Solution to a Monotone Inclusion Problem using the Relaxed Peaceman-Rachford Splitting Method: Convergence and its Rates

Chee-Khian Sim[*]

School of Mathematics and Physics
University of Portsmouth
Lion Gate Building, Lion Terrace
Portsmouth PO1 3HF

November 12, 2022

Abstract

We consider the convergence behavior using the relaxed Peaceman-Rachford splitting method to solve the monotone inclusion problem $0 \in (A + B)(u)$, where $A, B : \Re^n \rightrightarrows \Re^n$ are maximal β-strongly monotone operators, $n \geq 1$ and $\beta > 0$. Under a technical assumption, convergence of iterates using the method on the problem is proved when either A or B is single-valued, and the fixed relaxation parameter θ lies in the interval $(2+\beta, 2+\beta+\min\{\beta, 1/\beta\})$. With this convergence result, we address an open problem that is not settled in [22] on the convergence of these iterates for $\theta \in (2+\beta, 2+\beta+\min\{\beta, 1/\beta\})$. Pointwise convergence rate results and R-linear convergence rate results when θ lies in the interval $[2+\beta, 2+\beta+\min\{\beta, 1/\beta\})$ are also provided in the paper. Our analysis to achieve these results is atypical and hence novel. Numerical experiments on the weighted Lasso minimization problem are conducted to test the validity of the assumption.

Keywords. Relaxed Peaceman-Rachford splitting method; Maximal strong monotonicity; Convergence; Pointwise convergence rate; R-linear convergence rate.

1 Introduction

We consider the following monotone inclusion problem:

$$0 \in (A + B)(u), \tag{1}$$

where $A, B : \Re^n \rightrightarrows \Re^n$ are point-to-set operators that are maximal β-strongly monotone. When $\beta = 0$, A, B are maximal monotone operators, while when $\beta > 0$, A, B are maximal and strongly monotone with constant β in the usual sense. In our discussion in this paper, we always consider (1) having a solution and $\beta > 0$. With $\beta > 0$, then there exists only one solution to (1). Let this unique solution be given by u^*.

The relaxed Peaceman-Rachford (PR) splitting method is studied extensively in the literature, such as [3]-[5], [7]-[14], [16]-[18], [20, 22], to solve the monotone inclusion problem (1). The

---

[*]Email address: chee-khian.sim@port.ac.uk

method generates iterates $\{x_k\}$ using the following recursive relation:

$$x_k = x_{k-1} + \theta(J_{\rho B}(2J_{\rho A}(x_{k-1}) - x_{k-1}) - J_{\rho A}(x_{k-1})), \quad k \geq 1, \tag{2}$$

where x_0 is any point in $\Re^n$, $\theta > 0$ is a fixed relaxation parameter, $\rho > 0$ is an arbitrary scalar, and $J_T := (I + T)^{-1}$. In this paper, we always set ρ to be equal to 1. When $\theta = 1$, the method is also known as the Douglas-Rachford splitting method, while when $\theta = 2$, it is also called the Peaceman-Rachford splitting method. If the sequence $\{x_k\}$ generated by the relaxed PR splitting method converges to x^*, then the solution to (1) is given by $u^* = J_A(x^*)$.

