Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2023-0113

Vol. **xx**, No. x, pp. 1-39
x 2024

Convergence Analysis of Split-Douglas-Rachford Algorithm and a Novel Preconditioned ADMM with an Improved Condition

Maoran Wang[1], Xingju Cai[2] and Yongxin Chen[3,*]

[1] *School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, P.R. China*
[2] *School of Mathematical Sciences, Nanjing Normal University, Jiangsu Key Laboratory for NSLSCS, Nanjing 210023, P.R. China*
[3] *LMIB, School of Mathematical Sciences, Beihang University, Beijing 100191, P.R. China*

Received 22 September 2023; Accepted (in revised version) 9 March 2024

Abstract. For the primal-dual monotone inclusion problem, the split-Douglas-Rachford (SDR) algorithm is a well-known first-order splitting method. Similar to other first-order algorithms, the efficiency of SDR is greatly influenced by its step parameters. Therefore, expanding the range of stepsizes may lead to improved numerical performance. In this paper, we prove that the stepsize range of SDR can be expanded based on a series of properties of the product Hilbert space. Moreover, we present a concise counterexample to illustrate that the newly proposed stepsize range cannot be further enhanced. Furthermore, to bridge the theoretical gap in the convergence rate of SDR, we analyze the descent of SDR's fixed point residuals and provide the first proof of a sublinear convergence rate for the fixed point residuals. As an application, we utilize SDR to solve separable convex optimization problems with linear equality constraints and develop a novel preconditioned alternating direction method of multipliers (NP-ADMM). This method can handle scenarios where two linear operators are not identity operators. We propose strict convergence conditions and convergence rates for the preconditioned primal-dual split (P-PDS) method and NP-ADMM by demonstrating the relationship between SDR, P-PDS, and NP-ADMM. Finally, we conduct four numerical experiments to verify the computational efficiency of these algorithms and demonstrate that the performance of these algorithms has been significantly improved with the improved conditions.

AMS subject classifications: 90C25, 90C30, 90C33, 90C47

Key words: Primal-dual monotone inclusion problem, split-Douglas-Rachford algorithm, preconditioned ADMM, improved convergence condition, sublinear convergence rate.

*Corresponding author. *Email addresses:* 230901010@njnu.edu.cn (M.R. Wang), caixingju@njnu.edu.cn (X.J. Cai), chenyongxin@buaa.edu.cn (Y.X. Chen)

1. Introduction

Let $\mathcal{G}$ and $\mathcal{H}$ be two real Hilbert spaces. Let $B : \mathcal{G} \to 2^{\mathcal{G}}$ and $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximal monotone operators, and let Id be the identity operator. Let $L : \mathcal{H} \to \mathcal{G}$ be a continuous nonzero linear operator. $L^* : \mathcal{G} \to \mathcal{H}$ is the adjoint of L. This paper focuses on the primal-dual monotone inclusion problem, which can be described as follows:

$$\text{Find} \quad (u,x) \in \mathcal{G} \times \mathcal{H} \quad \text{s.t.} \begin{cases} 0 \in Ax + L^*u, \\ 0 \in B^{-1}u - Lx. \end{cases} \tag{1.1}$$

The solution set of (1.1) is defined as

$$\mathbf{Z} := \big\{(u,x) \in \mathcal{G} \times \mathcal{H} \mid 0 \in B^{-1}u - Lx, 0 \in Ax + L^*u\big\}.$$

We assume $\mathbf{Z}$ is nonempty. Problem (1.1) finds wide applications in various fields including variational inequalities [20], optimization [37], economics and traffic theory [22], signal and image processing [12], and differential inclusion [3,38].

A classical instance of problem (1.1) is the following monotone inclusion problem:

$$\begin{cases} 0 \in \partial f(x) + L^*u, \\ 0 \in \partial h^*(u) - Lx, \end{cases} \tag{1.2}$$

where $f : \mathcal{H} \to (-\infty, \infty]$ and $h : \mathcal{G} \to (-\infty, \infty]$ are proper lower semicontinuous convex functions. Problem (1.2) can be equivalently written as the following convex-concave saddle point problem:

$$\min_{x \in \mathcal{H}} \max_{u \in \mathcal{G}} \big\{ f(x) + \langle u, Lx \rangle - h^*(u) \big\}. \tag{1.3}$$

Moreover, if (u^*, x^*) is a solution to problem (1.2), then x^* is a solution to the following convex optimization problem:

$$\min_{x \in \mathcal{H}} f(x) + h(Lx), \tag{1.4}$$

and u^* is a solution to the dual problem of (1.4)

$$\min_{u \in \mathcal{K}} f^*(-L^*u) + h^*(u). \tag{1.5}$$

For solving problem (1.3), the Arrow-Hurwicz-Uzawa algorithm (AHUA) was first proposed in [2]. The recursion of AHUA is described as

$$\begin{cases} y_{k+1} = \arg\min_{y \in \mathcal{G}} \left\{ h^*(y) + \dfrac{1}{2\sigma} \|y - (y_k + \sigma L x_k)\|^2 \right\}, \\ x_{k+1} = \arg\min_{x \in \mathcal{H}} \left\{ f(x) + \dfrac{1}{2\tau} \|x - (x_k - \tau L^* y_{k+1})\|^2 \right\}. \end{cases}$$

The convergence of AHUA can be guaranteed by choosing sufficiently small primal step τ and dual step σ, or by assuming strong convexity of the function. However, excessively small steps may lead to reduced computational efficiency of the algorithm. To address this limitation, Chambolle and Pock [11] improved AHUA and introduced the primal-dual hybrid gradient algorithm (PDHG)

$$\begin{cases} y_{k+1} = \arg\min_{y\in\mathcal{G}} \left\{ h^*(y) + \frac{1}{2\sigma}\|y - (y_k + \sigma L x_k)\|^2 \right\}, \\ x_{k+1} = \arg\min_{x\in\mathcal{H}} \left\{ f(x) + \frac{1}{2\tau}\|x - (x_k - \tau L^*(2y_{k+1} - y_k))\|^2 \right\}. \end{cases}$$

The convergence of PDHG can be established when the condition $\tau\sigma\|L\|^2 < 1$ is satisfied [11]. He and Yuan [27] simplified the proof of PDHG's convergence in [11] by interpreting PDHG as a type of variable metric proximal point algorithm. Additionally, they proposed a primal-dual-based contraction method, which can be summarized as follows:

$$\begin{cases} \tilde{y}_k = \arg\min_{y\in\mathcal{G}} \left\{ h^*(y) + \frac{1}{2\sigma}\|y - (y_k + \sigma L x_k)\|^2 \right\}, \\ \bar{y}_k = \tilde{y}_k + \theta(\tilde{y}_k - y_k), \\ \tilde{x}_k = \arg\min_{x\in\mathcal{H}} \left\{ f(x) + \frac{1}{2\tau}\|x - (x_k - \tau L^*\bar{y}_k)\|^2 \right\}, \\ u_{k+1} = u_k - \gamma\alpha_k H^{-1}M(u_k - \tilde{u}_k), \end{cases} \tag{1.6}$$

where $u := (y, x)$ and the choices of α_k, H, and M are detailed in [27]. He and Yuan [27] proved that the convergence of (1.6) can be guaranteed by appropriately selecting the values of σ, τ, and θ. Moreover, their numerical experiments also showed that (1.6) has better numerical performance than PDHG. For the special saddle point problem

$$\min_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} \left\{ \langle y, Ax\rangle + \frac{\nu}{2}\|By - b\|^2 \right\}, \tag{1.7}$$

where $\nu > 0, \mathcal{Y}$ and $\mathcal{X}$ are nonempty closed convex sets. In [9, 28], for solving (1.7), the ranges of σ, τ, and θ in (1.6) can be further expanded to improve computational efficiency. Condat [15] improved the parameter condition of PDHG in finite-dimensional Hilbert spaces to

$$\tau\sigma\|L\|^2 \leq 1. \tag{1.8}$$

For a proper lower semicontinuous convex function $f : \mathcal{H} \to [-\infty, +\infty)$, the subdifferential ∂f is a maximal monotone operator [4, Theorem 20.25]. However, not all maximal monotone operators are subdifferential operators [37]. Therefore, problem (1.2) is a special case of problem (1.1). For instance, certain variational inequality problems [20], equilibrium problems [22], and differential inclusion problems [3, 38] can be expressed in the form of problem (1.1), but not in the form of problem (1.2). To solve the more general problem (1.1), an efficient split algorithm is the primal-dual splitting (PDS) method proposed in [42]. This method generates a sequence in $\mathcal{H} \times \mathcal{G}$ using the iterative scheme

$$\begin{cases} x_{n+1} = J_{\tau A}(x_n - \tau L^* v_n), \\ v_{n+1} = J_{\sigma B^{-1}}\big(v_n + \sigma L(2x_{n+1} - x_n)\big) \end{cases} \tag{1.9}$$

with initial point $(x_0, v_0) \in \mathcal{H} \times \mathcal{G}$ and positive numbers τ, σ satisfying $\tau\sigma\|L\|^2 < 1$. For problem (1.2), PDS coincides with PDHG. To avoid the estimation of the operator norm of L and accelerate the convergence, the preconditioned versions of PDS and PDHG (P-PDS, P-PDHG) were proposed in [14, 35], respectively. In detail, τId and σId are generalized to Y and Σ, respectively, where Y and Σ are two strongly monotone self-adjoint continuous linear operators. In [14,35], the convergence of P-PDHG and P-PDS can be guaranteed when Σ and Y meet

$$\big\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}}\big\| < 1, \tag{1.10}$$

where $\Sigma^{\frac{1}{2}}$ and $Y^{\frac{1}{2}}$ are square roots of Σ and Y, respectively, i.e., $\Sigma = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}$, $Y = Y^{\frac{1}{2}}Y^{\frac{1}{2}}$.

Very recently, for a special case of problem (1.3), by setting $h^*(u) = \langle u, b\rangle$, the convergence condition of PDHG can be improved to

$$\tau\sigma\|L\|^2 < \frac{4}{3}, \tag{1.11}$$

by He *et al.* [23]. Next, for the more general case where h^* is not necessarily linear, the global convergence of PDHG in finite-dimensional Hilbert spaces can be demonstrated under condition (1.11) by Li and Yan [31]. The convergence of P-PDHG in finite-dimensional Hilbert spaces was proved by Jiang *et al.* [29] and Ma *et al.* [34] under two conditions similar to (1.11).

For solving the problem (1.1) in general Hilbert spaces, the split-Douglas-Rachford (SDR) algorithm was proposed by Briceño-Arias and Roldán [6]. The recursion of the SDR reads as

$$\begin{cases} v_n = \Sigma(Id - J_{\Sigma^{-1}B})(Lx_n + \Sigma^{-1}u_n), \\ z_n = x_n - YL^* v_n, \\ x_{n+1} = J_{YA} z_n, \\ u_{n+1} = \Sigma L(x_{n+1} - x_n) + v_n. \end{cases} \tag{1.12}$$

The weak convergence of the sequence generated by SDR was proved by [6] under the condition

$$\big\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}}\big\| \le 1, \tag{1.13}$$

where $\Sigma^{\frac{1}{2}}, Y^{\frac{1}{2}}$ are square roots of Σ and Y, respectively.

Motivation and our contribution. It is widely recognized that the convergence conditions of P-PDHG can be enhanced [29, 34], and SDR can be reduced to P-PDHG for problem (1.2) [6]. This leads to some natural questions: Can the convergence condition of SDR be improved? What is the convergence rate of SDR in terms of fixed point residuals? In this paper, we address these questions by examining the weaker convergence conditions of SDR and the convergence rate of SDR concerning fixed point residuals. Our main contributions are as follows. Building upon the insights from [29, 31, 34], we refine the convergence condition (1.13) of SDR to

$$\left\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\right\|^2 < \frac{4}{3}. \tag{1.14}$$

Furthermore, we demonstrate that (1.14) cannot be further improved, i.e., the inequality (1.14) cannot be replaced by "$\leq$", as evidenced by a counterexample in finite-dimensional Hilbert spaces. A summary of the algorithms and convergence conditions mentioned above is provided in Table 1. The results established in this paper are highlighted in bold, and the symbol "$-$" indicates that A and B are only maximal monotone and not necessarily the subdifferentials of proper lower semicontinuous convex functions.

The numerical experiments in Section 5 demonstrate a significant improvement in the performance of SDR when using the improved condition (1.14). Furthermore, based on certain properties of product Hilbert spaces, we establish the sublinear convergence rate of SDR in terms of fixed point residuals under the condition (1.13). To the best of our knowledge, our convergence rate result is the first to address the sublinear and non-ergodic convergence rate of SDR in terms of fixed point residuals. Finally, we propose a novel preconditioned ADMM (NP-ADMM) by applying SDR to solve the following separable convex optimization problem with linear equality constraints:

$$\begin{aligned} \min_{x\in\mathcal{H}, w\in\mathcal{K}} \quad & f(x) + g(w) \\ \text{s.t.} \quad & Lx + Jw = b, \end{aligned}$$

Table 1: The comparison of the convergence conditions of some primal-dual algorithms.

Algorithm	A	B	Σ, Y	$\mathcal{H}, \mathcal{G}$
PDHG [11]	∂f	∂h	$\sigma Id,\ \tau Id,\ \tau\sigma\|L\|^2 < 1$	finite dimension
Condat [15]	∂f	∂h	$\sigma Id,\ \tau Id,\ \tau\sigma\|L\|^2 (<) \leq 1$	(in)finite dimension
He *et al.* [23]	∂f	$\operatorname{dom} B = \{b\}$	$\sigma Id,\ \tau Id,\ \tau\sigma\|L\|^2 < \frac{4}{3}$	finite dimension
Li and Yan [31]	∂f	∂h	$\sigma Id,\ \tau Id,\ \tau\sigma\|L\|^2 < \frac{4}{3}$	finite dimension
P-PDHG [35]	∂f	∂h	$\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\| < 1$	finite dimension
Jiang *et al.* [29]	∂f	∂h	$\Sigma = \sigma Id,\ \sigma\|LY^{\frac{1}{2}}\|^2 < \frac{4}{3}$	finite dimension
Ma *et al.* [34]	∂f	∂h	$\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2 < \frac{4}{3}$	finite dimension
PDS [42]	$-$	$-$	$\sigma Id,\ \tau Id,\ \tau\sigma\|L\|^2 < 1$	infinite dimension
SDR [6]	$-$	$-$	$\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2 \leq 1$	infinite dimension
This paper (SDR)	$-$	$-$	$\boldsymbol{\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2 < \frac{4}{3}}$	**infinite dimension**

Table 2: The comparison of iPADMM and NP-ADMM.

Algorithm	$\mathcal{H}, \mathcal{G}, \mathcal{K}$	J, b
iPADMM [34] [32]	finite dimension, $\mathcal{K} = \mathcal{G}$	$J = -Id$, $b = 0$
NP-ADMM	infinite dimension	$-$

where $f : \mathcal{H} \to (-\infty, \infty]$ and $g : \mathcal{G} \to (-\infty, \infty]$ denote proper lower semicontinuous convex functions, $L : \mathcal{H} \to \mathcal{G}$ and $J : \mathcal{K} \to \mathcal{G}$ are linear continuous operators, and $\mathcal{G}, \mathcal{H}$ and $\mathcal{K}$ are real Hilbert spaces. Unlike the indefinite proximal ADMM (iPADMM) proposed in [32, 34], NP-ADMM can handle more general situations where both L and J are not Id, meaning that NP-ADMM can solve more general problems than iPADMM. We summarize the main differences between iPADMM and NP-ADMM in Table 2. Under some assumptions, we prove the global convergence and sublinear convergence rates of NP-ADMM, and its tight convergence condition is also proposed.

This paper is organized as follows. Section 2 presents the notations and preliminaries. In Section 3, we prove the convergence of SDR under condition (1.14) and propose a counter-example to demonstrate the tightness of condition (1.14). Additionally, we establish the sublinear convergence rate of SDR in terms of fixed point residuals under condition (1.13). Section 4 introduces a novel preconditioned ADMM (NP-ADMM) method for solving separable convex optimization problems with linear equality constraints, and provides theoretical results on the convergence and rates of NP-ADMM. In Section 5, we conduct numerical experiments to demonstrate the efficiency of SDR and NP-ADMM with the improved condition (1.14). Finally, Section 6 concludes the paper.

2. Notations and preliminaries

In this paper, $\mathcal{G}$ and $\mathcal{H}$ denote two real Hilbert spaces with the inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. The symbols "$\to$" and "$\rightharpoonup$" denote strong and weak convergence, respectively. For a continuous nonzero linear operator $L : \mathcal{H} \to \mathcal{G}$, we denote its adjoint operator by $L^* : \mathcal{G} \to \mathcal{H}$, the range of L by $\operatorname{ran} L$, and the kernel of L by $\ker L$. Id represents the identity operator.

