{n+1}) \ge \langle \nabla f(x_{n+1}), x_n - x_{n+1} \rangle + \frac{1}{2L_f}\|\nabla f(x_{n+1}) - \nabla f(x_n)\|^2.$$

Similar constructed to (i), we derive the following inequality

$$\begin{aligned} &\frac{1}{L_f}\left(\Phi(x_n, y_n, z_n, w_n) - \Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1})\right) \\ &\ge \left(-\frac{1}{2\gamma L_f} + \frac{1}{\nu\gamma L_f}\right)\|x_{n+1} - x_n\|^2 \\ &\quad -\left(\frac{\gamma}{\nu L_f} - \frac{1}{2L_f^2}\right)\|\nabla f(x_{n+1}) - \nabla f(x_n)\|^2 + \frac{\tau}{2L_f}\|w_{n+1} - w_n\|^2. \end{aligned}$$

δ can be expressed as

$$\frac{\delta}{2L_f}=\begin{cases}-\frac{1}{2\gamma L_f}+\frac{1}{\nu\gamma L_f} & \frac{\gamma}{\nu L_f}-\frac{1}{2L_f^2}<0,\\ -\frac{1}{2\gamma L_f}+\frac{1}{\nu\gamma L_f}-\left(\frac{\gamma}{\nu L_f}-\frac{1}{2L_f^2}\right)L_f^2 & \text{otherwise.}\end{cases}$$

since $\nu\in(0,2)$, we can verify that for any γ satisfying $\gamma<\frac{1}{L_f}$, the constant δ is strictly positive.

a) If $0<\gamma L_f<\frac{\nu}{2}<1$,

$$\frac{\delta}{2L_f}=\frac{2-\nu}{2\nu\gamma L_f}>\frac{2-\nu}{2\nu}>0.$$

b) If $\frac{\nu}{2}\le\gamma L_f<1$,

$$\frac{\delta}{2L_f}=\frac{2-\nu}{2\nu\gamma L_f}+\frac{1}{2}-\frac{\gamma L_f}{\nu}>\frac{2-\nu}{2\nu}-\frac{1}{\nu}+\frac{1}{2}=0.$$

Therefore, when $\gamma<\frac{1}{L_f}$, it ensures that $\delta>0$.

In summary, let $\left[\rho_f\right]_+=\max\left\{\rho_f,0\right\}$, the range of γ can be expressed as $\gamma<\min\left\{\frac{1}{L_f},\frac{2-\nu}{2\left[\rho_f\right]_+}\right\}$, then sequence $\left\{\Phi\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^{\mathbb{A}}}$ is nonincreasing.

□

Theorem 7 (Subsequence convergence). Suppose that Assumption 1 hold, and let $\left\{\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^*}$ be a sequence generated by Algorithm 1 with stepsize γ as in Theorem 6. The following hold

Suppose that F is coercive. Then the sequence $\left\{\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^*}$ is bounded, when $n\to+\infty$, *we have* $\|x_{n+1}-x_n\|\to 0$, $\|y_{n+1}-y_n\|\to 0$, $\|z_{n+1}-z_n\|\to 0$, $\|z_{n+1}-x_{n+1}\|\to 0$, and $\|w_{n+1}-w_n\|\to 0$. For any cluster point $\left(x^*,y^*,z^*,w^*\right)$, we have $x^*=z^*$,

$$0\in\nabla f\left(z^*\right)+\partial h\left(z^*\right)-\partial g\left(z^*\right),$$

and

$$\lim_{n\to+\infty}\Phi\left(x_n,y_n,z_n,w_n\right)=\Phi\left(x^*,y^*,z^*,w^*\right)=\lim_{n\to+\infty}F\left(z_n\right)=F\left(z^*\right).$$