Given A, B maximal β-strongly monotone operators, and $\{x_k\}$ a sequence generated by the relaxed PR splitting method (2) with $\rho = 1$, it is well-known that when $\beta = 0$, the sequence $\{x_k\}$ generated converges for $\theta \in (0, 2)$ (see for example [2, 15]), while in [22], it is shown that when $\beta > 0$, the sequence $\{x_k\}$ converges for $\theta \in (0, 2 + \beta]$. Furthermore, in [22], an instance of (1) is given for nonconvergence of $\{x_k\}$ for any $\theta \geq 2$ when $\beta = 0$. This instance also shows nonconvergence of $\{x_k\}$ for any $\theta \geq 2 + \beta + \min\{\beta, 1/\beta\}$ when $\beta > 0$. When $\beta > 0$, we see from [22] that the convergence behavior of iterates generated by (2) to solve (1) for $\theta \in (2 + \beta, 2 + \beta + \min\{\beta, 1/\beta\})$ is not known, and to the best of our knowledge, has not been studied previously in the literature. This paper attempts to fill the gap on this by investigating the convergence behavior of iterates $\{x_k\}$ generated by (2) to solve (1) when $\theta \in (2+\beta, 2+\beta+\min\{\beta, 1/\beta\})$. We show that an accumulation point, x^{**}, of $\{x_k\}$ has $J_A(x^{**})$ that solves (1) over this range of θ, under a technical assumption. As a consequence, if A or B is single-valued, then $\{x_k\}$ converges to a limit point x^*, where $J_A(x^*)$ solves (1). Note that for $n = 1$, not having this technical assumption leads to trivial consideration. We believe that the assumption for convergence is merely technical and is not really needed for convergence to occur. Also, through a numerical study in Section 6, we find that the assumption is always satisfied. We further believe that the convergence analysis for θ beyond $2 + \beta$ needs to be atypical and convergence is not shown in this range in the literature, for example, in [1, 9], where the focus of these papers is also different from ours. Our analysis to achieve these results is based on finding explicit solutions to small dimensional optimization problems and small systems of inequalities as detailed in Section 2. To add further to these contributions, we are able to provide pointwise convergence rate and R-linear convergence rate results using (2) to solve (1) for $\theta \in [2+\beta, 2+\beta+\min\{\beta, 1/\beta\})$, complementing results in [10]-[14], [16]-[18], [20] and [22].

This paper is divided into several sections. In Section 2, we state and prove some technical results that are needed in a latter section, Section 4, to prove convergence of $\{x_k\}$. Section 3 introduces transformations on the relaxed Peaceman-Rachford (PR) splitting method (2) that prepares us for analysis in Sections 4 and 5. Section 4 is on convergence of $\{x_k\}$, while Section 5 investigates pointwise convergence rate and R-linear convergence rate of $\{x_k\}$. Finally, Section 6 provides numerical results using (2) to solve the weighted Lasso minimization problem.

1.1 Conventions and Notations

$0 \cdot \infty$ is defined and is a real number, not necessarily zero. What value it takes is dependent on the context.

Given $x = (x_1, \ldots, x_n)^T, y = (y_1, \ldots, y_n)^T \in (\Re \cup \{\infty\})^n$, $x \odot y := (x_1 y_1, \ldots, x_n y_n)^T$.

Given $\alpha = (\alpha_1, \ldots, \alpha_n)^T \in (\Re \cup \{\infty\})^n$, $\|\alpha\|_\infty := \sup\{|\alpha_i| \; ; \; \alpha_i \in \Re, i = 1, \ldots, n\}$.

Given $x \in \Re^n$, $\|x\|$ stands for the 2-norm of x, while $\|x\|_1$ stands for the 1-norm of x.

$e \in \Re^n$ is the vector in $\Re^n$ of all ones.

2   Technical Results

Proposition 2.1 *Let* $0 < \epsilon < \min\{\beta, 1/\beta\}$, *and* $x, y \in \Re$. *We have*

$$\max_{xy \geq x^2} \left[\frac{1-\beta}{1+\beta} x^2 + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} xy + \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} y^2 \right] = 0,$$

and is attained when $x = y = 0$.

Proof: If $x = 0$, then it is clear that the objective function in the above maximization problem is less than or equal to zero, and is equal to zero when y is also equal to zero, since $\epsilon < \min\{\beta, 1/\beta\} \leq \beta$. Hence, we can assume that $x \neq 0$. The proposition is then proved if we can show that the following holds:

$$\max_{z \geq 1} \left[\frac{1-\beta}{1+\beta} + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} z + \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} z^2 \right] < 0.$$

Let

$$f(z) := \frac{1-\beta}{1+\beta} + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} z + \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} z^2.$$

The maximum of f over $\Re$ is attained at $\frac{(2+\beta+\epsilon)(3\beta-\epsilon-2)}{4(1+\beta)(\beta-\epsilon)}$, which can be checked easily to be less than 1, since $\epsilon < \min\{\beta, 1/\beta\}$. Hence, as f is concave over $\Re$, we have

$$\max_{z \geq 1} \left[\frac{1-\beta}{1+\beta} + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} z + \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} z^2 \right] = \frac{1-\beta}{1+\beta} + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} + \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2}.$$

The latter is less than zero as $0 < \epsilon < \min\{\beta, 1/\beta\}$. $\qquad\qquad\square$

Proposition 2.2 *Let* $0 < \epsilon < \min\{\beta, 1/\beta\}$, *and* $\alpha, x, y \in \Re$, *where* $|y| \leq 1$, *be such that*

$$xy \quad \geq \quad 1, \quad (3)$$

$$-\left(1 + \frac{2\beta(\alpha - 1)}{2 + \beta + \epsilon}\right) xy - \left(\frac{(1+\beta)(\alpha - 1)}{2 + \beta + \epsilon}\right)\left(1 + \frac{(1+\beta)(\alpha - 1)}{2 + \beta + \epsilon}\right) x^2 + \frac{1 - \beta}{1 + \beta} \quad \geq \quad 0, \quad (4)$$

then $|\alpha| \leq 1$.