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a set-value operator. The domain, range, graph, and the zero point set of A are denoted by $\operatorname{dom} A = \{x \in \mathcal{H} \mid Ax \neq \varnothing\}$, $\operatorname{ran} A = \bigcup_{x \in \operatorname{dom} A} Ax$, $\operatorname{gra} A = \{(x, y) \in \mathcal{H} \times \mathcal{H} \mid y \in Ax, x \in \operatorname{dom} A\}$, and $\operatorname{zer} A = \{x \in \mathcal{H} \mid 0 \in Ax\}$, respectively. The inverse of A is denoted by $A^{-1} : y \to \{x \mid y \in Ax\}$. The operator A is monotone, if for every $(x, y), (z, w) \in \operatorname{gra} A$, we have $\langle y - w, x - z \rangle \geq 0$. The operator A is μ-strongly monotone (μ >0), if for every $(x, y), (z, w) \in \operatorname{gra} A$, we have $\langle y - w, x - z \rangle \geq \mu \|x - z\|^2$. The operator A is maximal monotone if it is monotone and there exists no monotone operator $B : \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} B$ properly contains $\operatorname{gra} A$. The resolvent of a maximal monotone operator A is denoted by $J_A = (Id + A)^{-1}$, and we have $\operatorname{dom} J_A = \mathcal{H}$ [4, Theorem 21.1]. Using the definition of the resolvent, we can directly obtain that

$$y = J_A x \quad \Leftrightarrow \quad x - y \in Ay.$$

For any given self-adjoint continuous linear operator $S : \mathcal{H} \to \mathcal{H}$, we define $\| \cdot \|_S^2 = \langle \cdot, \cdot \rangle_S$, where $\langle \cdot, \cdot \rangle_S : (x, y) \mapsto \langle Sx, y \rangle$ is symmetric and bilinear, but for $x \in \mathcal{H}$, $\langle x, x \rangle_S$ can be negative. For every x, y, z, and w in $\mathcal{H}$, we have

$$\|x + y\|_S^2 = \|x\|_S^2 + \|y\|_S^2 + 2\langle x, y \rangle_S, \tag{2.1}$$

and

$$2\langle x - y, z - w \rangle_S = \|x - w\|_S^2 + \|y - z\|_S^2 - \|x - z\|_S^2 - \|y - w\|_S^2. \tag{2.2}$$

Further, if S is a self-adjoint, monotone, and continuous linear operator, then $\langle \cdot, \cdot \rangle_S$ is a semi-inner product on $\mathcal{H}$. If S is a self-adjoint, strongly monotone and continuous linear operator, then $\langle \cdot, \cdot \rangle_S$ can be an inner product on $\mathcal{H}$, and the topologies of $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ and $(\mathcal{H}, \langle \cdot, \cdot \rangle_S)$ are equivalent.

We define $\Gamma_0(\mathcal{H})$ as the set of proper lower semicontinuous convex functions $f : \mathcal{H} \to (-\infty, \infty]$. Let $f \in \Gamma_0(\mathcal{H})$, then the subdifferential of f which is defined by

$$\partial f : x \mapsto \big\{ u \in \mathcal{H} \mid f(y) \geq f(x) + \langle u, y - x \rangle, \forall y \in \mathcal{H} \big\}$$

is maximal monotone. The set of minimizers of f, denoted by $\arg\min_{x \in \mathcal{H}} f(x)$, is represented by $\operatorname{zer} \partial f$. The resolvent of ∂f is also called the proximal operator of f, which is also denoted by $\operatorname{prox}_f : x \mapsto \arg\min_{y \in \mathcal{H}} \{ f(y) + \|y - x\|^2/2 \}$. Generally, we denote

$$\operatorname{prox}_f^Y : x \quad \mapsto \quad \arg\min_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2}\|y - x\|_Y^2 \right\}, \tag{2.3}$$

where $Y : \mathcal{H} \to \mathcal{H}$ is a strongly monotone self-adjoint linear operator. We have $\operatorname{prox}_f = \operatorname{prox}_f^{Id}$ and $\operatorname{prox}_f^Y = J_{Y^{-1}\partial f}$ [4, Proposition 24.24]. The Fenchel conjugate of f is defined by $f^* : u \mapsto \sup_{x \in \mathcal{H}} \{ \langle u, x \rangle - f(x) \}$. We have $f^* \in \Gamma_0(\mathcal{H})$ and $\partial f^* = (\partial f)^{-1}$, if $f \in \Gamma_0(\mathcal{H})$. It follows from [4, Proposition 23.34]

$$J_{Y^{-1}A} + Y^{-1} J_{YA^{-1}} Y = Id, \tag{2.4}$$

that

$$\operatorname{prox}_f^Y + Y^{-1} \operatorname{prox}_{f^*}^{Y^{-1}} Y = Id, \tag{2.5}$$

where (2.4) is called the generalized Moreau identity. Given a nonempty convex set C, let $\operatorname{sri} C$ represent the strong relative interior of C. If C is a nonempty closed convex set, we denote the indicator function of C as $I_C \in \Gamma_0(\mathcal{H})$ and the normal cone of C as $\mathcal{N}_C = \partial I_C$. Using the definition of resolvent of $\mathcal{N}_C$, we have $J_{\mathcal{N}_C} = P_C$, where P_C is the projection operator onto the closed convex set C.

Lemma 2.1 (Opial Lemma [4, Lemma 2.47]). *Let $\{c_n\}_{n=1}^{\infty}$ be a sequence in real Hilbert space $\mathcal{E}$ and let C be a nonempty subset of $\mathcal{E}$. Suppose that for any given $c \in C$, $\{\|c_n - c\|\}_{n=1}^{\infty}$ converges and that every weak sequential cluster point of $\{c_n\}_{n=1}^{\infty}$ belongs to C. Then $\{c_n\}_{n=1}^{\infty}$ converges weakly to a point in C.*

Theorem 2.1 (Fenchel-Rockafellar Duality Theorem [4, Lemma 15.23]). *Let $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, and let $L : \mathcal{H} \to \mathcal{G}$ be a continuous linear operator such that*

$$0 \in \operatorname{sri}\big(\operatorname{dom} g - L(\operatorname{dom} f)\big).$$

Then $\inf(f + g \circ L)(\mathcal{H}) = -\min(f^* \circ (-L^*) + g^*)(\mathcal{G})$.

Lemma 2.2 (Cauchy-Bunyakowsky-Schwarz Inequality [16, Proposition 1.4]). *If $\langle \cdot, \cdot \rangle$ is a semi-inner product on $\mathcal{H}$, then*

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

for all x and y in $\mathcal{H}$, where $\| \cdot \| := \sqrt{\langle \cdot, \cdot \rangle}$.

Lemma 2.3 ([5, Proposition 2.7]). *Let $\mathbf{x} = (u, x) \in \mathcal{G} \oplus \mathcal{H}$. Denote $\mathbf{M} : \mathbf{x} \to (B^{-1}u, Ax)$, $\mathbf{S} : \mathbf{x} \to (-Lx, L^*u)$, then $\mathbf{Z} = \operatorname{zer}(\mathbf{M} + \mathbf{S})$ and $\mathbf{M} + \mathbf{S}$ is maximal monotone in $\mathcal{G} \oplus \mathcal{H}$, where $\mathcal{G} \oplus \mathcal{H}$ is the Hilbert direct sum of $\mathcal{G}$ and $\mathcal{H}$ equipped with the inner product $\langle \mathbf{x_1}, \mathbf{x_2} \rangle_{\mathcal{G}\oplus\mathcal{H}} := \langle u_1, u_2\rangle + \langle x_1, x_2\rangle$, where $\mathbf{x}_i = (u_i, x_i) \in \mathcal{G} \oplus \mathcal{H}, u_i \in \mathcal{G}, x_i \in \mathcal{H}, i = 1, 2$.*

The following result involves the properties of monotone linear operators.

Lemma 2.4. *Let $\mathcal{G}$ and $\mathcal{H}$ be real Hilbert spaces. Let $L : \mathcal{H} \to \mathcal{G}$ be a continuous linear operator. Let $\Sigma : \mathcal{G} \to \mathcal{G}$ and $Y : \mathcal{H} \to \mathcal{H}$ be self-adjoint strongly monotone continuous linear operators, and α is a given positive number. Then, the following assertions are equivalent:*

(i) $\big\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}}\big\|^2 < \dfrac{1}{\alpha}$.

(ii) $Y^{-1} - \alpha L^* \Sigma L$ *is strongly monotone.*

Proof. Since Σ and Y are self-adjoint strongly monotone continuous linear operators, there exist self-adjoint strongly monotone continuous linear operators $\Sigma^{\frac{1}{2}}$ and $Y^{\frac{1}{2}}$ such that $\Sigma = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}$, $Y = Y^{\frac{1}{2}}Y^{\frac{1}{2}}$ [30, Theorem 4.3]. Furthermore, $\Sigma, \Sigma^{\frac{1}{2}}, Y$ and $Y^{\frac{1}{2}}$ are bijections.

(ii) $\Rightarrow$ (i). Since $Y^{-1} - \alpha L^*\Sigma L$ is strongly monotone, there exists a positive number μ, such that

$$\big\langle (Y^{-1} - \alpha L^*\Sigma L)x, x\big\rangle \geq \mu\|x\|^2 > 0, \quad \forall x \in \mathcal{H}\backslash\{0\}.$$

Noting that $Y^{-\frac{1}{2}}$ is a bijection and combining the following identity:

$$\big\langle (Y^{-1} - \alpha L^*\Sigma L)x, x\big\rangle = \big\|Y^{-\frac{1}{2}}x\big\|^2 \left(1 - \alpha\frac{\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}} Y^{-\frac{1}{2}} x\|^2}{\|Y^{-\frac{1}{2}}x\|^2}\right), \tag{2.6}$$

we get

$$1 - \alpha\frac{\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}} Y^{-\frac{1}{2}} x\|^2}{\|Y^{-\frac{1}{2}}x\|^2} \geq \frac{\mu\|x\|^2}{\|Y^{-\frac{1}{2}}x\|^2}$$

$$\geq \mu \inf_{x\in\mathcal{H}\backslash\{0\}} \frac{\|x\|^2}{\|Y^{-\frac{1}{2}}x\|^2} = \frac{\mu}{\|Y^{-\frac{1}{2}}\|^2} > 0, \quad \forall x \in \mathcal{H}\backslash\{0\}. \tag{2.7}$$

By the arbitrariness of x, (2.7) implies $\sup_{y\in\mathcal{H}\backslash\{0\}} \|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}y\|/\|y\| < \sqrt{1/\alpha}$.

(i) $\Rightarrow$ (ii). Using (2.6), for any $x \in \mathcal{H}\backslash\{0\}$, we have

$$\begin{aligned}
&\big\langle (Y^{-1} - \alpha L^*\Sigma L)x, x\big\rangle \\
&= \|Y^{-\frac{1}{2}}x\|^2 \left(1 - \alpha\frac{\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}Y^{-\frac{1}{2}}x\|^2}{\|Y^{-\frac{1}{2}}x\|^2}\right) \\
&\geq \big\langle x, Y^{-1}x\big\rangle\big(1 - \alpha\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2\big) \\
&\geq \eta\big(1 - \alpha\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2\big)\|x\|^2 \\
&= \xi\|x\|^2, \quad \forall x \in \mathcal{H}\backslash\{0\},
\end{aligned}$$

where η is the strong monotonicity constant of Y^{-1} and $\xi = \eta(1 - \alpha\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2) > 0$ from (i). Thus the proof is complete. □

Remark 2.1. Here we introduce a special case of Lemma 2.4. Let $\mathcal{H} = \mathbb{R}^n$, $\mathcal{G} = \mathbb{R}^m$, $L = (l_{ij})_{m\times n} \in \mathbb{R}^{m\times n}$, $Y = \mathrm{diag}(\tau_1, \cdots, \tau_n)$ and $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_m)$, where

$$\tau_j = \frac{t}{\sum_{i=1}^m |l_{ij}|^{2-\alpha}}, \quad \sigma_i = \frac{s}{\sum_{j=1}^n |l_{ij}|^{\alpha}}.$$

We can prove $\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2 \leq st$ for $\alpha \in [0, 2]$ and $s, t > 0$, by using [35, Lemma 2].

For further properties of monotone operators, functional analysis, and convex analysis, the readers are referred to [4, 16, 36].

3. Algorithm, convergence analysis and rates

In this section, we will prove the convergence of the split-Douglas-Rachford (SDR) algorithm under the improved convergence condition (1.14). We will also demonstrate the tightness of the condition (1.14) through a specific example. The iterative scheme for SDR is outlined in Algorithm 3.1.

Algorithm 3.1 Spilt-Douglas-Rachford Algorithm.

Let $\Sigma : \mathcal{G} \to \mathcal{G}$ and $Y : \mathcal{H} \to \mathcal{H}$ be strongly monotone self-adjoint continuous linear operators. For given (x_k, u_k), the new iterate (x_{k+1}, u_{k+1}) is generated by the following recursion:

$$\tilde{u}_k = J_{\Sigma B^{-1}}\left(u_k + \Sigma L x_k\right), \tag{3.1}$$

$$x_{k+1} = J_{YA}(x_k - YL^*\tilde{u}_k), \tag{3.2}$$

$$u_{k+1} = \tilde{u}_k - \Sigma L(x_k - x_{k+1}). \tag{3.3}$$

Remark 3.1. Using (3.1) and (3.3), we have

$$\tilde{u}_{k+1} = J_{\Sigma B^{-1}}\big(\tilde{u}_k + \Sigma L(2x_{k+1} - x_k)\big).$$

Hence, the equivalence of P-PDS and SDR is obtained.

3.1. Convergence of SDR under the improved condition (1.14)

To propose an improved convergence condition for Algorithm 3.1, we will present some useful results.

Lemma 3.1. *For any given* $(u^*, x^*) \in \mathbf{Z}$, *the sequence* $\{(u_k, \tilde{u}_k, x_k)\}_{k=1}^{\infty}$ *generated by Algorithm* 3.1 *satisfies the following inequality:*

$$\begin{aligned} &\|x_{k+1} - x^*\|_{Y^{-1}}^2 + \|u_{k+1} - u^*\|_{\Sigma^{-1}}^2 \\ \le\ &\|x_k - x^*\|_{Y^{-1}}^2 + \|u_k - u^*\|_{\Sigma^{-1}}^2 - \|x_{k+1} - x_k\|_K^2 - \|u_k - \tilde{u}_k\|_{\Sigma^{-1}}^2, \end{aligned} \tag{3.4}$$

where $K = Y^{-1} - L^*\Sigma L$.

Proof. For any given $(u^*, x^*) \in \mathbf{Z}$, using the definitions of J_{YA} and $J_{\Sigma B^{-1}}$, the monotonicity of A and B, (3.1) and (3.2), we have

$$\big\langle \tilde{u}_k - u^*, \Sigma^{-1}(u_k - \tilde{u}_k) + L(x_k - x^*)\big\rangle \ge 0, \tag{3.5}$$

$$\big\langle L^*(\tilde{u}_k - u^*) + Y^{-1}(x_{k+1} - x_k), x^* - x_{k+1}\big\rangle \ge 0. \tag{3.6}$$

Summing (3.5) and (3.6), we obtain that

$$\begin{aligned} 0 \le\ &\big\langle \tilde{u}_k - u^*, \Sigma^{-1}(u_k - \tilde{u}_k) + L(x_k - x^*)\big\rangle \\ &+ \big\langle L^*(\tilde{u}_k - u^*) + Y^{-1}(x_{k+1} - x_k), x^* - x_{k+1}\big\rangle \\ =\ &\big\langle \tilde{u}_k - u^*, \Sigma^{-1}(u_k - \tilde{u}_k)\big\rangle + \big\langle L^*(u^* - \tilde{u}_k), x_{k+1} - x_k\big\rangle \\ &+ \big\langle x_{k+1} - x^*, Y^{-1}(x_k - x_{k+1})\big\rangle \\ =\ &\frac{1}{2}\left(\|u_k - u^*\|_{\Sigma^{-1}}^2 - \|u_k - \tilde{u}_k\|_{\Sigma^{-1}}^2 - \|\tilde{u}_k - u^*\|_{\Sigma^{-1}}^2\right) - \big\langle L^*(\tilde{u}_k - u^*), x_{k+1} - x_k\big\rangle \\ &+ \frac{1}{2}\left(\|x_k - x^*\|_{Y^{-1}}^2 - \|x_{k+1} - x^*\|_{Y^{-1}}^2 - \|x_k - x_{k+1}\|_{Y^{-1}}^2\right), \end{aligned} \tag{3.7}$$

where the first equality follows from the definition of the adjoint operator and the last equality follows from (2.2). Obviously, we have

$$\begin{aligned} &\|\tilde{u}_k - u^*\|_{\Sigma^{-1}}^2 + 2\langle L^*(\tilde{u}_k - u^*), x_{k+1} - x_k\rangle + \|x_{k+1} - x_k\|_{L^*\Sigma L}^2 \\ =\ &\|\tilde{u}_k - u^* + \Sigma L(x_{k+1} - x_k)\|_{\Sigma^{-1}}^2. \end{aligned}$$

Using (3.3), we can obtain that

$$\|\tilde{u}_k - u^* + \Sigma L(x_{k+1} - x_k)\|_{\Sigma^{-1}}^2 = \|u_{k+1} - u^*\|_{\Sigma^{-1}}^2. \tag{3.8}$$

Substituting (3.7) into (3.8) and simplifying it, we can obtain (3.4), which completes the proof. ☐

Revisiting Lemma 3.1, we obtain the following result.