Proof. First, we show that the sequence $\left\{\Phi\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^{\mathbb{A}}}$ is bounded. Since f is L_f-smooth function, the descent lemma yields the following inequalities

$$f\left(z_n\right)+\left\langle\nabla f\left(x_n\right),x_n-z_n\right\rangle-\frac{L_f}{2}\|x_n-z_n\|^2\le f\left(x_n\right),$$

$$f(z_n)+\langle \nabla f(x_n), x_n - z_n \rangle + \frac{L_f}{2}\|x_n - z_n\|^2 \ge f(x_n),$$

Combining the above two inequalities, we have

$$\begin{aligned} &f(z_n)+\langle \nabla f(x_n), x_n - z_n \rangle - \frac{L_f}{2}\|x_n - z_n\|^2 \\ &\le f(x_n) \le f(z_n)+\langle \nabla f(x_n), x_n - z_n \rangle + \frac{L_f}{2}\|x_n - z_n\|^2. \end{aligned} \tag{18}$$

According to Lemma 5(i) and $y_{n+1} = y_n + \nu(z_{n+1} - x_{n+1})$, we know that

$$\gamma \nabla f(x_n) = y_{n-1} - x_n = y_n - \nu(z_n - x_n) - x_n = (y_n - z_n) - (1-\nu)(x_n - z_n),$$

which implies that

$$\nabla f(x_n) = \frac{1}{\gamma}(y_n - z_n) - \frac{1-\nu}{\gamma}(x_n - z_n). \tag{19}$$

Now, we estimate the term $\langle \nabla f(x_n), x_n - z_n \rangle$ in (18) with (19), we derive that

$$\begin{aligned} \langle \nabla f(x_n), x_n - z_n \rangle &= \left\langle \frac{1}{\gamma}(y_n - z_n) - \frac{1}{\gamma}(1-\nu)(x_n - z_n), x_n - z_n \right\rangle \\ &= \frac{1}{\gamma}\langle y_n - z_n, x_n - z_n \rangle - \frac{1-\nu}{\gamma}\|x_n - z_n\|^2 \\ &= \frac{1}{2\gamma}\|y_n - z_n\|^2 + \frac{1}{2\gamma}\|x_n - z_n\|^2 - \frac{1}{2\gamma}\|x_n - y_n\|^2 \\ &\quad - \frac{1-\nu}{\gamma}\|x_n - z_n\|^2. \end{aligned} \tag{20}$$

Since g^* is the Fenchel conjugate of g, by Proposition 1,

$$g^*(w_n) - \langle w_n, z_n \rangle \ge -g(z_n). \tag{21}$$

Combining (18) with (20) and (21), we get

$$\begin{aligned} \Phi(x_n, y_n, z_n, w_n) &= f(x_n) + h(z_n) + g^*(w_n) - \langle w_n, z_n \rangle + \frac{1}{2\gamma}\|x_n - y_n\|^2 \\ &\quad - \frac{1}{2\gamma}\|y_n - z_n\|^2 + \frac{1-\nu}{\gamma}\|x_n - z_n\|^2 \\ &= f(x_n) + h(z_n) + g^*(w_n) - \langle w_n, z_n \rangle \\ &\quad + \langle \nabla f(x_n), z_n - x_n \rangle + \frac{1}{2\gamma}\|x_n - z_n\|^2 \\ &\ge f(z_n) + h(z_n) - g(z_n) + \langle \nabla f(x_n), z_n - x_n \rangle \\ &\quad + \frac{1}{2\gamma}\|x_n - z_n\|^2 + \langle \nabla f(x_n), x_n - z_n \rangle - \frac{L_f}{2}\|x_n - z_n\|^2 \\ &= f(z_n) + h(z_n) - g(z_n) + \left(\frac{1}{2\gamma} - \frac{L_f}{2}\right)\|x_n - z_n\|^2 \\ &= F(z_n) + \left(\frac{1}{2\gamma} - \frac{L_f}{2}\right)\|x_n - z_n\|^2. \end{aligned} \tag{22}$$