Proof: Rearranging the left-hand side of (4), and letting $z = 1 - \alpha$, the left-hand side of (4) is given by the following function:

$$g(z) := -\frac{(1+\beta)^2 x^2}{(2+\beta+\epsilon)^2} z^2 + \frac{2\beta xy + (1+\beta)x^2}{2 + \beta + \epsilon} z - xy + \frac{1 - \beta}{1 + \beta}.$$

It is easy to check that the above quadratic function of z has its maximum point to be

$$z^* = \frac{(2 + \beta + \epsilon)(2\beta y + (1+\beta)x)}{2(1+\beta)^2 x}.$$

3

Since $|y| \leq 1$, $xy \geq 1$ and $0 < \epsilon < \min\{\beta, 1/\beta\} \leq 1$, we can check that

$$0 < z^* \leq 2.$$

For $z < 0$, since $xy \geq 1$, we have

$$g(z) < g(0) = -xy + \frac{1-\beta}{1+\beta} \leq 0.$$

For $z > 2$, we have

$$
\begin{aligned}
g(z) < g(2) &= \frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} x^2 + \frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} xy + \frac{1-\beta}{1+\beta} \\
&= x^2 \left[\frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} + \frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} \left(\frac{y}{x}\right) + \frac{1-\beta}{1+\beta}\left(\frac{1}{x^2}\right) \right] \\
&\leq x^2 \left[\frac{2(1+\beta)(\epsilon - \beta)}{(2+\beta+\epsilon)^2} y^2 + \frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} \left(\frac{y}{x}\right) + \frac{1-\beta}{1+\beta}\left(\frac{1}{x^2}\right) \right] \leq 0,
\end{aligned}
$$

where the second inequality holds since $|y| \leq 1$ and $\epsilon < \beta$, while the third inequality follows from $xy \geq 1$ and Proposition 2.1. Therefore, when $xy \geq 1$, for (4) to hold, that is, for $g(z) \geq 0$, we must have $0 \leq z \leq 2$, which is equivalent to $|\alpha| \leq 1$. $\qquad\square$

Proposition 2.3 *Let $x, z, x_1 \in \Re_+$ and $y, y_1 \in \Re$ satisfy*

$$
\begin{aligned}
y_1 - \frac{1-\beta}{1+\beta} x_1 - \frac{1-\beta}{1+\beta} x - \frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} y + \frac{2(1+\beta)(\beta - \epsilon)}{(2+\beta+\epsilon)^2} z &\leq 0, \\
x - y + x_1 - y_1 &\leq 0, \\
y^2 - xz &\leq 0,
\end{aligned}
$$

where $0 < \epsilon < \min\{\beta, 1/\beta\}$. Then $x = y = z = x_1 = y_1 = 0$.

Proof: Consider the following minimization problem:

$$
\min_{x,z,x_1 \in \Re_+, y, y_1 \in \Re} y_1 - \frac{1-\beta}{1+\beta} x_1 - \frac{1-\beta}{1+\beta} x - \frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} y + \frac{2(1+\beta)(\beta - \epsilon)}{(2+\beta+\epsilon)^2} z \tag{5}
$$

subject to

$$
\begin{aligned}
x - y + x_1 - y_1 &\leq 0, \tag{6} \\
y^2 - xz &\leq 0. \tag{7}
\end{aligned}
$$

Let $(x^*, y^*, z^*, x_1^*, y_1^*)$ be an optimal solution to the above minimization problem (5)-(7). By the Fritz-John condition [19] (see also [21]), there exist $\lambda_i \in \Re, i = 0, \ldots, 5$, such that the following holds:

$$
\lambda_0 \begin{pmatrix} -\frac{1-\beta}{1+\beta} \\ -\frac{3\beta - \epsilon - 2}{2+\beta+\epsilon} \\ \frac{2(1+\beta)(\beta - \epsilon)}{(2+\beta+\epsilon)^2} \\ -\frac{1-\beta}{1+\beta} \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -z^* \\ 2y^* \\ -x^* \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_5 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} = 0, \tag{8}
$$