Lemma 3.2. *Let* $\{(u_k, x_k)\}_{k=1}^{\infty}$ *be the sequence generated by Algorithm* 3.1*, where the operator* $Y - 3L^*\Sigma L/4$ *is strongly monotone. Then, it holds that*

$$\begin{aligned} &\|x_{k+1} - x^*\|_{Y^{-1}}^2 + \|u_{k+1} - u^*\|_{\Sigma^{-1}}^2 + \frac{1}{2}\|x_k - x_{k+1}\|_P^2 \\ \leq &\|x_k - x^*\|_{Y^{-1}}^2 + \|u_k - u^*\|_{\Sigma^{-1}}^2 + \frac{1}{2}\|x_k - x_{k-1}\|_P^2 \\ &- \left(\|x_k - x_{k+1}\|_V^2 + \left\|u_k - u_{k+1} + \frac{1}{2}\Sigma L(x_{k+1} - x_k)\right\|_{\Sigma^{-1}}^2\right) \end{aligned} \tag{3.9}$$

for any given $(u^*, x^*) \in \mathbf{Z}$*, where* $P = Y^{-1} - L^*\Sigma L/2$.

Proof. According to (3.3), we have

$$\begin{aligned} \|u_k - \tilde{u}_k\|_{\Sigma^{-1}}^2 &= \|u_{k+1} - u_k + \Sigma L(x_k - x_{k+1})\|_{\Sigma^{-1}}^2 \\ &= \|u_{k+1} - u_k\|_{\Sigma^{-1}}^2 + \|x_k - x_{k+1}\|_{L^*\Sigma L}^2 \\ &\quad + 2\langle u_{k+1} - u_k, L(x_k - x_{k+1})\rangle. \end{aligned} \tag{3.10}$$

Using (3.4) and (3.10), we obtain that

$$\begin{aligned} &\|x_{k+1} - x^*\|_{Y^{-1}}^2 + \|u_{k+1} - u^*\|_{\Sigma^{-1}}^2 \\ \leq &\|x_k - x^*\|_{Y^{-1}}^2 + \|u_k - u^*\|_{\Sigma^{-1}}^2 - \left(\|x_{k+1} - x_k\|_K^2 + \|u_k - \tilde{u}_k\|_{\Sigma^{-1}}^2\right) \\ = &\|x_k - x^*\|_{Y^{-1}}^2 + \|u_k - u^*\|_{\Sigma^{-1}}^2 - \|x_{k+1} - x_k\|_{Y^{-1}}^2 - \|u_{k+1} - u_k\|_{\Sigma^{-1}}^2 \\ &+ 2\langle u_{k+1} - u_k, L(x_{k+1} - x_k)\rangle. \end{aligned} \tag{3.11}$$

By a simple manipulation, we conclude that

$$\begin{aligned} &\|u_{k+1} - u_k\|_{\Sigma^{-1}}^2 + 2\langle u_k - u_{k+1}, L(x_{k+1} - x_k)\rangle + \|x_{k+1} - x_k\|_{Y^{-1}}^2 \\ = &\left\|u_{k+1} - u_k + \frac{1}{2}\Sigma L(x_{k+1} - x_k)\right\|_{\Sigma^{-1}}^2 + \|x_{k+1} - x_k\|_{Y^{-1} - \frac{1}{4}L^*\Sigma L}^2 \\ &+ \langle u_k - u_{k+1}, L(x_{k+1} - x_k)\rangle. \end{aligned} \tag{3.12}$$

It follows from (3.2) and the definition of J_{YA} that

$$\begin{aligned} x_k - x_{k+1} - YL^*\tilde{u}_k &\in YAx_{k+1}, \\ x_{k-1} - x_k - YL^*\tilde{u}_{k-1} &\in YAx_k. \end{aligned}$$

Using the monotonicity of A and the existence of Y^{-1}, we have

$$\left\langle Y^{-1}(x_k - x_{k+1}) - L^*\tilde{u}_k - Y^{-1}(x_{k-1} - x_k) + L^*\tilde{u}_{k-1}, x_{k+1} - x_k\right\rangle \geq 0. \tag{3.13}$$

Utilizing (3.3), we obtain

$$u_{k+1} - u_k = \tilde{u}_k - \tilde{u}_{k-1} - \Sigma L(x_k - x_{k+1}) + \Sigma L(x_{k-1} - x_k). \tag{3.14}$$

Substituting (3.14) into (3.13), we get

$$\begin{aligned}&\langle L^*(u_{k+1}-u_k), x_{k+1}-x_k\rangle\\&\leq -\|x_{k+1}-x_k\|_K^2+\langle x_{k+1}-x_k, K(x_k-x_{k-1})\rangle,\end{aligned}\tag{3.15}$$

where $K = Y^{-1} - L^*\Sigma L$. Substituting (3.15) into (3.12), we have

$$\begin{aligned}&\|u_{k+1}-u_k\|_{\Sigma^{-1}}^2+2\langle u_k-u_{k+1}, L(x_{k+1}-x_k)\rangle+\|x_{k+1}-x_k\|_{Y^{-1}}^2\\ \geq& \left\|u_{k+1}-u_k+\frac{1}{2}\Sigma L(x_{k+1}-x_k)\right\|_{\Sigma^{-1}}^2+\|x_{k+1}-x_k\|_{Y^{-1}-\frac{1}{4}L^*\Sigma L}^2+\|x_{k+1}-x_k\|_K^2\\ &+\langle x_k-x_{k+1}, K(x_k-x_{k-1})\rangle.\end{aligned}\tag{3.16}$$

Because $V := Y^{-1} - 3L^*\Sigma L/4$ is strongly monotone, we can get $\langle\cdot,\cdot\rangle_V$ is an inner product. Σ is a strongly monotone self-adjoint continuous linear operator, thus $\langle\cdot,\cdot\rangle_{L^*\Sigma L}$ is a semi-inner product. By Lemma 2.2, the last term in (3.16) implies that

$$\begin{aligned}&\langle x_k-x_{k+1}, K(x_k-x_{k-1})\rangle\\ =&\left\langle x_k-x_{k+1}, \left(V-\frac{L^*\Sigma L}{4}\right)(x_k-x_{k-1})\right\rangle\\ \geq& -\|x_k-x_{k+1}\|_V\|x_{k-1}-x_k\|_V\\ &-\frac{1}{4}\|x_k-x_{k+1}\|_{L^*\Sigma L}\|x_{k-1}-x_k\|_{L^*\Sigma L}\\ \geq& -\frac{1}{2}\|x_k-x_{k+1}\|_V^2-\frac{1}{8}\|x_k-x_{k+1}\|_{L^*\Sigma L}^2\\ &-\frac{1}{2}\|x_{k-1}-x_k\|_V^2-\frac{1}{8}\|x_k-x_{k-1}\|_{L^*\Sigma L}^2\\ =& -\frac{1}{2}\|x_k-x_{k+1}\|_P^2-\frac{1}{2}\|x_k-x_{k-1}\|_P^2,\end{aligned}\tag{3.17}$$

where $P = Y^{-1} - L^*\Sigma L/2$. Substituting (3.16) and (3.17) into (3.11) and by simple manipulations, we immediately get (3.9). Hence, we complete the proof. □

Combining the above lemmas, the weak convergence of the sequence generated by Algorithm 3.1 with condition (1.14) is presented.

Theorem 3.1. *Let $(u_0, x_0) \in \mathcal{G}\times\mathcal{H}$ and $\Sigma : \mathcal{G} \to \mathcal{G}$, $Y : \mathcal{H} \to \mathcal{H}$ be strongly monotone self-adjoint continuous linear operators such that $V := Y^{-1} - 3L^*\Sigma L/4$ is strongly monotone. The sequence $\{(u_k, x_k)\}_{k=0}^{\infty}$ generated by Algorithm 3.1 satisfies the following statements:*

(i) $\sum_{k\in\mathbb{N}}\|x_{k+1}-x_k\|_V^2<+\infty,\quad \sum_{k\in\mathbb{N}}\left\|u_k-u_{k+1}+\frac{1}{2}\Sigma L(x_{k+1}-x_k)\right\|_{\Sigma^{-1}}^2<\infty.$

(ii) There exists $(u, x) \in \mathbf{Z}$ satisfying $x_k \rightharpoonup x, u_k \rightharpoonup u$ in $\mathcal{H}$ and $\mathcal{G}$, respectively.

Proof. Firstly, by rearranging the terms of (3.9), we can deduce

$$\begin{aligned}
&\|x_k - x_{k+1}\|_V^2 + \left\| u_k - u_{k+1} + \frac{1}{2}\Sigma L(x_{k+1} - x_k) \right\|_{\Sigma^{-1}}^2 \\
&\le \|x_k - x^*\|_{Y^{-1}}^2 - \|x_{k+1} - x^*\|_{Y^{-1}}^2 + \|u_k - u^*\|_{\Sigma^{-1}}^2 - \|u_{k+1} - u^*\|_{\Sigma^{-1}}^2 \\
&\quad + \frac{1}{2}\|x_k - x_{k-1}\|_P^2 - \frac{1}{2}\|x_k - x_{k+1}\|_P^2. \qquad (3.18)
\end{aligned}$$

Therefore, it follows from (3.18) that

$$\sum_{k\in\mathbb{N}} \|x_{k+1} - x_k\|_V^2 < +\infty, \quad \sum_{k\in\mathbb{N}} \left\| u_k - u_{k+1} + \frac{1}{2}\Sigma L(x_{k+1} - x_k) \right\|_{\Sigma^{-1}}^2 < \infty.$$

Then we turn to prove (ii), we define $\mathcal{X}$ as the product space of $\mathcal{G}$ and $\mathcal{H}$ equipped with the inner product

$$\langle \mathbf{x_1}, \mathbf{x_2} \rangle_{\mathcal{X}} := \langle u_1, \Sigma^{-1} u_2 \rangle + \langle x_1, Y^{-1} x_2 \rangle,$$

where $\Sigma : \mathcal{G} \to \mathcal{G}$ and $Y : \mathcal{H} \to \mathcal{H}$ are strongly monotone, self-adjoint, and continuous linear operators, $\mathbf{x}_i = (u_i, x_i) \in \mathcal{X}, u_i \in \mathcal{G}, x_i \in \mathcal{H}, i = 1, 2$. We denote the associated norm by $\|\cdot\|_{\mathcal{X}} = \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{X}}}$. We can verify that $(\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}})$ is a Hilbert space and the topology of $\mathcal{X}$ and $\mathcal{G} \oplus \mathcal{H}$ is equivalent.

For every $k \in \mathbb{N}$, denote $\mathbf{x}_k = (u_k, x_k)$. We can rewrite (3.9) as

$$\begin{aligned}
&\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_{\mathcal{X}}^2 + \frac{1}{2}\|x_k - x_{k+1}\|_P^2 \\
&\le \|\mathbf{x}_k - \mathbf{x}^*\|_{\mathcal{X}}^2 + \frac{1}{2}\|x_k - x_{k-1}\|_P^2 \\
&\quad - \left(\|x_k - x_{k+1}\|_V^2 + \left\| u_k - u_{k+1} + \frac{1}{2}\Sigma L(x_{k+1} - x_k) \right\|_{\Sigma^{-1}}^2 \right), \qquad (3.19)
\end{aligned}$$

where $\mathbf{x}^* := (u^*, x^*)$ is an arbitrarily given element in $\mathbf{Z}$.

Because $V = Y^{-1} - 3L^*\Sigma L/4$ is strongly monotone, we obtain that $\|\cdot\|_V$ is a norm in $\mathcal{H}$. Using the conclusion of (i) and (3.3), we have

$$x_k - x_{k+1} \to 0, \quad u_k - u_{k+1} \to 0, \quad u_k - \tilde{u}_k \to 0, \quad k \to \infty \qquad (3.20)$$

in $\mathcal{H}$ and $\mathcal{G}$, respectively. We know that the sequence $\{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_{\mathcal{X}}^2 + \|x_k - x_{k+1}\|_P^2/2\}_{k=1}^\infty$ is convergent from (3.19). Because $V := Y^{-1} - 3L^*\Sigma L/4$ is strongly monotone, we get $\|\Sigma^{\frac{1}{2}} L Y^{\frac{1}{2}}\|^2 < 4/3 < 2$ and $P = Y^{-1} - L^*\Sigma L/2$ is strongly monotone by utilizing Lemma 2.4. Further, we can yield that $\|\cdot\|_P$ is a norm in $\mathcal{H}$ and that $\{\|x_k - x_{k+1}\|_P^2\}_{k=1}^\infty$ converges to 0. Thus, we immediately have that $\{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_{\mathcal{X}}^2\}_{k=1}^\infty$ converges for any given (u^*, x^*) in $\mathbf{Z}$. Hence, $\{\mathbf{x}_k\}_{k=1}^\infty$ is a bounded sequence in $\mathcal{X}$. Furthermore, we can claim that the sequence $\{\mathbf{x}_k\}_{k=1}^\infty$ is also bounded in $\mathcal{G} \oplus \mathcal{H}$ from the equivalence of $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{G}\oplus\mathcal{H}}$.

Next, let $(\tilde{u},\tilde{x}) \in \mathcal{G}\oplus\mathcal{H}$ be a weak sequential cluster point of $\{\mathbf{x}_k\}_{k=1}^{\infty}$. Without loss of generality, we may assume $\mathbf{x}_{k_j} \rightharpoonup (\tilde{u},\tilde{x})$, $j\to\infty$. It is clear from the definition of the inner product of $\mathcal{G}\oplus\mathcal{H}$, (3.4) and (3.20) that

$$\tilde{u}_{k_j} \rightharpoonup \tilde{u}, \quad x_{k_j+1} \rightharpoonup \tilde{x}, \quad j\to\infty. \tag{3.21}$$

Utilizing the definitions of J_{YA} and $J_{\Sigma B^{-1}}$ and the iterative scheme of Algorithm 3.1, we can get

$$0 \in B^{-1}\tilde{u}_k - Lx_{k+1} + \Sigma^{-1}(\tilde{u}_k - u_k) + L(x_{k+1}-x_k), \tag{3.22}$$

$$0 \in Ax_{k+1} + L\tilde{u}_k + Y^{-1}(x_{k+1}-x_k). \tag{3.23}$$

By using the definition of $\mathbf{M}+\mathbf{S}$, we get

$$\big(\Sigma^{-1}(u_{k_j}-\tilde{u}_{k_j}) + L(x_{k_j}-x_{k_j+1}), Y^{-1}(x_{k_j}-x_{k_j+1})\big) \in (\mathbf{M}+\mathbf{S})(\tilde{u}_{k_j}, x_{k_j+1}). \tag{3.24}$$

We conclude from the continuity of Σ^{-1}, Y^{-1} and L, the maximal monotonicity of $\mathbf{M}+\mathbf{S}$ in $\mathcal{G}\oplus\mathcal{H}$ and (3.21) that

$$(0,0) \in (\mathbf{M}+\mathbf{S})(\tilde{u},\tilde{x}). \tag{3.25}$$

Therefore, we deduce from the convergence of $\{\|\mathbf{x}_{k+1}-\mathbf{x}^*\|_{\mathcal{X}}^2\}_{k=1}^{\infty}$, (3.25), Lemma 2.1 and the definition of the inner product of $\mathcal{G}\oplus\mathcal{H}$ that there exists $(u,x)\in\mathbf{Z}$ such that $u_k \rightharpoonup u, x_k \rightharpoonup x$ in $\mathcal{G}$ and $\mathcal{H}$, respectively. Thus we have obtained the result of the theorem. □

3.2. An example of the condition (1.14) tightness

Next, we claim that condition (1.14) is tight to ensure the convergence of the sequence generated by Algorithm 3.1.