From Lemma 5(ii), we have

$$z_{n-1} \in \partial g^*(w_n) + \tau(w_n - w_{n-1}).$$

And because of Proposition 1(ii),

$$g^*(w_n) + g(z_{n-1} - \tau(w_n - w_{n-1})) = \langle w_n, z_{n-1} - \tau(w_n - w_{n-1}) \rangle,$$

subtracting $\langle w_n, z_n \rangle$ from both sides of the equation, then rearrange it

$$\begin{aligned} g^*(w_n) - \langle w_n, z_n \rangle &= \langle w_n, z_{n-1} - \tau(w_n - w_{n-1}) \rangle \\ &\quad - g(z_{n-1} - \tau(w_n - w_{n-1})) - \langle w_n, z_n \rangle, \\ &= \langle w_n, z_{n-1} - z_n - \tau(w_n - w_{n-1}) \rangle \\ &\quad - g(z_{n-1} - \tau(w_n - w_{n-1})). \end{aligned} \tag{23}$$

Then, by the convexity of g, taking $t_n \in \partial g(z_n)$ yields

$$g(z_{n-1} - \tau(w_n - w_{n-1})) \geq g(z_n) + \langle t_n, z_{n-1} - \tau(w_n - w_{n-1}) - z_n \rangle,$$

that is, $-g(z_{n-1} - \tau(w_n - w_{n-1})) \leq -g(z_n) + \langle t_n, z_n - z_{n-1} + \tau(w_n - w_{n-1}) \rangle$.

Combining (23), we obtain

$$g^*(w_n) - \langle w_n, z_n \rangle \leq -g(z_n) + \langle t_n - w_n, z_n - z_{n-1} + \tau(w_n - w_{n-1}) \rangle, \tag{24}$$

Then, combining (18) with (20) and (24), they become

$$\begin{aligned} \Phi(x_n, y_n, z_n, w_n) &\leq f(z_n) + \langle \nabla f(x_n), x_n - z_n \rangle + \frac{L_f}{2}\|x_n - z_n\|^2 \\ &\quad + h(z_n) - g(z_n) + \langle t_n - w_n, z_n - z_{n-1} + \tau(w_n - w_{n-1}) \rangle \\ &\quad + \langle \nabla f(x_n), z_n - x_n \rangle + \frac{1}{2\gamma}\|x_n - z_n\|^2 \\ &= F(z_n) + \langle t_n - w_n, z_n - z_{n-1} + \tau(w_n - w_{n-1}) \rangle \\ &\quad + \left(\frac{1}{2\gamma} + \frac{L_f}{2}\right)\|x_n - z_n\|^2. \end{aligned} \tag{25}$$

Since $\gamma < \min\left\{\frac{1}{L_f}, \frac{2-\nu}{2[\rho_f]_+}\right\}$, the size of γ is always less than $\frac{1}{L_f}$, $\frac{1}{2\gamma} - \frac{L_f}{2} > 0$ is always holds. Since F is a proper lower semicontinuous coercive function, it is bounded below [18], according to (22) that the sequence $\{\Phi(x_n, y_n, z_n, w_n)\}_{n\in\mathbb{N}^*}$ is bounded below. From (i) above, $\{\Phi(x_n, y_n, z_n, w_n)\}_{n\in\mathbb{N}^*}$ is nonincreasing. It means that it is convergent.

From Theorem 6, the following inequality holds

$$\Phi(x_n, y_n, z_n, w_n) - \Phi(x_{n+1}, y_{n+1}, z_{n+1}, w_{n+1}) \geq \frac{\delta}{2}\|x_{n+1} - x_n\|^2 + \frac{\tau}{2}\|w_{n+1} - w_n\|^2,$$

Summing the above inequality from $n = 1$ to N and using the telescoping property, we obtain

$$\frac{\delta}{2}\sum_{n=1}^{N}\left\|x_{n+1}-x_n\right\|^2+\frac{\tau}{2}\sum_{n=1}^{N}\left\|w_{n+1}-w_n\right\|^2$$
$$\leq \Phi\left(x_1,y_1,z_1,w_1\right)-\Phi\left(x_{N+1},y_{N+1},z_{N+1},w_{N+1}\right),$$