4

with

$$\lambda_0 \geq 0, \quad (\lambda_0, \ldots, \lambda_5) \neq 0, \tag{9}$$

$$\lambda_1(x^* - y^* + x_1^* - y_1^*) = 0, \quad \lambda_1 \geq 0, \quad x^* - y^* + x_1^* - y_1^* \leq 0, \tag{10}$$

$$\lambda_2((y^*)^2 - x^*z^*) = 0, \quad \lambda_2 \geq 0, \quad (y^*)^2 - x^*z^* \leq 0, \tag{11}$$

$$\lambda_3 x^* = 0, \quad \lambda_3 \geq 0, \quad x^* \geq 0, \tag{12}$$

$$\lambda_4 z^* = 0, \quad \lambda_4 \geq 0, \quad z^* \geq 0, \tag{13}$$

$$\lambda_5 x_1^* = 0, \quad \lambda_5 \geq 0, \quad x_1^* \geq 0. \tag{14}$$

Case 1: $\lambda_0 \neq 0$.

We show that this case is impossible from (8)-(14). WLOG, let $\lambda_0 = 1$. Then, from the last "row" in (8), we get that $\lambda_1 = 1$. From the first to the third "row" in (8) and $\lambda_1 = 1$, we get

$$\lambda_2 z^* + \lambda_3 = \frac{2\beta}{1+\beta}, \tag{15}$$

$$\lambda_2 y^* = \frac{2\beta}{2+\beta+\epsilon}, \tag{16}$$

$$\lambda_2 x^* + \lambda_4 = \frac{2(1+\beta)(\beta-\epsilon)}{(2+\beta+\epsilon)^2}. \tag{17}$$

Since the right-hand side of (16) is positive, we have $\lambda_2 \neq 0, y^* \neq 0$. Hence, from the equality in (11), we get that $(y^*)^2 = x^*z^*$, $x^* \neq 0$. Since $x^* \neq 0$, we have from the equality in (12) that $\lambda_3 = 0$. Hence, (15) becomes

$$\lambda_2 z^* = \frac{2\beta}{1+\beta}. \tag{18}$$

On the other hand, multiplying both sides of the equality in (17) by z^*, using $(y^*)^2 = x^*z^*$ and the equality in (13), we get

$$\lambda_2(y^*)^2 = \frac{2(1+\beta)(\beta-\epsilon)}{(2+\beta+\epsilon)^2} z^*.$$

Hence,

$$z^* = \frac{(2+\beta+\epsilon)^2}{2(1+\beta)(\beta-\epsilon)} \lambda_2(y^*)^2. \tag{19}$$

Substituting the expression for z^* in (19) into (18), we can solve for $\lambda_2 y^*$ to be

$$\lambda_2 y^* = \pm \frac{2\sqrt{\beta(\beta-\epsilon)}}{2+\beta+\epsilon}.$$

Comparing the above expression for $\lambda_2 y^*$ with that in (16) leads to a contradiction. Hence, λ_0 cannot be nonzero.

Case 2: $\lambda_0 = 0$.

In this case, (8) becomes

$$\lambda_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -z^* \\ 2y^* \\ -x^* \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_5 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} = 0. \tag{20}$$

We immediately observe from (20) that $\lambda_1 = \lambda_5 = 0$. Also note that not all $\lambda_i = 0$ is equal to zero, $i = 2, 3$ and 4.

If $\lambda_2 \neq 0$. Then from the third "row" in (20), $x^* \geq 0$ and $\lambda_4 \geq 0$, we have that $x^* = \lambda_4 = 0$. It then follows from the last inequality in (11) that $y^* = 0$. Furthermore, from the first "row" in (20), we have $z^* = 0$ since $z^* \geq 0$ and $\lambda_3 \geq 0$.

Note that λ_3 has to be zero since if it is positive, then this leads to a contradiction in the first "row" of (20) as $z^* \geq 0$ and $\lambda_2 \geq 0$.

If $\lambda_4 \neq 0$. Then from the third "row" in (20), we have $-\lambda_2 x^* = \lambda_4 > 0$, which is impossible. Hence, $\lambda_4 = 0$.