Example 3.1. We use Algorithm 3.1 with $\Sigma = Id$ and $Y = 4Id/3$ for solving the following convex-concave saddle point problem in real lines $\mathbb{R}$ [26],

$$\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} \big\{x + I_{\mathbb{R}_+}(x) - xy + y - I_{\mathbb{R}_+}(y)\big\},$$

which has a unique saddle point $(1,1)$. In this problem, we can verify that $A = 1+\mathcal{N}_{\mathbb{R}_+}$, $B^{-1} = \mathcal{N}_{\mathbb{R}_+} - 1$ and $L = -Id$. In this setting, we choose $\Sigma = Id$, $Y = 4Id/3$, thus the iterative scheme of Algorithm 3.1 can be rewritten as

$$\begin{cases} \tilde{u}_k = \max\{u_k - x_k + 1, 0\}, \\ x_{k+1} = \max\left\{x_k + \dfrac{4}{3}(\tilde{u}_k - 1), 0\right\}, \\ u_{k+1} = \tilde{u}_k + (x_k - x_{k+1}). \end{cases}$$

Choosing the initial point $(u_0,x_0) = (7/8, 5/4)$, we obtain $(u_1,x_1) = (9/8, 3/4)$, (u_2,x_2) $= (7/8,5/4) = (u_0,x_0)$. Thus, the sequence $\{(u_k,x_k)\}$ generated by Algorithm 3.1 with $\|\Sigma^{\frac{1}{2}}LY^{\frac{1}{2}}\|^2 = 4/3$ does not converge. Combining Remark 3.1 and Example 3.1, we can conclude that the condition (1.14) is tight for P-PDS.

3.3. Convergence rate of SDR

In the final part of Section 3, we will examine the convergence rate of SDR. We define the fixed point residual of SDR as

$$r_k := \|u_{k+1} - u_k\|^2_{\Sigma^{-1}} + \|x_{k+1} - x_k\|^2_{Y^{-1}}.$$

If $r_k = 0$, then it can be inferred from (3.22) and (3.23) that (u_k, x_k) is a solution point of problem (1.1). Hence, we can use r_k as an indicator to evaluate the convergence speed of SDR.

Lemma 3.3 ([6, Theorem 3.3]). *Let the sequence* $\{(u_k, x_k)\}_{k=1}^{\infty}$ *be generated by Algorithm* 3.1 *with* (1.13)*, it holds*

$$\begin{aligned} &\|x_{k+1} - x^*\|^2_{Y^{-1}} + \|u_{k+1} - u^*\|^2_{\Sigma^{-1}} + \|x_{k+1} - x_k\|^2_K \\ \le\, &\|x_k - x^*\|^2_{Y^{-1}} + \|u_k - u^*\|^2_{\Sigma^{-1}} + \|x_k - x_{k-1}\|^2_K \\ &- \|x_{k+1} - x_k\|^2_{Y^{-1}} - \|u_{k+1} - u_k\|^2_{\Sigma^{-1}} \end{aligned} \tag{3.26}$$

for any given $(u^*, x^*) \in \mathbf{Z}$*, where* $K = Y^{-1} - L^*\Sigma L$*.*

To simplify the proof of the convergence rate of SDR, we introduce the following notations. For the given Σ, Y, and L in Algorithm 3.1, let $Q, H, M : \mathcal{G} \oplus \mathcal{H} \to \mathcal{G} \oplus \mathcal{H}$ be operators that satisfy the following conditions:

$$Q(u, x) = \big(\Sigma^{-1}u + Lx, Y^{-1}x\big), \tag{3.27}$$

$$H(u, x) = \big(\Sigma^{-1}u, Y^{-1}x\big), \tag{3.28}$$

$$M(u, x) = \big(u + \Sigma Lx, x\big) \tag{3.29}$$

for all $(u, x) \in \mathcal{G} \oplus \mathcal{H}$.

From the definitions of Q, H and M and the continuity and linearity of Σ, Y, L, and L^*, it can be inferred that Q, H and M are continuous linear operators. Further, we can get the following results.

Lemma 3.4. *Suppose* Q, H*, and* M *are defined by* (3.27)*,* (3.28)*, and* (3.29)*, respectively. Then we have* $Q = HM$ *and* $(Q^* + Q - M^*HM)(u, x) = (\Sigma^{-1}u, (Y^{-1} - L^*\Sigma L)x)$ *for any* $(u, x) \in \mathcal{G} \oplus \mathcal{H}$*.*

Proof. For any $(u, x) \in \mathcal{G} \oplus \mathcal{H}$,

$$\begin{aligned} HM(u, x) &= H\big(u + \Sigma Lx, x\big) \\ &= \big(\Sigma^{-1}(u + \Sigma Lx), Y^{-1}x\big) \\ &= \big(\Sigma^{-1}u + Lx, Y^{-1}x\big) \\ &= Q(u, x). \end{aligned}$$

For arbitrary pairs (u_1, x_1) and $(u_2, x_2) \in \mathcal{G} \oplus \mathcal{H}$, utilizing the definitions of Q and inner product in $\mathcal{G} \oplus \mathcal{H}$, we can obtain that

$$\begin{aligned}
&\big\langle Q(u_1, x_1), (u_2, x_2)\big\rangle \\
=& \big\langle \Sigma^{-1}u_1, u_2\big\rangle + \big\langle Lx_1, u_2\big\rangle + \big\langle Y^{-1}x_1, x_2\big\rangle \\
=& \big\langle u_1, \Sigma^{-1}u_2\big\rangle + \big\langle x_1, L^*u_2\big\rangle + \big\langle x_1, Y^{-1}x_2\big\rangle \\
=& \big\langle (u_1, x_1), (\Sigma^{-1}u_2, Y^{-1}x_2 + L^*u_2)\big\rangle.
\end{aligned}$$

Thus we obtain $Q^*(u, x) = (\Sigma^{-1}u_2, Y^{-1}x_2 + L^*u_2)$. Similarly, we can verify $M^*(u, x) = (u, x + L^*\Sigma u)$. Hence, we can yield

$$\begin{aligned}
&(Q^* + Q - M^*HM)(u, x) \\
=& \big(2\Sigma^{-1}u + Lx, 2Y^{-1}x + L^*u\big) \\
&- \big(\Sigma^{-1}u + Lx, Y^{-1}x + L^*u + L^*\Sigma Lx\big) \\
=& \big(\Sigma^{-1}u, Y^{-1}x - L^*\Sigma Lx\big).
\end{aligned}$$

Thus the proof is complete. □

Lemma 3.5. *Let* $\{(u_k, \tilde{u}_k, x_k)\}_{k=1}^{\infty}$ *be the sequence generated by Algorithm* 3.1*, then it holds*

$$(0, 0) \in (\mathbf{M} + \mathbf{S})(\tilde{u}_k, x_{k+1}) + Q(\tilde{u}_k - u_k, x_{k+1} - x_k), \tag{3.30}$$

$$(u_{k+1}, x_{k+1}) = (u_k, x_k) + M(\tilde{u}_k - u_k, x_{k+1} - x_k). \tag{3.31}$$

Proof. Based on (3.1)-(3.3), and the definition of the resolvent of the maximal monotone operator, we can obtain directly

$$\begin{aligned}
&0 \in B^{-1}\tilde{u}_k - Lx_{k+1} + \Sigma^{-1}(\tilde{u}_k - u_k) + L(x_{k+1} - x_k), \\
&0 \in Ax_{k+1} + L\tilde{u}_k + Y^{-1}(x_{k+1} - x_k).
\end{aligned}$$

By using the definitions of $\mathbf{M} + \mathbf{S}, Q$, and M, we complete the proof of (3.30) and (3.31). □

Lemma 3.6. *Let* $\{(u_k, \tilde{u}_k, x_k)\}_{k=1}^{\infty}$ *be the sequence generated by Algorithm* 3.1*, we have the following inequality:*

$$\begin{aligned}
&\|u_{k+2} - u_{k+1}\|_{\Sigma^{-1}}^2 + \|x_{k+2} - x_{k+1}\|_{Y^{-1}}^2 \\
\le& \|u_{k+1} - u_k\|_{\Sigma^{-1}}^2 + \|x_{k+1} - x_k\|_{Y^{-1}}^2 - \|u_k - \tilde{u}_k - (u_{k+1} - \tilde{u}_{k+1})\|_{\Sigma^{-1}}^2 \\
&- \|x_k - x_{k+1} - (x_{k+1} - x_{k+2})\|_K^2,
\end{aligned} \tag{3.32}$$

where $K := Y^{-1} - L^*\Sigma L$.

Proof. Throughout the proof of this lemma, we use $\langle\cdot,\cdot\rangle$ for $\langle\cdot,\cdot\rangle_{\mathcal{G}\oplus\mathcal{H}}$. We note that (3.30) holds for $k:=k+1$, thus, we have

$$(0,0)\in(\mathbf{M}+\mathbf{S})(\tilde{u}_{k+1},x_{k+2})+Q(\tilde{u}_{k+1}-u_{k+1},x_{k+2}-x_{k+1}).$$

Using the monotonicity of $\mathbf{M}+\mathbf{S}$ in $\mathcal{G}\oplus\mathcal{H}$, we deduce

$$\begin{aligned}&\big\langle(\tilde{u}_k-\tilde{u}_{k+1},\tilde{x}_k-\tilde{x}_{k+1}),Q(\tilde{u}_{k+1}-u_{k+1},\tilde{x}_{k+1}-x_{k+1})\big\rangle\\ \geq\ &\big\langle(\tilde{u}_k-\tilde{u}_{k+1},\tilde{x}_k-\tilde{x}_{k+1}),Q(\tilde{u}_k-u_k,\tilde{x}_k-x_k)\big\rangle.\end{aligned}\tag{3.33}$$

We note that for all $(u,x)\in\mathcal{G}\oplus\mathcal{H}$, the equality

$$\begin{aligned}&\langle(u,x),Q(u,x)\rangle=\langle Q^*(u,x),(u,x)\rangle\\ =\ &\frac{1}{2}\langle(Q^*+Q)(u,x),(u,x)\rangle=\frac{1}{2}\|(u,x)\|^2_{Q^*+Q}\end{aligned}\tag{3.34}$$

always holds. Utilizing (3.33) and (3.34), we yield

$$\begin{aligned}&\big\langle(u_k-u_{k+1},x_k-x_{k+1}),Q(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\big\rangle\\ =\ &\frac{1}{2}\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{Q+Q^*}\\ &+\big\langle(\tilde{u}_k-\tilde{u}_{k+1},\tilde{x}_k-\tilde{x}_{k+1}),Q(\tilde{u}_{k+1}-u_{k+1},\tilde{x}_{k+1}-x_{k+1})\big\rangle\\ &-\big\langle(\tilde{u}_k-\tilde{u}_{k+1},\tilde{x}_k-\tilde{x}_{k+1}),Q(\tilde{u}_k-u_k,\tilde{x}_k-x_k)\big\rangle\\ \geq\ &\frac{1}{2}\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{Q+Q^*}.\end{aligned}\tag{3.35}$$

Substituting (3.31) in the left-hand side of (3.35), we obtain

$$\begin{aligned}&\big\langle M(u_k-\tilde{u}_k,x_k-\tilde{x}_k),Q(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\big\rangle\\ \geq\ &\frac{1}{2}\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{Q+Q^*}.\end{aligned}\tag{3.36}$$

Noting $Q=HM$, let

$$\begin{aligned}x&=(u_k-\tilde{u}_k,x_k-\tilde{x}_k),\\ y&=(0,0),\\ z&=(u_k-\tilde{u}_k,x_k-\tilde{x}_k),\\ w&=(u_{k+1}-\tilde{u}_{k+1},x_{k+1}-\tilde{x}_{k+1}),\end{aligned}$$

$S=M^*HM$ in (2.2), then we have

$$\begin{aligned}&2\big\langle M(u_k-\tilde{u}_k,x_k-\tilde{x}_k),Q(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\big\rangle\\ =\ &\|(u_k-\tilde{u}_k,x_k-\tilde{x}_k)\|^2_{M^*HM}-\|(u_{k+1}-\tilde{u}_{k+1},x_{k+1}-\tilde{x}_{k+1})\|^2_{M^*HM}\\ &+\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{M^*HM}.\end{aligned}\tag{3.37}$$

Combining (3.36) and (3.37), we can deduce that

$$\begin{aligned}&\|(u_k-\tilde{u}_k,x_k-\tilde{x}_k)\|^2_{M^*HM}-\|(u_{k+1}-\tilde{u}_{k+1},x_{k+1}-\tilde{x}_{k+1})\|^2_{M^*HM}\\&\geq\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{Q+Q^*-M^*HM}.\end{aligned}\tag{3.38}$$

Substituting (3.31) in (3.38), we have

$$\begin{aligned}&\|(u_{k+1}-u_k,x_{k+1}-x_k)\|^2_H-\|(u_{k+2}-u_{k+1},x_{k+2}-x_{k+1})\|^2_H\\&\geq\|(u_k-\tilde{u}_k-u_{k+1}+\tilde{u}_{k+1},x_k-\tilde{x}_k-x_{k+1}+\tilde{x}_{k+1})\|^2_{Q+Q^*-M^*HM}.\end{aligned}$$

Using the definition of the inner product of $\mathcal{G}\oplus\mathcal{H}$, we directly obtain the conclusion of this lemma. □

Now, we present the sublinear convergence rate of SDR from the above lemmas.

Theorem 3.2. *Let* $\{(u_k,x_k)\}_{k=1}^{\infty}$ *be the sequence generated by Algorithm* 3.1 *with* (1.13)*, then we have*

$$\|u_{k+1}-u_k\|^2_{\Sigma^{-1}}+\|x_{k+1}-x_k\|^2_{Y^{-1}}=o(1/k).\tag{3.39}$$

Proof. Denote

$$a_k:=\|u_{k+1}-u_k\|^2_{\Sigma^{-1}}+\|x_{k+1}-x_k\|^2_{Y^{-1}}.$$

Since $Y^{-1}-L^*\Sigma L$ is monotone and (3.32) holds, we have that the sequence $\{a_k\}_{k=1}^{\infty}$ is decreasing. Further, we can rewrite (3.26) as

$$a_k\leq\Delta_k-\Delta_{k+1},\quad\forall k\in\mathbb{N}_+,\tag{3.40}$$

where

$$\Delta_k=\|x_k-x^*\|^2_{Y^{-1}}+\|u_k-u^*\|^2_{\Sigma^{-1}}+\|x_k-x_{k-1}\|^2_K.$$

For any given positive integer n, summing k from 1 to n on both sides of the inequality, we have

$$\sum_{k=1}^{n}a_k\leq\Delta_1<\infty.\tag{3.41}$$

For any real number r, the largest integer that is not greater than the number r is denoted by $\lfloor r\rfloor$. It can be seen from the decreasing sequence $\{a_k\}$ and (3.41) that

$$0\leq\frac{n}{2}a_n\leq\sum_{k=\lfloor\frac{n}{2}\rfloor}^{n}a_k\to 0,\quad n\to\infty,\tag{3.42}$$

which implies the conclusion of this theorem. □

Remark 3.2. In finite-dimensional Hilbert spaces, Ma *et al.* [34] established the sublinear convergence rate, in terms of fixed point residuals, of P-PDHG under the condition (1.10). Therefore, under the condition (1.14), which is weaker than (1.10), Theorem 3.2 yields the same result regarding the sublinear convergence rate in general Hilbert spaces.

4. A novel preconditioned ADMM and its tight convergence conditions and rates

4.1. A special case: A linear operator is the negative identity operator

Initially, we investigate the numerical algorithm for solving the following convex optimization problem in the Hilbert space $\mathcal{H}$:

$$\min_{x \in \mathcal{H}} f(x) + h(Lx). \tag{4.1}$$

Assumption 4.1. For problem (4.1), we make the following two assumptions:

(A1) $0 \in \operatorname{sri}(L \operatorname{dom} f - \operatorname{dom} h)$.

(A2) The solution set of problem (4.1) is nonempty.

Remark 4.1. Problem (4.1) can be rewritten as

$$\min_{x \in \mathcal{H}} \sup_{u \in \mathcal{G}} \big\{ f(x) + \langle u, Lx \rangle - h^*(u) := \Phi(x, u) \big\}. \tag{4.2}$$

Under Assumption 4.1, we get the solution set of dual problem of (4.1) is nonempty and the duality gap is zero by using Theorem 2.1. From [4, Theorem 19.1], we obtain that

$$\begin{cases} 0 \in \partial f(x) + L^* u, \\ 0 \in \partial h^*(u) - Lx, \end{cases}$$

where x is a solution to the primal problem and u is a solution to the dual problem. Thus, the set of saddle points of Φ is nonempty. Further, if (x, u) is a saddle point of Φ, then x is a solution to the problem (4.1) and u is a solution to the dual problem of (4.1).