Taking $N\to+\infty$, we obtain

$$\frac{\delta}{2}\sum_{n=1}^{+\infty}\left\|x_{n+1}-x_n\right\|^2+\frac{\tau}{2}\sum_{n=1}^{+\infty}\left\|w_{n+1}-w_n\right\|^2$$
$$\leq \Phi\left(x_1,y_1,z_1,w_1\right)-\lim_{n\to+\infty}\Phi\left(x_n,y_n,z_n,w_n\right)<+\infty.$$

From the above inequality, it shows that

$$\sum_{n=1}^{+\infty}\left\|x_{n+1}-x_n\right\|^2<+\infty,\ \sum_{n=1}^{+\infty}\left\|w_{n+1}-w_n\right\|^2<+\infty,$$

which implies $\left\|x_{n+1}-x_n\right\|\to 0$, $\left\|w_{n+1}-w_n\right\|\to 0$. According to Lemma 5(iv) and (v), $\left\|y_{n+1}-y_n\right\|\to 0$, $\left\|z_{n+1}-z_n\right\|\to 0$.

Since the sequence $\left\{\Phi\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^*}$ is nonincreasing, and according to (22), we can deduce that F has an upper bound. Furthermore, since F is bounded below, we can further conclude that $\left\{F\left(z_n\right)\right\}_{n\in\mathbb{N}^*}$ and $\left\{\left\|x_n-z_n\right\|\right\}_{n\in\mathbb{N}^*}$ is bounded. Moreover, recall that F is coercive, we can derive that $\left\{z_n\right\}_{n\in\mathbb{N}^*}$ is bounded, then $\left\{x_n\right\}_{n\in\mathbb{N}^*}$ is also bounded. By Lemma 4(i), $\left\{y_n\right\}_{n\in\mathbb{N}^*}$ is bounded. According to Lemma 4(ii) and Proposition 1(ii), we obtain

$$w_{n+1}\in\partial g\left(z_n-\tau\left(w_{n+1}-w_n\right)\right).$$

Since $t_n\in\partial g\left(z_n\right)$, $\left\{z_n\right\}_{n\in\mathbb{N}^*}$ is bounded and $\left\|w_{n+1}-w_n\right\|\to 0$ as $n\to+\infty$, it follows that the sequence $\left\{w_n\right\}_{n\in\mathbb{N}^*}$ and $\left\{t_n\right\}_{n\in\mathbb{N}^*}$ are bounded.

Let $\left(x^*,y^*,z^*,w^*\right)$ be a cluster point of the sequence $\left\{\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^*}$.

Since $\left\{\Phi\left(x_n,y_n,z_n,w_n\right)\right\}_{n\in\mathbb{N}^*}$ is bounded, there exists a subsequence $\left\{\left(x_{k_n},y_{k_n},z_{k_n},w_{k_n}\right)\right\}_{n\in\mathbb{N}^{\mathbb{A}}}$ converges to $\left(x^*,y^*,z^*,w^*\right)$. It is known that as $n\to+\infty$,

$$x_{n+1}-x_n\to 0,\ y_{n+1}-y_n\to 0,\ z_{n+1}-z_n\to 0\ \text{and}\ w_{n+1}-w_n\to 0.$$

From the iteration of y_{n+1} in **Algorithm 1**, we have $x_{n+1}-z_{n+1}=-\frac{1}{\nu}\left(y_{n+1}-y_n\right)$, hence $x_{n+1}-z_{n+1}\to 0$ as $n\to+\infty$, we obtain $x^*=z^*$, then

$$\lim_{n\to+\infty}\left(x_{k_n-1},y_{k_n-1},z_{k_n-1}\right)=\lim_{n\to+\infty}\left(x_{k_n-2},y_{k_n-2},z_{k_n-2}\right)=\left(x^*,y^*,z^*\right),\tag{26}$$