Hence, in this case, we have $x^* = y^* = z^* = 0$, with $\lambda_2 > 0$ and $\lambda_1 = \lambda_3 = \lambda_4 = \lambda_5 = 0$. Finally, since $x_1^* \leq y_1^*$, we observe that the optimal value to the minimization problem (5)-(7) is greater than or equal to $\frac{2\beta}{1+\beta} x_1^* (\geq 0)$. In fact, it is equal to zero with $x_1^* = y_1^* = 0$.

In conclusion, the optimal value of the minimization problem (5)-(7) is zero with optimal solution $x^* = y^* = z^* = x_1^* = y_1^* = 0$. The proposition then follows. $\qquad\square$

3 Preliminaries to Convergence Analysis

We observe a few facts in this paragraph. Recall that if $\{x_k\}$ generated by the relaxed PR splitting method (2) converges to x^* for a given $\theta > 0$, and if we let $J_A(x^*) = u^*$, then $u^* \in \Re^n$ is a solution to (1), which is unique. Hence, there exists $z^* \in \Re^n$ such that $z^* \in A(u^*), -z^* \in B(u^*)$, since u^* is a solution to (1). Note that z^* that satisfies $z^* \in A(u^*), -z^* \in B(u^*)$ is unique if either A or B is single-valued. It is also easy to see in this case that $x^* = u^* + z^*$. We assume[1] $u^* = 0$ without loss of generality from now onwards. We can do this because by letting $A_1(\cdot) = A(\cdot + u^*)$ and $B_1(\cdot) = B(\cdot + u^*)$, we observe that $z^* \in A_1(0), -z^* \in B_1(0)$, and the relaxed PR splitting method (2) using (A, B) and using (A_1, B_1) generate sequence with corresponding terms in each sequence differing from each other by u^*.

Let $A_0 := A - \beta I$ and $B_0 := B - \beta I$. Then, A_0 and B_0 are maximal monotone operators from $\Re^n$ to $\Re^n$. In terms of A_0, B_0, (2) is given by

$$\bar{x}_k = \bar{x}_{k-1} + \frac{\theta}{1+\beta} \left(J_{\frac{1}{1+\beta} B_0} \left(\frac{2}{1+\beta} J_{\frac{1}{1+\beta} A_0}(\bar{x}_{k-1}) - \bar{x}_{k-1} \right) - J_{\frac{1}{1+\beta} A_0}(\bar{x}_{k-1}) \right), \quad k \geq 1, \quad (21)$$

where $\bar{x}_k := \frac{1}{1+\beta} x_k$. Letting $\bar{A}_0 := \frac{1}{1+\beta} A_0$ and $\bar{B}_0 := \frac{1}{1+\beta} B_0$, we can rewrite (21) as

$$\bar{x}_k = \bar{x}_{k-1} + \frac{\theta}{1+\beta} \left(J_{\bar{B}_0} \left(\frac{2}{1+\beta} J_{\bar{A}_0}(\bar{x}_{k-1}) - \bar{x}_{k-1} \right) - J_{\bar{A}_0}(\bar{x}_{k-1}) \right), \quad k \geq 1. \quad (22)$$

Note that $\bar{A}_0$ and $\bar{B}_0$ are maximal monotone operators, and that $\{x_k\}$ converges to x^* if and only if $\{\bar{x}_k\}$ converges to $\bar{x}^* := \frac{1}{1+\beta} x^*$. Furthermore, we have $J_{\bar{A}_0}(\bar{x}^*) = 0, J_{\bar{B}_0}(-\bar{x}^*) = 0$, that is,

$$\bar{x}^* \in \bar{A}_0(0), \quad -\bar{x}^* \in \bar{B}_0(0). \quad (23)$$

Note that suppose A or B is single-valued, then if there exists $\bar{x}^{**} \in \Re^n$ that satisfies (23), we have $\bar{x}^{**} = \bar{x}^*$.

---

[1]Note that because of this assumption on u^*, we have $x^* = z^*$, when A or B is single-valued.

We consider $\theta = 2 + \beta + \epsilon$, where $0 \le \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, in this paper.