The problem (4.2) is a special case of (1.1), by setting $A = \partial f$ and $B = \partial h$. Hence, under assumptions (A1) and (A2), we have

$$\mathbf{Z} := \big\{ (u, x) \mid 0 \in \partial h^*(u) - Lx, 0 \in \partial f(x) + L^* u \big\} \neq \varnothing$$

and Algorithm 3.1 reduces to the following iterative scheme:

$$\tilde{u}_k = \operatorname{prox}_{h^*}^{\Sigma^{-1}} \left(u_k + \Sigma L x_k \right), \tag{4.3}$$

$$x_{k+1} = \operatorname{prox}_f^{Y^{-1}} (x_k - Y L^* \tilde{u}_k), \tag{4.4}$$

$$u_{k+1} = \tilde{u}_k - \Sigma L (x_k - x_{k+1}). \tag{4.5}$$

The following result reveals the relationship between the solutions to the two equivalent forms of the problem (4.1).

Lemma 4.1. *Consider the following two convex-concave saddle point problems:*

$$\min_{x\in\mathcal{H}}\max_{u\in\mathcal{G}}\big\{f(x)+\langle u,Lx\rangle-h^*(u):=\Phi(x,u)\big\},\tag{4.6}$$

$$\min_{x\in\mathcal{H},y\in\mathcal{G}}\max_{\lambda\in\mathcal{G}}\big\{f(x)+h(y)+\langle\Sigma^{\frac{1}{2}}\lambda,Lx-y\rangle:=\mathcal{L}(x,y,\lambda)\big\}.\tag{4.7}$$

Suppose that Assumption 4.1 *holds, then the set of saddle points of* Φ *is nonempty. For an arbitrarily given* (x,u)*, which is a saddle point of* Φ*, there exists* $\lambda\in\mathcal{G}$ *such that* (x,Lx,λ) *is a saddle point of* $\mathcal{L}$ *and* $u=\Sigma^{\frac{1}{2}}\lambda$*.*

Proof. Based on the discussion in Remark 4.1, we have that the set of saddle points of Φ is nonempty. Clearly, saddle points of $\mathcal{L}$ are also solutions to the following equations:

$$\begin{cases}0\in\partial f(x)+L^*\Sigma^{\frac{1}{2}}\lambda,\\ 0\in\partial h(y)-\Sigma^{\frac{1}{2}}\lambda,\\ y=Lx.\end{cases}$$

Letting (x,u) be a saddle point of Φ, we obtain that

$$\begin{cases}0\in\partial f(x)+L^*u,\\ 0\in\partial h^*(u)-Lx.\end{cases}$$

Setting $y:=Lx$ and $\lambda:=\Sigma^{-\frac{1}{2}}u$, it is clear that (x,y,λ) is a saddle point of $\mathcal{L}$ from $\partial h=(\partial h^*)^{-1}$. Thus we complete the proof. ☐

Remark 4.2. We observe that

$$\mathcal{L}(x,y,\lambda):=f(x)+h(y)+\big\langle\Sigma^{\frac{1}{2}}\lambda,Lx-y\big\rangle$$

is the Lagrangian function of the following problem:

$$\begin{aligned}\min_{x\in\mathcal{H},y\in\mathcal{G}}\quad & f(x)+h(y)\\ \text{s.t.}\quad & \Sigma^{\frac{1}{2}}Lx-\Sigma^{\frac{1}{2}}y=0.\end{aligned}\tag{4.8}$$

Since $\Sigma^{\frac{1}{2}}$ is a bijection, the equivalence of (4.8) and (4.1) can be obtained. Further, if (x,y,λ) is a saddle point of $\mathcal{L}$, then x is a solution to (4.1) and $y=Lx$.

The preconditioned alternating direction method of multipliers (P-ADMM) for solving (4.1) is given in Algorithm 4.1.

Algorithm 4.1 The Preconditioned Alternating Direction Method of Multipliers (P-ADMM).

Let $\Sigma:\mathcal{G}\to\mathcal{G}$ and $Y:\mathcal{H}\to\mathcal{H}$ be strongly monotone self-adjoint continuous linear operators. For given (x_k,λ_k,y_k), the new iterate $(x_{k+1},\lambda_{k+1},y_{k+1})$ is generated by the following recursion:

$$x_{k+1} = \arg\min_{x\in\mathcal{H}} \left\{ f(x) + \frac{1}{2}\left\|Lx - y_k + \Sigma^{-\frac{1}{2}}\lambda_k\right\|_\Sigma^2 + \frac{1}{2}\|x - x_k\|_{Y^{-1}-L^*\Sigma L}^2 \right\}, \quad (4.9)$$

$$\lambda_{k+1} = \lambda_k + \Sigma^{\frac{1}{2}}(Lx_{k+1} - y_k), \quad (4.10)$$

$$y_{k+1} = \arg\min_{y\in\mathcal{K}} \left\{ h(y) + \frac{1}{2}\left\|Lx_{k+1} - y + \Sigma^{-\frac{1}{2}}\lambda_{k+1}\right\|_\Sigma^2 \right\}. \quad (4.11)$$

Remark 4.3. The sequence $\{x_k\}_{k=1}^\infty$ generated by P-ADMM is well defined, as we can verify the objective function in the subproblem (4.9) is strongly convex.

Lemma 4.2 ([34, Lemma 3.1])**.** *Let $\Sigma : \mathcal{G} \to \mathcal{G}$ and $Y : \mathcal{H} \to \mathcal{H}$ be strongly monotone self-adjoint, then following statements hold:*

(i) Let $\{(x_k, \tilde{u}_k)\}_{k=1}^\infty$ be the sequence satisfying by (4.3)-(4.5) *and denote*

$$y_{k+1} = \Sigma^{-1}(\tilde{u}_k - \tilde{u}_{k+1}) + L(2x_{k+1} - x_k), \quad (4.12)$$

$$\lambda_{k+1} = \Sigma^{-\frac{1}{2}}\tilde{u}_k + \Sigma^{\frac{1}{2}}L(x_{k+1} - x_k). \quad (4.13)$$

Then the sequence $\{(x_k, y_k, \lambda_k)\}_{k=1}^\infty$ is generated by Algorithm 4.1.

(ii) Let $\{(x_k, y_k, \lambda_k)\}_{k=1}^\infty$ be the sequence generated by Algorithm 4.1 *and set*

$$\tilde{u}_k = \Sigma^{\frac{1}{2}}\lambda_k + \Sigma(Lx_k - y_k). \quad (4.14)$$

Then the sequence $\{(x_k, \tilde{u}_k)\}_{k=1}^\infty$ satisfies (4.3)-(4.5).

By utilizing Lemmas 4.1 and 4.2, we can directly obtain the weak convergence of P-ADMM.

Theorem 4.1. *Suppose Assumption* 4.1 *holds. Let $\{(x_k, y_k, \lambda_k)\}_{k=1}^\infty$ be the sequence generated by Algorithm* 4.1*, where the operator $Y^{-1} - 3L^*\Sigma L/4$ is strongly monotone. Then, the following statements hold:*

(i) There exists a saddle point (x^*, Lx^*, λ^*) of $\mathcal{L}$, such that $x_k \rightharpoonup x^*, y_k \rightharpoonup Lx^*, \lambda_k \rightharpoonup \lambda^*$, where

$$\mathcal{L}(x, y, \lambda) = f(x) + h(y) + \left\langle \Sigma^{\frac{1}{2}}\lambda, Lx - y \right\rangle.$$

(ii) If $Y^{-1} - L^*\Sigma L$ is monotone, then

$$\|x_{k+1} - x_k\|_{Y^{-1}}^2 + \|Lx_{k+1} - y_k\|_\Sigma^2 = o(1/k).$$

(iii) The convergence condition (1.14) *is tight.*

Proof. (i) Let $\{(x_k, y_k, \lambda_k)\}$ be the sequence generated by Algorithm 4.1. By using the conclusions of Theorem 3.1 and Lemma 4.2, there exists a point (x^*, u^*) which is the solution to (4.6) such that $x_k \rightharpoonup x^*$ and $\tilde{u}_k := \Sigma^{\frac{1}{2}}\lambda_k + \Sigma(Lx_k - y_k) \rightharpoonup u^*$. Using Lemma 4.1, for (x^*, u^*), there exists $\lambda^* \in \mathcal{G}$ such that (x^*, Lx^*, λ^*) is a saddle point of $\mathcal{L}$ and $u^* = \Sigma^{\frac{1}{2}}\lambda^*$. Hence, we yield that $\Sigma^{\frac{1}{2}}\lambda_k + \Sigma(Lx_k - y_k) \rightharpoonup u^* = \Sigma^{\frac{1}{2}}\lambda^*$. Combining (4.10), we obtain

$$\Sigma^{\frac{1}{2}}\lambda_{k+1} + \Sigma L(x_k - x_{k+1}) \rightharpoonup \Sigma^{\frac{1}{2}}\lambda^*. \tag{4.15}$$

Based on Theorem 3.1, we get $x_k - x_{k+1} \to 0$. It follows from (4.15) and $\Sigma^{\frac{1}{2}}$ is a bijection that $\lambda_k \rightharpoonup \lambda^*$. Since $\lambda_{k+1} = \lambda_k + \Sigma^{\frac{1}{2}}(Lx_{k+1} - y_k)$, we get

$$y_k = Lx_k + \Sigma^{-\frac{1}{2}}(\lambda_k - \lambda_{k+1}) \rightharpoonup Lx^*.$$

(ii) Let $\{(x_k, y_k, \lambda_k)\}_{k=1}^{\infty}$ be the sequence generated by Algorithm 4.2. Using Theorem 4.2, we have $\{(x_k, \tilde{u}_k)\}_{k=1}^{\infty}$ are generated by (4.3)-(4.5), where $\tilde{u}_k := \Sigma^{\frac{1}{2}}\lambda_k + \Sigma(Lx_k - y_k)$. Combining (4.5) and (4.10), we get

$$\begin{aligned} u_{k+1} &= \tilde{u}_k + \Sigma L(x_{k+1} - x_k) \\ &= \Sigma^{\frac{1}{2}}\lambda_k + \Sigma(Lx_{k+1} - y_k) \\ &= \Sigma^{\frac{1}{2}}\lambda_{k+1}. \end{aligned}$$

Thus,

$$\|u_{k+1} - u_k\|_{\Sigma^{-1}}^2 = \|Lx_{k+1} - y_k\|_{\Sigma}^2.$$

Using Theorem 3.2, we obtain immediately

$$\|x_{k+1} - x_k\|_{Y^{-1}}^2 + \|Lx_{k+1} - y_k\|_{\Sigma}^2 = o(1/k).$$

(iii) Considering the following optimization problem in [26]:

$$\begin{aligned} &\min\ x \\ &\text{s.t.}\ \ x \geq 0,\ \ x \geq 1. \end{aligned} \tag{4.16}$$

We know that the problem (4.16) has and only has a solution $x^* = 1$, and the dual problem of (4.16) is

$$\begin{aligned} &\max \quad \lambda \\ &\text{s.t.}\ \ 0 \leq \lambda \leq 1. \end{aligned} \tag{4.17}$$

It is clear that the problem (4.17) has a unique solution $\lambda^* = 1$. We can show that the corresponding saddle point problem of (4.16) is the convex-concave saddle point problem in Example 3.1

$$\min_{x \in \mathbb{R}} \max_{y \in \mathbb{R}} \big\{x + I_{\mathbb{R}_+}(x) - xy + y - I_{\mathbb{R}_+}(y)\big\}.$$

Hence, using Example 3.1 and Theorem 4.2, we can get that the condition (1.14) for ensuring convergence of P-ADMM can not be improved. The proof is complete. ☐

4.2. The general case: No requirements for two linear operators

Next, we deal with more general formulations which involve two linear operators and three real Hilbert spaces. Let $\mathcal{G}, \mathcal{H}, \mathcal{K}$ be real Hilbert spaces, $f \in \Gamma_0(\mathcal{H})$, and $g \in \Gamma_0(\mathcal{K})$. Let $L : \mathcal{H} \to \mathcal{G}$ and $J : \mathcal{K} \to \mathcal{G}$ be linear continuous operators. Consider the following convex optimization problem:

$$\begin{aligned} \min_{x\in\mathcal{H}, w\in\mathcal{K}} \quad & f(x) + g(w) \\ \text{s.t.} \quad & Lx + Jw = b. \end{aligned} \tag{4.18}$$

Assumption 4.2. For problem (4.18), we make the following assumptions:

(A1′) $b \in \operatorname{sri}(L \operatorname{dom} f + J \operatorname{dom} g)$.

(A2) The solution set to the problem (4.18) is nonempty.

(A3) $0 \in \operatorname{sri}(\operatorname{dom} g^* - \operatorname{ran} J^*)$.

Remark 4.4. Set $\mathcal{K} = \mathcal{G}$, $J = -Id$, $b = 0$, then (4.1) is a special case of problem (4.18). In this case, (A1′) reduces to (A1) and (A3) always holds.

Similar to the discussion in the Lemma 4.1, we can get the following result.

Lemma 4.3. *Consider the following two convex-concave saddle point problems:*

$$\min_{x\in\mathcal{H}} \max_{u\in\mathcal{G}} \big\{ f(x) + \langle u, Lx - b\rangle - g^*(-J^*u) := \tilde{\Phi}(x,u) \big\}, \tag{4.19}$$

$$\min_{x\in\mathcal{H}, w\in\mathcal{K}} \max_{\lambda\in\mathcal{G}} \big\{ f(x) + g(w) + \langle \Sigma^{\frac{1}{2}}\lambda, Lx + Jw - b\rangle := \tilde{\mathcal{L}}(x,w,\lambda) \big\}. \tag{4.20}$$

Suppose that Assumption 4.2 *holds, then the set of saddle points of* $\tilde{\Phi}$ *is nonempty. For an arbitrarily given* (x,u) *which is a saddle point of* $\tilde{\Phi}$*, there exist* $w \in \mathcal{K}$ *and* $\lambda \in \mathcal{G}$ *such that* (x,w,λ) *is a saddle point of* $\tilde{\mathcal{L}}$*,* $Lx + Jw = b$*, and* $u = \Sigma^{\frac{1}{2}}\lambda$*.*

Proof. Note that problem (4.18) can be rewritten as

$$\min_{x\in\mathcal{H}} \Big\{ f(x) + \min_{-Jw = Lx - b} g(w) \Big\}. \tag{4.21}$$

We set

$$h := \big((-J) \triangleright g\big)(\cdot - b) : y \;\mapsto \min_{-Jw = y - b} g(w),$$

hence (4.21) can be equivalently written as

$$\min_{x\in\mathcal{H}} \big\{ f(x) + h(Lx) \big\}. \tag{4.22}$$

Using [4, Corollary 15.28], we have $(-J) \triangleright g = (g^* \circ (-J)^*)^*$. Further, $h = ((-J) \triangleright g)(\cdot - b) \in \Gamma_0(\mathcal{G})$. (A1′) $\Leftrightarrow 0 \in \operatorname{sri}(L \operatorname{dom} f + J \operatorname{dom} g - b) \Leftrightarrow 0 \in \operatorname{sri}(L \operatorname{dom} f - ((-J) \operatorname{dom} g +$

$b)) \Leftrightarrow 0 \in \operatorname{sri}(L\operatorname{dom} f - \operatorname{dom} h)$. By using Lemma 4.1 for problem (4.21), we can get $\tilde{\Phi}$ has saddle points. Saddle points of $\tilde{\mathcal{L}}$ are also solutions to the following inclusion problem:

$$\begin{cases} 0 \in \partial f(x) + L^*\Sigma^{\frac{1}{2}}\lambda, \\ 0 \in \partial g(w) + J^*\Sigma^{\frac{1}{2}}\lambda, \\ Lx + Jw = b. \end{cases} \tag{4.23}$$

Letting (x, u) be a saddle point of $\tilde{\Phi}$, we obtain

$$\begin{cases} 0 \in \partial f(x) + L^*u, \\ 0 \in \partial h^*(u) - Lx + b = -J\partial g^*(-J^*u) - Lx + b. \end{cases}$$

Then there exists $w \in \partial g^*(-J^*u)$ such that $Lx + Jw = b$. Further, we get

$$\begin{cases} 0 \in \partial f(x) + L^*u, \\ 0 \in \partial g(w) + J^*u, \\ Lx + Jw = b. \end{cases} \tag{4.24}$$

Comparing (4.23) with (4.24) and setting $\lambda := \Sigma^{-\frac{1}{2}}u$, it is clear that (x, w, λ) is a saddle point of $\tilde{\mathcal{L}}$. The proof is complete. $\square$

Remark 4.5. We note that

$$\tilde{\mathcal{L}}(x, y, \lambda) := f(x) + g(w) + \left\langle \Sigma^{\frac{1}{2}}\lambda, Lx + Jw - b \right\rangle$$

is the Lagrangian function of the following problem:

$$\begin{aligned} \min_{x\in\mathcal{H}, w\in\mathcal{K}} \quad & f(x) + g(w) \\ \text{s.t.} \quad & \Sigma^{\frac{1}{2}}(Lx + Jw - b) = 0. \end{aligned} \tag{4.25}$$

Since $\Sigma^{\frac{1}{2}}$ is a bijection, the equivalence of (4.18) and (4.25) can be observed. Further, if (x, y, λ) is a saddle point of $\tilde{\mathcal{L}}$, then (x, w) is a solution to (4.18) and $Lx + Jw = b$.