if $\tau>0$,

$$\lim_{n\to+\infty}w_{k_n-1}=\lim_{n\to+\infty}w_{k_n-2}=w^*.\tag{27}$$

From the iteration of z_{n+1}, we have

$$h\left(z_{k_n}\right)+\frac{1}{2\gamma}\left\|z_{k_n}-\left(2x_{k_n}-y_{k_{n-1}}+\gamma w_{k_n}\right)\right\|^2$$
$$\leq h\left(z^*\right)+\frac{1}{2\gamma}\left\|z^*-\left(2x_{k_n}-y_{k_{n-1}}+\gamma w_{k_n}\right)\right\|^2.$$

Similar for w_{n+1}, we have

$$g^*\left(w_{k_n}\right)-\left\langle w_{k_n}, z_{k_{n-1}}\right\rangle+\frac{\tau}{2}\left\|w_{k_n}-w_{k_{n-1}}\right\|^2$$
$$\leq g^*\left(w^*\right)-\left\langle w^*, z_{k_{n-1}}\right\rangle+\frac{\tau}{2}\left\|w^*-w_{k_{n-1}}\right\|^2.$$

Using (26) and (27), taking the limit yields

$$\limsup_{n\to+\infty} h\left(z_{k_n}\right)\leq h\left(z^*\right),\ \limsup_{n\to+\infty} g^*\left(w_{k_n}\right)\leq g^*\left(w^*\right).$$

Additionally, since h and g^* are lower semicontinuous, we know that $\liminf_{n\to+\infty} h\left(z_{k_n}\right)\geq h\left(z^*\right)$ and $\liminf_{n\to+\infty} g^*\left(w_{k_n}\right)\geq g^*\left(w^*\right)$. Consequently,

$$\lim_{n\to+\infty} h\left(z_{k_n}\right)=h\left(z^*\right),\ \lim_{n\to+\infty} g^*\left(w_{k_n}\right)=g^*\left(w^*\right).$$

Recall that Lemma 5(i) and (iii), we have

$$w_{k_n}\in\nabla f\left(x_{k_n}\right)-\frac{1}{\gamma}\left(x_{k_n}-z_{k_n}\right)+\partial h\left(z_{k_n}\right).$$

Since $x^*=z^*$, by taking the limit, we have the following result

$$w^*\in\nabla f\left(z^*\right)+\partial h\left(z^*\right). \tag{28}$$

Similarly, from Lemma 5(ii), it follows that

$$z_n\in\partial g^*\left(w_{n+1}\right)+\tau\left(w_{n+1}-w_n\right) \tag{29}$$

Taking the limit on both sides of (28), it yields $z^*\in\partial g^*\left(w^*\right)$. Using Proposition 1, it is not difficult to see that

$$g^*\left(w^*\right)-\left\langle w^*, z^*\right\rangle=-g\left(z^*\right)\ \text{and}\ w^*\in\partial g\left(z^*\right) \tag{30}$$

Therefore, combining (28) and (30) leads to

$$0\in\nabla f\left(z^*\right)+\partial h\left(z^*\right)-\partial g\left(z^*\right).$$

Then, according to (4), we have

$$\Phi\left(x_{k_n}, y_{k_n}, z_{k_n}, w_{k_n}\right)=f\left(x_{k_n}\right)+h\left(z_{k_n}\right)+g^*\left(w_{k_n}\right)-\left\langle w_{k_n}, z_{k_n}\right\rangle$$
$$+\frac{1}{2\gamma}\left\|x_{k_n}-y_{k_n}\right\|^2-\frac{1}{2\gamma}\left\|y_{k_n}-z_{k_n}\right\|^2+\frac{1-\nu}{\gamma}\left\|x_{k_n}-z_{k_n}\right\|^2.$$

Taking limit on both sides of the above equality and using the continuity of f, we deduce that

$$\begin{aligned}\lim_{n\to+\infty}\Phi\left(x_{k_n}, y_{k_n}, z_{k_n}, w_{k_n}\right)&=f\left(z^*\right)+h\left(z^*\right)+g^*\left(w^*\right)-\left\langle w^*, z^*\right\rangle\\&=\Phi\left(x^*, y^*, z^*, w^*\right).\end{aligned} \tag{31}$$