Let $\bar{u}_k := J_{\bar{A}_0}(\bar{x}_k)$, $\bar{v}_k := J_{\bar{B}_0}\left(\frac{2}{1+\beta}\bar{u}_k - \bar{x}_k\right)$. Hence,

$$
\begin{align}
\bar{x}_k - \bar{u}_k &\in \bar{A}_0(\bar{u}_k), \tag{24}\\
\bar{w}_k &\in \bar{B}_0(\bar{v}_k), \tag{25}
\end{align}
$$

where

$$
\bar{w}_k := \frac{2}{1+\beta}\bar{u}_k - \bar{x}_k - \bar{v}_k. \tag{26}
$$

By monotonicity of $\bar{A}_0$ and $\bar{B}_0$, for $k \ge 0$, we have by (23) - (25) that

$$
\begin{align}
\langle (\bar{x}_k - \bar{x}^*) - \bar{u}_k, \bar{u}_k \rangle &\ge 0, \tag{27}\\
\langle \bar{w}_k + \bar{x}^*, \bar{v}_k \rangle &\ge 0. \tag{28}
\end{align}
$$

Observe from (27) that

$$
\langle \bar{x}_k - \bar{x}^*, \bar{u}_k \rangle \ge \|\bar{u}_k\|^2. \tag{29}
$$

We can also write (28), using (26), as

$$
\langle \bar{x}_k - \bar{x}^*, \bar{v}_k \rangle \le \frac{2}{1+\beta}\langle \bar{u}_k, \bar{v}_k \rangle - \|\bar{v}_k\|^2. \tag{30}
$$

Inequalities (29) and (30) play important roles to arrive at the convergence and convergence rates results in this paper.

We end this section with the following, which we state without proof:

Proposition 3.1 *For all $k \ge 1$, we have $x_k - x_{k-1} = (1+\beta)(\bar{x}_k - \bar{x}_{k-1}) = \theta(\bar{v}_{k-1} - \bar{u}_{k-1})$.*

4 Convergence of $\{x_k\}$

We begin the section by stating the convention that the ith component of $\bar{x}_k, \bar{u}_k, \bar{v}_k$ is to be written as $\bar{x}_k^i, \bar{u}_k^i$ and $\bar{v}_k^i$ respectively, while the ith component of $\bar{x}^*$ is denoted by $\bar{x}_i^*$. We use the same convention for the ith component of x_k, u_k, v_k and x^*. Using this convention, we state a technical assumption on $\bar{x}_k$ that is to apply to the whole section:

Assumption 4.1 *For all $i = 1, \ldots, n$ and for all $k \ge K_0$, where $K_0 \ge 0$, we have $\bar{x}_k^i \ne \bar{x}_i^*$.*

The above assumption makes the analysis in this section possible.

Remark 4.2 *For $n = 1$, if Assumption 4.1 does not hold, that is, there exists $k_0 \ge K_0$ such that $\bar{x}_{k_0} = \bar{x}^*$. Then it is easy to see that for all $k \ge k_0$, we have $\bar{x}_k = \bar{x}^*$. Hence, not having Assumption 4.1 when $n = 1$ leads to a trivial situation.*

For $k \ge 1$, let us write $\bar{x}_k$ as $\alpha_{k-1} \odot (\bar{x}_{k-1} - \bar{x}^*) + \bar{x}^*$, where $\alpha_{k-1} \in (\Re \cup \{\infty\})^n$ (with the understanding that $\alpha_{k-1}^i = \infty$ if and only if $\bar{x}_{k-1}^i = \bar{x}_i^*$. Here, α_{k-1}^i is the ith component of

α_{k-1}.) From (22) and with $\theta = 2+\beta+\epsilon$, $\bar{v}_k$ defined by $J_{\bar{B}_0}\left(\frac{2}{1+\beta}\bar{u}_k - \bar{x}_k\right)$, where $\bar{u}_k = J_{\bar{A}_0}(\bar{x}_k)$, can then be written as

$$\bar{v}_k = \frac{1+\beta}{2+\beta+\epsilon}(\alpha_k - e) \odot (\bar{x}_k - \bar{x}^*) + \bar{u}_k, \tag{31}$$

and, by (31), $\bar{w}_k$ defined by (26) can be written as

$$\bar{w}_k = \frac{1-\beta}{1+\beta}\bar{u}_k - \bar{x}_k - \frac{1+\beta}{2+\beta+\epsilon}(\alpha_k - e) \odot (\bar{x}_k - \bar{x}^*). \tag{32}$$