For solving problem (4.18), we propose a novel preconditioned ADMM (NP-ADMM) as follows:

Algorithm 4.2 Novel Preconditioned ADMM for Solving Problem (4.18).

Let $\Sigma : \mathcal{G} \to \mathcal{G}$ and $Y : \mathcal{H} \to \mathcal{H}$ be strongly monotone self-adjoint continuous linear operators. For given (x_k, λ_k, w_k), the new iterate $(x_{k+1}, \lambda_{k+1}, w_{k+1})$ is generated by the following recursion:

$$x_{k+1} = \arg\min_{x\in\mathcal{H}} \left\{ f(x) + \frac{1}{2}\|Lx + Jw_k - b + \Sigma^{-\frac{1}{2}}\lambda_k\|_{\Sigma}^2 \right.$$

$$+\frac{1}{2}\|x-x_k\|^2_{Y^{-1}-L^*\Sigma L}\Bigg\}, \tag{4.26}$$

$$\lambda_{k+1}=\lambda_k+\Sigma^{\frac{1}{2}}(Lx_{k+1}+Jw_k-b), \tag{4.27}$$

$$w_{k+1}\in\arg\min_{w\in\mathcal{K}}\left\{g(w)+\frac{1}{2}\|Lx_{k+1}+Jw-b+\Sigma^{-\frac{1}{2}}\lambda_{k+1}\|^2_\Sigma\right\}. \tag{4.28}$$

The following result provides the existence of solutions to the subproblem (4.28) in Algorithm 4.2 under Assumption 4.2 (A3).

Lemma 4.4. *Suppose* $0\in\operatorname{sri}(\operatorname{dom}g^*-\operatorname{ran}J^*)$ *holds. Let* $\{(x_k,\lambda_k,w_k)\}_{k=1}^\infty$ *be the sequence generated by Algorithm 4.2, then* $\{w_k\}$ *is well defined. Moreover, denoting* $y_k=b-Jw_k$*, the sequence* $\{(y_k,x_k,\lambda_k)\}_{k=1}^\infty$ *satisfies*

$$x_{k+1}=\arg\min_{x\in\mathcal{H}}\left\{f(x)+\frac{1}{2}\|Lx-y_k+\Sigma^{-\frac{1}{2}}\lambda_k\|^2_\Sigma+\frac{1}{2}\|x-x_k\|^2_{Y^{-1}-L^*\Sigma L}\right\}, \tag{4.29}$$

$$\lambda_{k+1}=\lambda_k+\Sigma^{\frac{1}{2}}(Lx_{k+1}-y_k), \tag{4.30}$$

$$y_{k+1}=\arg\min_{y\in\mathcal{G}}\left\{((-J)\triangleright g)(y-b)+\frac{1}{2}\|y-Lx_{k+1}-\Sigma^{-\frac{1}{2}}\lambda_{k+1}\|^2_\Sigma\right\}. \tag{4.31}$$

Proof. Set

$$y_{k+1}=\arg\min_{y\in\mathcal{G}}\left\{((-J)\triangleright g)(y-b)+\frac{1}{2}\|y-Lx_{k+1}-\Sigma^{\frac{1}{2}}\lambda_{k+1}\|^2_\Sigma\right\}. \tag{4.32}$$

Since $0\in\operatorname{sri}(\operatorname{dom}g^*-\operatorname{ran}J^*)$, we get $(-J)\triangleright g=(g^*\circ(-J)^*)^*\in\Gamma_0(\mathcal{G})$ and $\{y_k\}$ is well defined. Using Fermat's Theorem, we have

$$0\in\partial\big((-J)\triangleright g\big)(y_{k+1}-b)+\Sigma(y_{k+1}-Lx_{k+1})-\Sigma^{\frac{1}{2}}\lambda_{k+1},$$

i.e.,

$$\begin{aligned}y_{k+1}-b&\in\partial\big(g^*\circ(-J^*)\big)\big(-\Sigma(y_{k+1}-Lx_{k+1})+\Sigma^{\frac{1}{2}}\lambda_{k+1}\big)\\&=-J\partial g^*\big(J^*(\Sigma(y_{k+1}-Lx_{k+1})-\Sigma^{\frac{1}{2}}\lambda_{k+1})\big).\end{aligned}$$

Hence, there exists $w_{k+1}\in\partial g^*(J^*(\Sigma(y_{k+1}-Lx_{k+1})-\Sigma^{\frac{1}{2}}\lambda_{k+1}))$ such that $y_{k+1}=b-Jw_{k+1}$. Further, we have

$$J^*\big(\Sigma(y_{k+1}-Lx_{k+1})-\Sigma^{\frac{1}{2}}\lambda_{k+1}\big)\in\partial g(w_{k+1}). \tag{4.33}$$

Substituting $y_{k+1}=b-Jw_{k+1}$ in (4.33), we get

$$0\in\partial g(w_{k+1})+J^*\big(\Sigma(Jw_{k+1}+Lx_{k+1}-b)+\Sigma^{\frac{1}{2}}\lambda_{k+1}\big).$$

Using Fermat's Theorem, we yield

$$w_{k+1} \in \arg\min_{w \in \mathcal{K}} \left\{ g(w) + \frac{1}{2} \|Jw + Lx_{k+1} - b + \Sigma^{-\frac{1}{2}} \lambda_{k+1}\|_{\Sigma}^2 \right\}, \tag{4.34}$$

thus (4.31) and the existence of w_k can be obtained. Substituting $y_{k+1} = b - Jw_{k+1}$ in (4.26) and (4.27), respectively, we can have (4.29) and (4.30). ☐

Remark 4.6. Let $\{(x_k, w_k, \lambda_k)\}_{k=1}^{\infty}$ be the sequence generated by Algorithm 4.2 and $y_k = b - Jw_k$. By using Lemma 4.4, we can obtain that $\{(x_k, y_k, \lambda_k)\}_{k=1}^{\infty}$ is the sequence generated by Algorithm 4.1 with $h = (-J \triangleright g)(\cdot - b)$.

Next, we propose the convergence and sublinear convergence rates of Algorithm 4.2.

Theorem 4.2. *Suppose Assumption* 4.2 *holds. Let* $\{(x_k, w_k)\}_{k=1}^{\infty}$ *be the sequence generated by Algorithm* 4.2*, where the operator* $Y^{-1} - 3L^*\Sigma L/4$ *is strongly monotone. Then, the following statements hold:*

(i) There exists a saddle point (x^*, w^*, λ^*) *of* $\tilde{\mathcal{L}}$*, such that* $x_k \rightharpoonup x^*$, $Jw_k \rightharpoonup Jw^*$, *and* $\lambda_k \rightharpoonup \lambda^*$.

(ii) Suppose $\operatorname{ran} J^* = \mathcal{K}$*, then we have* $w_k \rightharpoonup w^*$.

(iii) If $Y^{-1} - L^*\Sigma L$ *is monotone, then*

$$\|x_{k+1} - x_k\|_{Y^{-1}}^2 + \|Lx_{k+1} + Jw_k - b\|_{\Sigma}^2 = o(1/k).$$

(iv) The convergence condition (1.14) *is tight.*

Proof. (i) Using Theorem 3.1, Lemma 4.2, and Remark 4.6, we obtain that there exits a saddle point (x^*, u^*) of $\tilde{\Phi}$, which satisfies $x_k \rightharpoonup x^*$ and $\tilde{u}_k := \Sigma^{\frac{1}{2}} \lambda_k + \Sigma(Lx_k + Jw_k - b) \rightharpoonup u^*$. For (x^*, u^*), using Lemma 4.3, there exists $(w^*, \lambda^*) \in \mathcal{K} \times \mathcal{G}$ which satisfies (x^*, w^*, λ^*) is a saddle point of $\tilde{\mathcal{L}}$, $Lx^* + Jw^* = b$ and $u^* = \Sigma^{\frac{1}{2}} \lambda^*$. Combining

$$\lambda_{k+1} = \lambda_k + \Sigma^{\frac{1}{2}} (Lx_{k+1} + Jw_k - b),$$

we get

$$\Sigma^{\frac{1}{2}} \lambda_{k+1} + \Sigma(Lx_k - Lx_{k+1}) = \tilde{u}_k \rightharpoonup u^* = \Sigma^{\frac{1}{2}} \lambda^*.$$

From Theorem 3.1, we have $x_k - x_{k+1} \to 0$. Since $\Sigma^{\frac{1}{2}}$ is a bijection, we get $\lambda_k \rightharpoonup \lambda^*$. Using

$$\lambda_{k+1} = \lambda_k + \Sigma^{\frac{1}{2}} (Lx_{k+1} + Jw_k - b),$$

we obtain

$$Jw_k = \Sigma^{-\frac{1}{2}} (\lambda_{k+1} - \lambda_k) + b - Lx_{k+1} \rightharpoonup b - Lx^* = Jw^*.$$

(ii) If $\operatorname{ran} J^* = \mathcal{K}$, for any $v \in \mathcal{K}$, there exists $z \in \mathcal{K}$, such that $v = J^*z$. We conclude that

$$\langle w_k, v\rangle = \langle Jw_k, z\rangle \ \ \rightarrow \ \ \langle Jw^*, z\rangle = \langle w^*, v\rangle, \quad k \to \infty, \quad \forall v \in \mathcal{K}.$$

(iii) Setting $y_k := b - Jw_k$ and using Remark 4.6, we have that $\{(x_k, y_k, \lambda_k)\}_{k=1}^{\infty}$ is generated by Algorithm 4.1 with $h = ((-J) \triangleright g)(\cdot - b)$. Thus, by using Theorem 4.1, we get

$$\begin{aligned} &\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|Lx_{k+1} + Jw_k - b\|^2_{\Sigma} \\ = &\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|Lx_{k+1} - y_k\|^2_{\Sigma} = o(1/k). \end{aligned}$$

(iv) By setting $\mathcal{K} = \mathcal{G}$ and $J = -Id$, we can know Algorithm 4.1 is a special case of Algorithm 4.2. Thus, using Theorem 4.1, we can immediately obtain that the condition (1.14) is tight for Algorithm 4.2. The proof is complete. □

Remark 4.7. (i) Regarding problem (4.18), [8] introduced sublinear convergence rates (in the ergodic sense) for the residuals of function values and constraint violations in the generalized ADMM. In contrast to the conclusions in [8], our result concerning constraint violations is non-ergodic, and the proof methodology differs as well.

(ii) If we set $Y^{-1} = (\tau r)^{-1} Id$ and $\Sigma = \beta Id$ $(\tau, r, \beta > 0)$ and interchange the positions of x and w, Algorithm 4.2 can be reduced to the optimal linearized ADMM (OLADMM) proposed in [25]. In this scenario, condition (1.14) can degenerate to the result in [25]: $r > \beta\|L^{\mathrm{T}}L\|$ and $\tau \in (0.75, 1)$.

(iii) Particularly, in problem (4.25), if we set $\mathcal{G} = \mathcal{K}$, $g \equiv 0$, and J be a zero operator, then problem (4.25) can be reduced to a convex optimization problem with linear equality constraints

$$\begin{aligned} &\min_{x \in \mathcal{H}} \ f(x) \\ &\ \text{s.t.} \ \ Lx = b. \end{aligned} \tag{4.35}$$

Denote $w_0 = 0$, Algorithm 4.2 can be reduced to a kind of preconditioned linearized augmented Lagrangian method

$$x_{k+1} = \arg\min_{x \in \mathcal{H}} \left\{ f(x) + \frac{1}{2}\|Lx - b + \Sigma^{-\frac{1}{2}}\lambda_k\|^2_{\Sigma} + \frac{1}{2}\|x - x_k\|^2_{Y^{-1} - L^*\Sigma L} \right\}, \tag{4.36}$$

$$\lambda_{k+1} = \lambda_k + \Sigma^{\frac{1}{2}}(Lx_{k+1} - b). \tag{4.37}$$

Using Theorem 4.2, we can get directly the convergence and the convergence rate of the sequence $\{(x_k, \lambda_k)\}_{k=1}^{\infty}$ satisfying (4.36) and (4.37).

Corollary 4.1. *Suppose that* $b \in \operatorname{sri}(L \operatorname{dom} f)$ *and the solutions to the problem* (4.35) *exist. For a given initial point* $(x_0, \lambda_0) \in \mathcal{H} \times \mathcal{G}$*, let* $\{(x_k, \lambda_k)\}_{k=1}^{\infty}$ *be the sequence satisfying* (4.36) *and* (4.37)*. If* $Y^{-1} - 3L^*\Sigma L/4$ *is strongly monotone, then the following statements hold:*

(i) There exists a saddle point (x^*, λ^*) *of*

$$\bar{\mathcal{L}}(x,\lambda) := f(x) + \langle \Sigma^{\frac{1}{2}}\lambda, Lx - b\rangle$$

such that $x_k \rightharpoonup x^*$ *and* $\lambda_k \rightharpoonup \lambda^*$.

(ii) If $Y^{-1} - L^*\Sigma L$ *is monotone, then*

$$\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|Lx_{k+1} - b\|^2_{\Sigma} = o(1/k).$$

(iii) The convergence condition (1.14) *can not be improved.*

Remark 4.8. If we set $\Sigma = Id$, $Y = \tau Id$ $(\tau > 0)$, then (4.36) and (4.37) can be reduced to the linearized augmented Lagrangian method (LALM) proposed in [40] with penalty parameter $\beta = 1$ in augmented Lagrangian function

$$\mathcal{L}_\beta(x,\lambda) = f(x) + \langle\lambda, Lx - b\rangle + \frac{\beta}{2}\|Lx - b\|^2.$$

In this case, the stepsize range (1.14) is the same as the result in [24] $\tau\|L\|^2 < 4/3$.

5. Numerical experiments

This section focuses on testing four special applications of the monotone inclusion problem (1.1) and presenting the numerical results to support our theoretical statements in Sections 3 and 4.

The first and second experiments involve the use of Algorithm 3.1 to solve the classical traffic equilibrium problem and the basis pursuit problem, respectively. We compare the performance of Algorithm 3.1 under conditions (1.13) and (1.14). The third and fourth experiments provide a report on the numerical performance of Algorithm 4.2 for solving separable convex optimization problems with linear equality constraints. Specifically, we consider the covariance selection problem and the TV-L^1 denoising problem under conditions (1.13) and (1.14), respectively. Based on the aforementioned experiments, we have observed that the condition (1.14) can result in improved numerical performance. Our code is implemented in MATLAB 2021a.

5.1. Traffic equilibrium problem

Firstly, we consider a strongly connected transportation network $G(\mathcal{N}, \mathcal{A})$. The node set of this network is denoted $\mathcal{N}$ and the arc set is denoted $\mathcal{A}$. There are two different subsets of $\mathcal{N}$ that represent origin nodes set O and destination nodes set D, respectively. The set of origin-destination (OD) pairs is a given subset RS of $O \times D$. For every given OD pair $rs \in RS$, where r is a origin node, s is a destination node. For each $rs \in RS$, denote q_{rs} as the travel demand between the OD pair rs. Let P_{rs} denote the set of paths connecting the OD pair rs and let P be the union of P_{rs} for $rs \in RS$, i.e.