According to (30) and the convergence of the sequence $\left\{\Phi\left(x_n, y_n, z_n, w_z\right)\right\}_{n\in\mathbb{N}^*}$, we know

$$\begin{aligned}\lim_{n\to+\infty}\Phi\left(x_n, y_n, z_n, w_n\right)&=f\left(x^*\right)+h\left(z^*\right)+g^*\left(w^*\right)-\left\langle w^*, z^*\right\rangle\\&=f\left(z^*\right)+h\left(z^*\right)-g\left(z^*\right)\\&=F\left(z^*\right).\end{aligned}$$

From (22), (25), together with the boundedness of the sequences $\left\{w_n\right\}_{n\in\mathbb{N}^*}$

and $\{t_n\}_{n\in\mathbb{N}^*}$, and the fact that $z_{n+1}-z_n \to 0$, $w_{n+1}-w_n \to 0$, and $x_{n+1}-z_{n+1} \to 0$ as $n \to 0$, we have $\lim_{n\to+\infty} F(z_n) = \lim_{n\to+\infty} \Phi(x_n, y_n, z_n, w_n)$, thus the theorem is proved. □

Remark 3. Compared with the result of Pham *et al.*

$$\gamma < \frac{-\nu\rho_f + \sqrt{\nu^2\rho_f^2 + 8(2-\nu)L_f^2}}{4L_f^2},$$

the expression for γ that we derived, namely

$$\gamma < \min\left\{\frac{1}{L_f}, \frac{2-\nu}{2[\rho_f]_+}\right\},$$

is considerably simpler and more intuitive in form. This formulation clearly reveals the independent roles of the two key constants, $\frac{1}{L_f}$ reflects the restriction imposed by the smoothness of f, while $\frac{2-\nu}{2[\rho_f]_+}$ captures the limitation due to its hypoconvexity. The final step size must satisfy both constraints simultaneously. Moreover, one can intuitively see how changes in L_f or ρ_f affect the allowable step size range.

We now show that the step size range we obtain is always not smaller than that of Pham *et al.* First, if we require

$$\frac{-\nu\rho_f + \sqrt{\nu^2\rho_f^2 + 8(2-\nu)L_f^2}}{4L_f^2} \le \frac{1}{L_f},$$

then it follows that $\rho_f \ge -L_f$. Second, if we require

$$\frac{-\nu\rho_f + \sqrt{\nu^2\rho_f^2 + 8(2-\nu)L_f^2}}{4L_f^2} \le \frac{2-\nu}{2[\rho_f]_+},$$

then it follows that $\rho_f^2 \le L_f^2$, *i.e.*, $\rho_f \in [-L_f, L_f]$, which is exactly consistent with the assumption.

Thus, the new admissible set contains the old one. Notably, when $\nu = 1$, $\rho_f = L_f$, the range of γ is the same as that of Pham *et al.*

4. Conclusion

We revisit the Backward-Douglas-Rachford algorithm proposed by Pham *et al.* for solving generalized DC programming problems. While retaining the original assumptions, we adopt the analysis framework of Themelis *et al.* to derive a step size condition for BDRSM that has a larger range and a more concise expression. This step size range is always not smaller than that obtained by Pham *et al.*, and the two are equivalent in special cases. Moreover, under this step size range, we establish the subsequential convergence of the iterative sequence generated by

BDRSM.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Chuang, C., He, H. and Zhang, Z. (2021) A Unified Douglas-Rachford Algorithm for Generalized DC Programming. *Journal of Global Optimization*, **82**, 331-349. https://doi.org/10.1007/s10898-021-01079-y

[2] Tao, P.D. and Souad, E.B. (1986) Algorithms for Solving a Class of Nonconvex Optimization Problems. Methods of Subgradients. *North-Holland Mathematics Studies*, **129**, 249-271. https://doi.org/10.1016/s0304-0208(08)72402-2