$P := \bigcup_{rs\in RS} P_{rs}$. For a given $rs \in RS$ and $p \in P_{rs}$, h^p_{rs} represents the flow on path p. Let v_a be the flow on arc a. Setting total path flow vector $h = (h^p_{rs}, rs \in RS, p \in P_{rs})^\intercal$ and total arc flow vector $v := (v_a, a \in \mathcal{A})^\intercal$, we can get $v = \Delta h$, where $\Delta = (\delta_{a,p})$ is the arc-path incidence matrix with entries

$$\delta_{a,p} := \begin{cases} 1, & \text{if path } p \in P \text{ includes arc } a \in \mathcal{A}, \\ 0, & \text{otherwise.} \end{cases}$$

Setting total travel demand $q = (q_{rs}, rs \in RS)^\intercal$, Ω is the OD pair-path incidence matrix whose entries are

$$\omega_{p,rs} = \begin{cases} 1, & \text{if} \quad p \in P_{rs}, \\ 0, & \text{otherwise,} \end{cases}$$

then the feasible set of flow is

$$H := \{h \mid Fh \geq b\},$$

where

$$F = \begin{pmatrix} \Omega \\ -\Omega \\ I \end{pmatrix}, \quad b = \begin{pmatrix} q \\ -q \\ O \end{pmatrix}.$$

Let $t_a(v)$ be a travel time function about the total flow vector v on $a \in \mathcal{A}$, in general, the larger the traffic volume, the more time it takes for vehicles to pass, so it can be assumed that $t_a(v)$ is a continuous increasing function. T_p is the sum of the arc costs $t_a(v)$ on all the arcs a traversed by the path $p \in P$, i.e.

$$T_p(h) = \sum_{a\in\mathcal{A}} \delta_{a,p} t_a(v). \tag{5.1}$$

We can rewrite (5.1) as

$$T(h) := (T_p, p \in P)^\intercal = (\Delta)^\intercal t(v) = (\Delta)^\intercal t(\Delta h),$$

where $t(v) = (t_a(v), a \in \mathcal{A})^\intercal$. t_a is a continuous increasing function, hence, T is maximal monotone. Let c_{rs} be the (unknown) minimum travel costs between OD pair rs, $c = (c_{rs}, rs \in RS)^\intercal$, using Wardrop user equilibrium principle and [20, Proposition 1.4.8], the above traffic equilibrium problem can be converted into the following variational inequality:

$$(h - h^*)^\intercal T(h^*) \geq 0, \quad \forall h \in H. \tag{5.2}$$

By utilizing the Lagrange multiplier, the problem (5.2) is equivalent to the following primal-dual monotone inclusion problem:

$$\begin{cases} 0 \in Th - F^\intercal y, \\ 0 \in \mathcal{N}_{\mathbb{R}^m_+}(y) - b + Fh, \end{cases} \tag{5.3}$$

where $m = 3|RS|$. Utilizing Remark 2.1, we choose $L = -F$, $s = 1, t = \{0.9, 1, 1.1, 1.33\}$ and $\alpha = 1$, then the iterative scheme of Algorithm 3.1 for solving (5.3) reads as

$$\begin{cases} \tilde{y}_k = P_{\mathbb{R}^m_+}\big(y_k - \Sigma(Fh_k - b)\big), \\ h_{k+1} = (Id + YT)^{-1}\{h_k + YF^\intercal \tilde{y}_k\}, \\ y_{k+1} = \tilde{y}_k + \Sigma F(h_k - h_{k+1}). \end{cases} \tag{5.4}$$

We consider an example in [22]. In this example, $G(\mathcal{N}, \mathcal{A})$ consists of 9 nodes, 28 arcs, 72 OD pairs, and 1216 paths. The diagram of this network reads as Fig. 1.

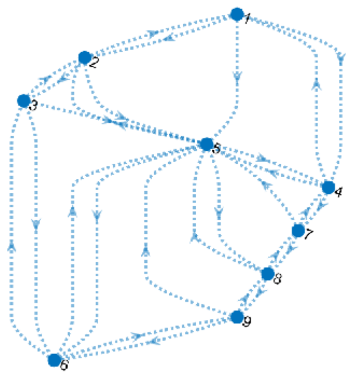

Figure 1: The network example $G(\mathcal{N}, \mathcal{A})$.

In this section, the arc travel time function follows the Bureau of Public Road (BPR) function

$$t_a(v) = \alpha_a \left[1 + 0.15\left(\frac{v_a}{C_a}\right)^4\right],$$

where α_a and C_a are free-flow travel time and capacity of link a, respectively.

We arrange and number the paths based on the sequence number of the origin point. We set the free-flow travel time as a sequence of equal differences about the sequence number of the paths, with an initial value of 15 and a tolerance of 1. Additionally, we set the capacity of every link a as a sequence of equal differences about the sequence number of the paths, with an initial value of 30 and a tolerance of 2.

For the traffic demand vector b between each OD pair, we number it based on the destination point of the OD pair. We set the traffic demand between the OD pairs as a sequence of equal differences about the sequence number of the OD pair, with an initial value of 100 and a tolerance of 10. The stopping criterion of Algorithm 3.1 is

$$\|h_{k+1} - h_k\|^2 + \|y_{k+1} - y_k\|^2 \le 10^{-5}.$$

We show the fixed point residual $\|h_{k+1} - h_k\|^2 + \|y_{k+1} - y_k\|^2$ and relative error $(\|h_{k+1} - h_k\|^2 + \|y_{k+1} - y_k\|^2)/(\|h_k\|^2 + \|y_k\|^2)$ curves with respect to iterations, respectively, in Fig. 2 and use "Iter." for the iteration numbers. From Fig. 2, we can observe that condition (1.14) has a significant acceleration effect compared to condition (1.13).

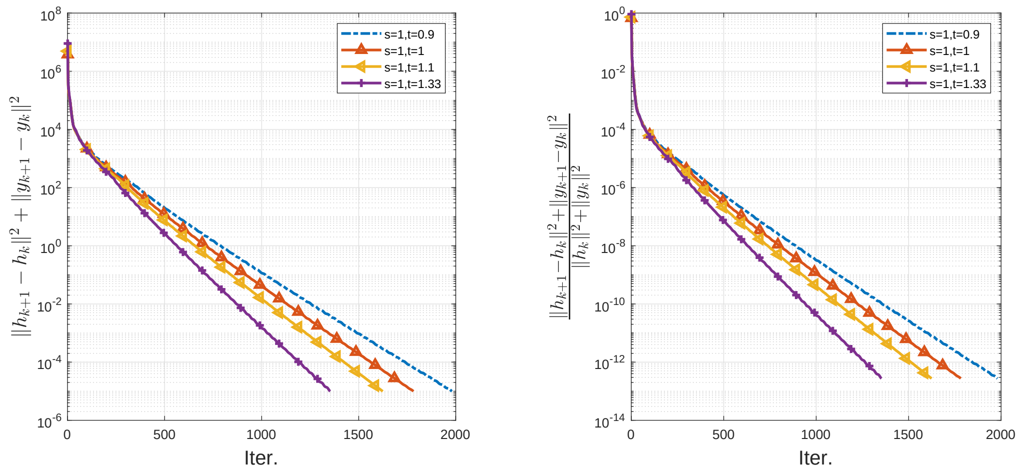

Figure 2: The comparison of the numerical performance of SDR with different Σ and Y for solving the traffic equilibrium problem (5.3).

5.2. Basis pursuit (BP)

Secondly, we consider the basis pursuit problem as follows:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \ \|x\|_1 \\ &\text{s.t.} \quad Lx = b, \end{aligned} \tag{5.5}$$

where $\|x\|_1 = \sum_{k=1}^{n} |x_k|$, $b \in \mathbb{R}^m$, and $L \in \mathbb{R}^{m\times n}$ is a data matrix. The BP problem (5.5) plays a key role in compressed sensing and statistical learning [7, 13]. Consider the primal-dual form of the BP problem (5.5) as

$$\min_{x\in\mathbb{R}^n} \max_{u\in\mathbb{R}^m} \|x\|_1 + \langle u, Lx\rangle - \langle u, b\rangle.$$

It is equivalent to the following monotone inclusion problem:

$$\begin{cases} 0 \in \partial\|x\|_1 + L^* u, \\ 0 = b - Lx. \end{cases}$$

Using Algorithm 3.1 to solve inclusion problem (5.2), the scheme of Algorithm 3.1 is

$$\begin{cases} \tilde{u}_k = u_k + \Sigma(Lx_k - b), \\ x_{k+1} = \arg\min\limits_{x\in\mathbb{R}^n} \left\{ \|x\|_1 + \dfrac{1}{2}\|x - x_k + Y L^* \tilde{u}_k\|_{Y^{-1}} \right\}, \\ u_{k+1} = \tilde{u}_k - \Sigma L(x_k - x_{k+1}). \end{cases} \tag{5.6}$$

We take $n = m = 3000$. To improve the accuracy of the test, we take three groups of random data for this numerical experiment and finally take the average value. We

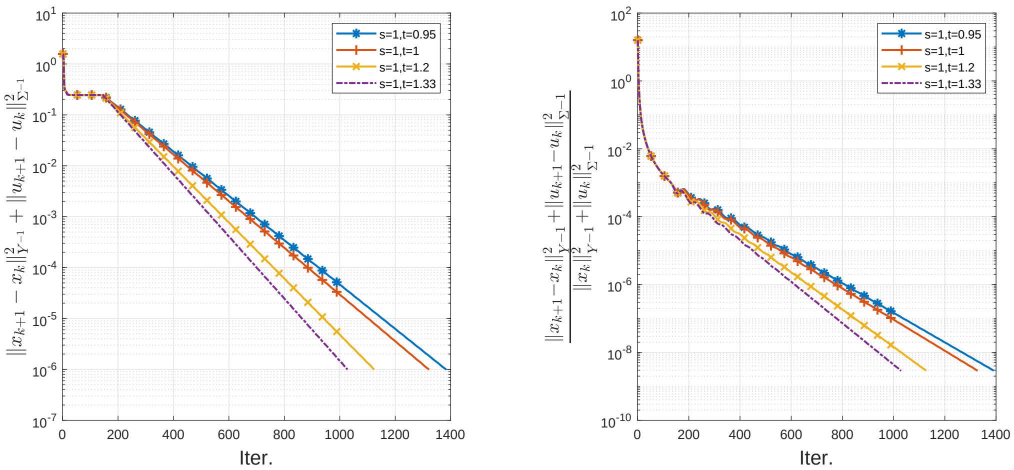

Figure 3: The comparison of the numerical performance of SDR with different stepsizes for solving problem (5.2).

set $L \in \mathbb{R}^{3000\times 3000}$ being a discrete cosine transform (DCT) [1], $x^* \in \mathbb{R}^{3000}$ whose entries satisfy standard normal distribution and the sparse density of x^* is 0.05. In this experiment, we denote $b = Lx^*$, and use $x_0 = \mathbf{0}$ and $u_0 = \mathbf{0}$ as the initial points. Using Remark 2.1, we choose $s = 1$, $t = \{0.95, 1, 1.2, 1.33\}$, and the stopping criterion of Algorithm 3.1 is

$$\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|u_{k+1} - u_k\|^2_{\Sigma^{-1}} \le 10^{-6}.$$

In Fig. 3, we plot the fixed point residual $\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|u_{k+1} - u_k\|^2_{\Sigma^{-1}}$ and relative error $(\|x_{k+1} - x_k\|^2_{Y^{-1}} + \|u_{k+1} - u_k\|^2_{\Sigma^{-1}})/(\|x_k\|^2_{Y^{-1}} + \|u_k\|^2_{\Sigma^{-1}})$ curves with respect to iterations, respectively. From Fig. 3, we can see that a larger t results in a small number of iterations under the same stopping criterion.

5.3. Covariance selection problem

Thirdly, we consider the covariance selection problem penalized by the l_1 norm

$$\min_{X\in\mathbb{S}^n_+} \langle S, X\rangle - \ln\big(\det(X)\big) + \rho e^T |X| e, \tag{5.7}$$

where S is a known real symmetric matrix and $\mathbb{S}^n_+$ represents the set of symmetric $n \times n$ positive semi-definite matrices. $\langle\cdot,\cdot\rangle$ is the standard trace inner product in $\mathbb{R}^{n\times n}$, $e \in \mathbb{R}^n$ denotes the vector whose elements are all 1, $|X|$ represents a real matrix whose elements are the absolute value of the corresponding elements of X. The covariance selection problem has a wide range of applications in various fields [17–19, 41], such as portfolio, speech recognition, gene network analysis, machine learning, and so on.

By introducing auxiliary variables Y, we can rewrite problem (5.7) as

$$\begin{aligned} \min_{X,Y}\ & \langle S, X\rangle - \ln\big(\det(X)\big) + I_{\mathbb{S}^n_+}(X) + \rho e^T |Y| e, \\ \text{s.t.}\ & X - Y = 0. \end{aligned} \tag{5.8}$$

It is clear that $\mathbb{S}^n_+$ is a closed and convex subset of $\mathbb{R}^{n\times n}$, therefore we have $I_{\mathbb{S}^n_+} \in \Gamma_0(\mathbb{R}^{n\times n})$. Utilizing Algorithm 4.2 to solve problem (5.8), we set $\Sigma = Y = \sigma Id$, then (1.14) reduces to $\sigma^2 < 4/3$. In this setting, the recursion of Algorithm 4.2 reads as:

$$\begin{cases} X_{k+1} = \arg\min_{X\in\mathbb{S}^n_+} \left\{ \langle S, X\rangle - \ln\big(\det(X)\big) + \frac{\sigma}{2}\left\| X - Y_k + \frac{1}{\sqrt{\sigma}} Z_k \right\|^2 \right. \\ \qquad\qquad\qquad\qquad\qquad\qquad \left. + \frac{1-\sigma^2}{2\sigma}\|X - X_k\|^2 \right\}, \\ Z_{k+1} = Z_k + \sqrt{\sigma}(X_{k+1} - Y_k), \\ Y_{k+1} = \arg\min_{Y\in\mathbb{S}^n_+} \left\{ \rho e^T|Y|e + \frac{\sigma}{2}\left\| X_{k+1} - Y + \frac{1}{\sqrt{\sigma}} Z_{k+1}\right\|^2 \right\}. \end{cases}$$

First, we consider the X_{k+1} subproblem. Because $-\ln x \to +\infty,\ x \to 0^+$, X_{k+1} must be a interior point of S^n_+, i.e. $X_{k+1} \in S^n_{++}$. Utilizing Fermat's Theorem, we can get that X_{k+1} satisfies

$$S - \sigma Y_k + \sqrt{\sigma} Z_k - \frac{1-t^2}{t} X_k + \frac{1}{\sigma} X_{k+1} - X_{k+1}^{-1} = 0. \tag{5.9}$$

For the sake of convenience, we set $F = S - \sigma Y_k + \sqrt{\sigma} Z_k - (1-t^2)X_k/t$ and let

$$F = G\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) G^{\mathsf{T}} \tag{5.10}$$

be the symmetric Schur decomposition of F, where G is an orthogonal matrix and λ_i $(i = 1, 2, \ldots, n)$ are eigenvalues of F. The symmetric Schur decomposition of the solution of (5.9) can be written into the following form [21]:

$$X_{k+1} = G\,\mathrm{diag}(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_n) G^{\mathsf{T}}, \tag{5.11}$$

where $\tilde{\lambda}_i$ $(i = 1, 2, \ldots, n)$ are to be determined. In fact, substituting (5.10) and (5.11) into (5.9), we obtain

$$\tilde{\lambda}_i = \frac{-\sigma\lambda_i + \sqrt{\sigma^2\lambda_i^2 + 4\sigma}}{2}, \quad i = 1, 2, \ldots, n.$$

Next, we solve Y_{k+1} subproblem. Using the definition of proximal operator and Moreau identity, we get

$$Y_{k+1} = X_{k+1} + \frac{1}{\sqrt{\sigma}} Z_{k+1} - \min\left\{ \max\left\{ X_{k+1} + \frac{1}{\sqrt{\sigma}} Z_{k+1}, -\frac{\rho}{\sigma} \right\}, \frac{\rho}{\sigma} \right\}.$$

In this experiment, we generate the data of the problem (5.8) identically as [33]. In detail $S = A^{-1} + lV - \min(\lambda_{\min}(A^{-1} + lV) - v, 0)Id$, where V is an independent and identically distributed uniform random symmetric matrix, A is a sparse invertible matrix with positive diagonal elements, Id is the $n \times n$ identity matrix, l and v are small

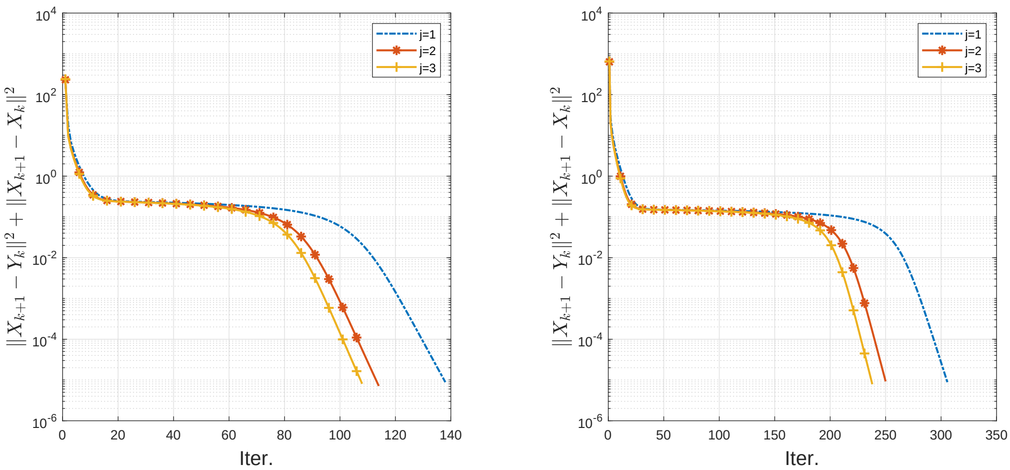

Figure 4: The comparison of NP-ADMM with different stepsizes for solving covariance selection problem (5.8), $n = 500$ (left), $n = 1000$ (right).

positive numbers. Similar to parameter setting in [33], we take the sparse density of A to be 0.01, $l = 0.15, v = 0.0001$, and $\rho = 0.5$. We set $\sigma = \tau = \sqrt{((2j-1)/2j)(4/3)}$, $(j = 1, 2, 3)$ and the stopping criterion is

$$\|X_{k+1} - X_k\|^2 + \|X_{k+1} - Y_k\|^2 < 10^{-5}.$$

To improve the accuracy of the test, we take three groups of random data for this test and finally take the average value. From Fig. 4, we draw the fixed point residual $\|X_{k+1}-Y_k\|^2+\|X_{k+1}-X_k\|^2$ curve with iterations in the situation of dimension $n = 500$ and $n = 1000$. As can be seen from Fig. 4, when the stepsizes of Algorithm 4.2 are close to the boundary of the condition (1.14), the number of iterations required is significantly reduced.

We present the more detailed numerical result of Algorithm 4.2 under different choices of Σ and a variety of dimensions n in Table 3 where "Time" is the computing time in seconds. We can observe that under the same dimension, the CPU time and the number of iterations of Algorithm 4.2 under condition (1.13) are about 30% higher than those under condition (1.14). This advantage of condition (1.14) is more obvious for large-scale situations of problem (5.8).

5.4. TV-L_1 denoising

Finally, we consider the following TV-L_1 denoising problem (discrete vision) for image denoising

$$\min_{x \in \mathbb{R}^N} \tau\|\mathcal{D}x\|_1 + \|x - q\|_1, \tag{5.12}$$

where q is the observed image, $\mathcal{D} : x \mapsto \mathcal{D}x := \begin{pmatrix} \mathcal{D}_1 x \\ \mathcal{D}_2 x \end{pmatrix}$ is the discrete gradient operator, and τ is a regularization parameter. In our experiment, we set $\tau = 1$ and input image q

Table 3: The numerical result of solving problem (5.8).

n	$j=1$		$j=2$		$j=3$	
	Iter.	Time(s)	Iter.	Time(s)	Iter.	Time(s)
100	32	0.7702	27	0.2308	**26**	**0.1894**
200	56	1.6931	47	1.2226	**44**	**1.1039**
300	82	6.5089	67	4.1391	**64**	**4.0714**
400	108	12.5759	89	10.6139	**85**	**9.8401**
500	137	25.1082	112	18.8840	**107**	**18.3507**
600	167	44.2045	137	37.1000	**130**	**33.2041**
700	199	71.3312	163	58.8336	**155**	**55.3372**
800	233	124.6614	191	101.4911	**181**	**96.0022**
900	267	192.5235	219	155.2754	**208**	**139.1944**
1000	304	373.0020	249	284.4315	**236**	**270.8189**

with 25% salt and pepper noise. We employ SNR(dB), i.e.,

$$\mathrm{SNR} = 20\lg \frac{\|x^*\|}{\|\tilde{x} - x^*\|}$$

to measure the quality of a recovered image where $\tilde{x}$ is a reconstructed image and x^* is the real image. To simplify the subproblem, we use the following equivalent form of (5.12) proposed in [39]:

$$\begin{aligned} &\min_{x,y,z} \ \tau\|y\|_1 + \|z-q\|_1 \\ &\text{s.t.} \ \begin{pmatrix} \mathcal{D} \\ Id \end{pmatrix} x + \begin{pmatrix} -Id & O \\ O & -Id \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \end{aligned} \tag{5.13}$$

We set

$$\begin{pmatrix} y \\ z \end{pmatrix} = w, \quad f(x) = 0, \quad g(w) = \tau\|y\|_1 + \|z-q\|_1, \quad L = \begin{pmatrix} \mathcal{D} \\ Id \end{pmatrix}, \quad J = \begin{pmatrix} -Id & O \\ O & -Id \end{pmatrix}.$$

Noting the structure of $\mathcal{D}$, we decomposed y as $y = \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}$. In this setting, we set $\Sigma = \sigma Id_N$, $Y = \sigma Id_{3N}$ $(\sigma > 0)$, thus, we can rewrite Algorithm 4.2 as

$$\begin{cases} x_{k+1} = x_k - \sigma^2\big(\mathcal{D}_1^\intercal \mathcal{D}_1 + \mathcal{D}_2^\intercal \mathcal{D}_2 + Id\big)x_k \\ \qquad\quad + \sigma^2\big(\mathcal{D}_1^\intercal y_k^1 + \mathcal{D}_2^\intercal y_k^2 + z_k - \mathcal{D}_1^\intercal \lambda_k^{y^1} - \mathcal{D}_2^\intercal \lambda_k^{y^2} - \lambda_k^z/\sqrt{\sigma}\big), \\ \lambda_{k+1}^{y^1} = \lambda_k^{y^1} + \sqrt{\sigma}\big(\mathcal{D}_1 x_{k+1} - y_k^1\big), \\ \lambda_{k+1}^{y^2} = \lambda_k^{y^2} + \sqrt{\sigma}\big(\mathcal{D}_2 x_{k+1} - y_k^2\big), \\ \lambda_{k+1}^z = \lambda_k^z + \sqrt{\sigma}(x_{k+1} - z_k), \\ y_{k+1}^1 = \mathrm{prox}_{\frac{\tau}{\sigma}\|\cdot\|_1}\big(\mathcal{D}_1 x_{k+1} + \lambda_{k+1}^{y_1}/\sqrt{\sigma}\big), \\ y_{k+1}^2 = \mathrm{prox}_{\frac{\tau}{\sigma}\|\cdot\|_1}\big(\mathcal{D}_2 x_{k+1} + \lambda_{k+1}^{y_2}/\sqrt{\sigma}\big), \\ z_{k+1}^1 = q + \mathrm{prox}_{\frac{1}{\sigma}\|\cdot\|_1}\big(x_{k+1} + \lambda_{k+1}^{y_1}/\sqrt{\sigma} - q\big). \end{cases}$$

Under this choice of Σ and Y, the condition (1.14) can be rewritten as

$$\sigma^2\|L\|^2 < \frac{4}{3}.$$

Next, we provide a range of σ by estimating $\|L\|$. For any given $x \in \mathbb{R}^N$, we have

$$\|Lx\|^2 = \|\mathcal{D}x\|^2 + \|x\|^2 \leq \big(\|\mathcal{D}\|^2 + 1\big)\|x\|^2.$$

Using $\|\mathcal{D}\| < \sqrt{8}$ [10], we can get

$$\|L\| = \sup_{x\in\mathbb{R}^N\setminus\{\mathbf{0}\}} \frac{\|Lx\|}{\|x\|} \leq 3.$$

Thus, we can set $\sigma = \sqrt{1/9 + (1-2^{-d})/27},\ d = 0,1,2,3$ in NP-ADMM. We take House (720×960) and Peppers (512×512) for testing. In Fig. 5, we show the SNR with respect to iterations.

In Table 4, we report more detailed numerical results, and the restored images are summarized in Fig. 6. These results in Table 4, Figs. 5 and 6 show that NP-ADMM

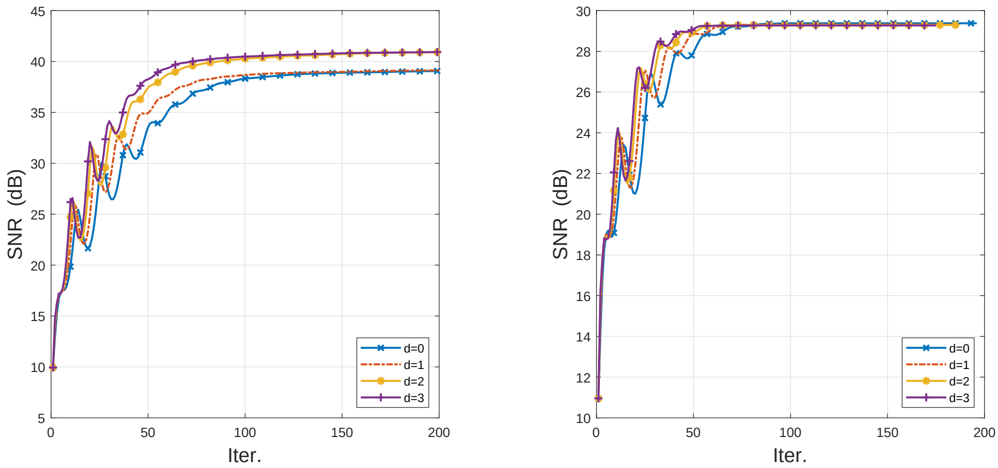

Figure 5: SNR with respect to iterations for solving TV-L_1 model (5.12). Left: House: 720×960 and right: Peppers: 512×512.

Table 4: Numerical results for solving problem (5.12).

Algorithm	House			Peppers		
	SNR	Time(s)	Iter.	SNR	Time(s)	Iter.
NP-ADMM $d=0$	39.0650	83.0778	171	29.2215	18.6319	87
NP-ADMM $d=1$	39.1329	81.3900	165	29.2471	17.3215	78
NP-ADMM $d=2$	40.9202	77.1009	154	29.2610	14.5727	66
NP-ADMM $d=3$	**40.9403**	**75.9415**	**150**	**29.2642**	**14.2776**	**65**

Figure 6: From left to right: original clean image, noisy image with 25% salt and pepper noise, image denoising using NP-ADMM ($d = 0, 1, 2, 3$). Test problem House (the first row), 720×960; test problem Peppers (the second row), 512×512.

can effectively solve TV-L_1 denoising problem (5.12). However, when the stepsize range approaches the boundary of (1.14), the number of iterations and iteration time gradually decrease and the advantage is more obvious for high-resolution images.

6. Conclusion

In this paper, we investigate the convergence of SDR for the primal-dual monotone inclusion problem and propose a novel preconditioned ADMM for solving separable convex optimization problems with linear equality constraints. We enhance the convergence condition of SDR. By constructing a concise counterexample, we demonstrate that the new range proposed in this paper cannot be improved any further. Additionally, to address the theoretical gap in the convergence rate of SDR, we establish a sublinear convergence rate in terms of fixed point residuals for the first time. Moreover, by utilizing SDR to solve separable convex optimization problems with linear equality constraints, we develop a novel preconditioned alternating direction method of multipliers (NP-ADMM) that can handle cases where the two linear operators are not identical. We also introduce tight convergence conditions and convergence rates for NP-ADMM. Finally, we conduct four numerical experiments to validate the computational efficiency of these algorithms and demonstrate that the performance of these algorithms has been significantly enhanced with our improved conditions.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grants 12131004 and 12126603.

References

[1] N. AHMED, T. NATARAJAN, AND K. R. RAO, *Discrete Cosine Transform*, IEEE Trans. Comput. 23 (1974), 90–93.

[2] K. J. ARROW, *Studies in Linear and Non-Linear Programming*, in: Stanford Mathematical Studies in the Social Science, Vol. 2, Stanford Unversity Press, 1958.

[3] H. ATTOUCH AND M. THÉRA, *A general duality principle for the sum of two operators*, J. Convex Anal. 3 (1996), 1–24.

[4] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books Math./Ouvrages Math. SMC, Springer, Cham, 2017.

[5] L. M. BRICEÑO-ARIAS AND P. L. COMBETTES, *A monotone + skew splitting model for composite monotone inclusions in duality*, SIAM J. Optim. 21 (2010), 1230–1250.

[6] L. M. BRICEÑO-ARIAS AND F. ROLDÁN, *Split-Douglas-Rachford for composite monotone inclusions and split-ADMM*, SIAM J. Optim. 31 (2021), 2987–3013.

[7] A. M. BRUCKSTEIN, D. L. DONOHO, AND M. ELAD, *From sparse solutions of systems of equations to sparse modeling of signals and images*, SIAM Rev. 51 (2009), 34–81.

[8] X. J. CAI AND D. R. HAN, *O(1/t) complexity analysis of the generalized alternating direction method of multipliers*, Sci. China Math. 62 (2019), 795–808.

[9] X. J. CAI, D. R. HAN, AND L. L. XU, *An improved first-order primal-dual algorithm with a new correction step*, J. Global Optim. 57 (2013), 1419–1428.

[10] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision 20 (2004), 89–97.

[11] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision 40 (2011), 120–145.

[12] A. CHAMBOLLE AND T. POCK, *An introduction to continuous optimization for imaging*, Acta Numer. 25 (2016), 161–319.

[13] S. S. CHEN AND D. SAUNDERS, *Atomic decomposition by basis pursuit*, SIAM Rev. 43 (2001), 129–159.

[14] P. L. COMBETTES AND B. C. VŨ, *Variable metric forward-backward splitting with applications to monotone inclusions in duality*, Optimization 9 (2014), 1289–1318.

[15] L. CONDAT, *A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, J. Optim. Theory Appl. 158 (2013), 460–479.

[16] J. B. CONWAY, *A Course in Functional Analysis*, Springer, 2007.

[17] D. R. COX AND N. WERMUTH, *Multivariate Dependencies: Models, Analysis and Interpretation*, Chapman and Hall, 1996.

[18] A. D'ASPREMONT, O. BANERJEE, AND L. EL GHAOUI, *First-order methods for sparse covariance selection*, SIAM J. Matrix Anal. Appl. 30 (2008), 56–66.

[19] A. DEMPSTER, *Covariance selection*, Biometrics 28 (1972), 157–175.

[20] F. FACCHINEI AND J. S. PANG, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II,* Springer, 2003.

[21] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, 1996.

[22] D. R. HAN AND HONG. K. LO, *Solving non-additive traffic assignment problems: A descent method for co-coercive variational inequalities*, European J. Oper. Res. 159 (2004), 529–544.

[23] B. S. HE, F. MA, S. J. XU, AND X. M. YUAN, *A generalized primal-dual algorithm with improved convergence condition for saddle point problems*, SIAM J. Imaging Sci. 15 (2022), 1157–1183.

[24] B. S. HE, F. MA, AND X. M. YUAN, *Optimal proximal augmented Lagrangian method and its application to full Jacobian splitting for multi-block separable convex minimization problems*, IMA J. Numer. Anal. 40 (2019), 1–29.

[25] B. S. HE, F. MA, AND X. M. YUAN, *Optimally linearizing the alternating direction method*

of multipliers for convex programming, Comput. Optim. Appl. 75 (2020), 361–388.

[26] B. S. HE, Y. F. YOU, AND X. M. YUAN, *On the convergence of primal-dual hybrid gradient algorithm*, SIAM J. Imaging Sci. 7 (2014), 2526–2537.

[27] B. S. HE AND X. M. YUAN, *Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective*, SIAM J. Imaging Sci. 5 (2012), 119–149.

[28] H. J. HE, J. DESAI, AND K. WANG, *A primal-dual prediction-correction algorithm for saddle point optimization*, J. Global Optim. 66 (2016), 573–583.

[29] F. JIANG, Z. Y. ZHANG, AND H. J. HE, *Solving saddle point problems: A landscape of primal-dual algorithm with larger stepsizes*, J. Global Optim. 85 (2022), 821–846.

[30] S. LANG, *Real and Functional Analysis*, Springer, 1993.

[31] Y. LI AND M. YAN, *On the improved conditions for some primal-dual algorithms*, arXiv: 2201.00139, 2022.

[32] Y. L. LIU, Y. B. XU, AND W. T. YIN, *Acceleration of primal-dual methods by preconditioning and simple subproblem procedures*, J. Sci. Comput. 86 (2021), 1–34.

[33] Z. LU, *Smooth optimization approach for covariance selection*, SIAM J. Control Optim. 19 (2009), 1807–1827.

[34] Y. M. MA, X. J. CAI, B. JIANG, AND D. R. HAN, *Understanding the convergence of the preconditioned PDHG method: A view of indefinite proximal ADMM*, J. Sci. Comput. 94(3) (2023), 1–39.

[35] T. POCK AND A. CHAMBOLLE, *Diagonal preconditioning for first order primal-dual algorithms in convex optimization*, in: IEEE International Conference on Computer Vision, (2011), 1762–1769.

[36] F. RIESZ AND B. S. NAGY, *Functional Analysis*, Ferderick Ungar Publishing Co, 1955.

[37] E. K. RYU AND W. T. YIN, *Large-Scale Convex Optimiztion: Algorithms & Analysis via Monotone Operators*, Cambridge University Press, 2022.

[38] R. E. SHOWALTER, *Montone Operators in Banach Space and Nonlinear Partial Differential Equations*, Mathematical Surveys and Monographs 49, AMS, 1997.

[39] E. Y. SIDKY, J. H. JØRGENSEN, AND X. C. PAN, *Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm*, Phys. Med. Biol. 57 (2012), 3065–3091.

[40] J. F. YANG AND X. M. YUAN, *Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization*, Math. Comp. 82 (2012), 301–329.

[41] X. M. YUAN, *Alternating Direction Method for Covariance Selection Models*, J. Sci. Comput. 51 (2012), 261–273.

[42] B. C. VŨ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Adv. Comput. Math. 38 (2013), 667–681.