[3] Le Thi, H.A. and Pham Dinh, T. (2018) DC Programming and DCA: Thirty Years of Developments. *Mathematical Programming*, **169**, 5-68. https://doi.org/10.1007/s10107-018-1235-y

[4] Tao, P.D. and An, L.T.H. (1998) A D.C. Optimization Algorithm for Solving the Trust-Region Subproblem. *SIAM Journal on Optimization*, **8**, 476-505. https://doi.org/10.1137/s1052623494274313

[5] Sun, T., Yin, P., Cheng, L. and Jiang, H. (2017) Alternating Direction Method of Multipliers with Difference of Convex Functions. *Advances in Computational Mathematics*, **44**, 723-744. https://doi.org/10.1007/s10444-017-9559-3

[6] Pham, T.N., Dao, M.N., Amjady, N. and Shah, R. (2025) A Proximal Splitting Algorithm for Generalized DC Programming with Applications in Signal Recovery. *European Journal of Operational Research*, **326**, 42-53. https://doi.org/10.1016/j.ejor.2025.04.034

[7] Douglas, J. and Rachford, H.H. (1956) On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables. *Transactions of the American Mathematical Society*, **82**, 421-439. https://doi.org/10.1090/s0002-9947-1956-0084194-4

[8] Lions, P.L. and Mercier, B. (1979) Splitting Algorithms for the Sum of Two Nonlinear Operators. *SIAM Journal on Numerical Analysis*, **16**, 964-979. https://doi.org/10.1137/0716071

[9] Li, G. and Pong, T.K. (2016) Douglas-Rachford Splitting for Nonconvex Optimization with Application to Nonconvex Feasibility Problems. *Mathematical Programming*, **159**, 371-401. https://doi.org/10.1007/s10107-015-0963-5

[10] Themelis, A. and Patrinos, P. (2020) Douglas-Rachford Splitting and ADMM for Nonconvex Optimization: Tight Convergence Results. *SIAM Journal on Optimization*, **30**, 149-181. https://doi.org/10.1137/18m1163993

[11] Themelis, A., Stella, L. and Patrinos, P. (2022) Douglas-Rachford Splitting and ADMM for Nonconvex Optimization: Accelerated and Newton-Type Linesearch Algorithms. *Computational Optimization and Applications*, **82**, 395-440. https://doi.org/10.1007/s10589-022-00366-y

[12] Tran Dinh, Q. and Pham, N. H. and Phan, D. and Nguyen, L. (2021) FedDR—Randomized Douglas-Rachford Splitting Algorithms for Nonconvex Federated Composite Optimization. arXiv: 2103.03452.

[13] Patrinos, P., Stella, L. and Bemporad, A. (2014) Douglas-Rachford Splitting: Complexity Estimates and Accelerated Variants. 53*rd IEEE Conference on Decision and*

Control, Los Angeles, 15-17 December 2014, 4234-4239.
https://doi.org/10.1109/cdc.2014.7040049

[14] Kurdyka, K. (1998) On Gradients of Functions Definable in O-Minimal Structures. *Annales de l'institut Fourier*, **48**, 769-783.

[15] Łojasiewicz, S. (1963) Une propriété topologique des sous-ensembles analytiques reels. *Les Équations aux Dérivées Partielles*, **117**, 87-89.

[16] Banert, S. and Boţ, R.I. (2018) A General Double-Proximal Gradient Algorithm for D.C. Programming. *Mathematical Programming*, **178**, 301-326.
https://doi.org/10.1007/s10107-018-1292-2

[17] Nesterov, Y. (2004) Introductory Lectures on Convex Optimization: A Basic Course, Applied Optimization. Kluwer Academic Publishers.

[18] Rockafellar, R.T. and Wets, R.J.B. (1998) Variational Analysis. Springer.
https://doi.org/10.1007/978-3-642-02431-3