Using (31), (32), we can write (28) in the following way:

$$\sum_{i=1}^{n}\left[-\left(1 + \frac{2\beta(\alpha_k^i - 1)}{2+\beta+\epsilon}\right)\bar{u}_k^i(\bar{x}_k^i - \bar{x}_i^*)\right.$$
$$\left. -\left(\frac{(1+\beta)(\alpha_k^i - 1)}{2+\beta+\epsilon}\right)\left(1 + \frac{(1+\beta)(\alpha_k^i - 1)}{2+\beta+\epsilon}\right)(\bar{x}_k^i - \bar{x}_i^*)^2\right] + \frac{1-\beta}{1+\beta}\|\bar{u}_k\|^2 \geq 0, \tag{33}$$

where $0 < \epsilon < \min\{\beta, 1/\beta\}$.

Remark 4.3 *With $\bar{x}_k$ given by $\alpha_{k-1} \odot (\bar{x}_{k-1} - \bar{x}^*) + \bar{x}^*$, we deduce from Assumption 4.1 that for all $i = 1, \ldots, n$ and for all $k \geq K_0$, we have $\alpha_k^i \neq 0$ or ∞. WLOG, we let K_0 in Assumption 4.1 to be equal to zero from now onwards. When $n \geq 2$, if we allow $\alpha_k^i = 0$ or ∞ for some $i = 1, \ldots, n$, and for infinity many k, then our approach to proving convergence of iterates generated by the relaxed PR splitting method (2) for $\theta = 2+\beta+\epsilon$, where $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, in this section, has to be further refined, and may not be able to carry through. When $n = 1$, however, having $\alpha_k^1 = 0$ or ∞ for some $k \geq 0$ leads trivially to $\bar{x}_j = \bar{x}^*$ for all $j \geq k+1$.*

The next two lemmas enable us to prove Theorem 4.6 on the convergence behavior of $\{x_k\}$.

Lemma 4.4 *Given $\bar{x}_k = \alpha_{k-1} \odot (\bar{x}_{k-1} - \bar{x}^*) + \bar{x}^*$, where $\{\bar{x}_k\}$ is generated by (22) with $\theta = 2+\beta+\epsilon$, $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, and $\bar{x}_0 \in \Re^n$, we have $\{\bar{x}_k\}$ is bounded. Furthermore, for each $i = 1, \ldots, n$, if $\{\bar{x}_{j_k}^i\}$ is a convergent sequence of $\{\bar{x}_k^i\}$, then if $\lim_{k\to\infty} \bar{x}_{j_k}^i \neq \bar{x}_i^*$, we have $\lim_{k\to\infty} |\alpha_k^i| = 1$.*

Proof: Suppose $\{\bar{x}_k\}$ is unbounded. For each $i = 1, \ldots, n$, and $N \geq 1$, let $I_N^i := \{k \geq 0 ; |\bar{x}_k^i| \geq N\}$. Since we assume $\{\bar{x}_k\}$ is unbounded, the set $U \subseteq \{1, \ldots, n\}$ defined to be such that for each $i \in U$, $\bigcap_{N=1}^{\infty} I_N^i \neq \emptyset$, is nonempty.

We observe that in order for (29), (33) to hold, there must exist $i_0 \in U$ and $N_0 \geq 1$ such that for all $k \in I_{N_0}^{i_0}$, we have

$$(\bar{x}_k^{i_0} - \bar{x}_{i_0}^*)\bar{u}_k^{i_0} \geq \|\bar{u}_k\|^2,$$

$$-\left(1 + \frac{2\beta(\alpha_k^{i_0} - 1)}{2+\beta+\epsilon}\right)\bar{u}_k^{i_0}(\bar{x}_k^{i_0} - \bar{x}_{i_0}^*)$$
$$-\left(\frac{(1+\beta)(\alpha_k^{i_0} - 1)}{2+\beta+\epsilon}\right)\left(1 + \frac{(1+\beta)(\alpha_k^{i_0} - 1)}{2+\beta+\epsilon}\right)(\bar{x}_k^{i_0} - \bar{x}_{i_0}^*)^2 + \frac{1-\beta}{1+\beta}\|\bar{u}_k\|^2 \geq 0.$$

From the above two inequalities, by Proposition 2.2, where we let $x = (\bar{x}_k^{i_0} - \bar{x}_{i_0}^*)/\|\bar{u}_k\|, y = \bar{u}_k^{i_0}/\|\bar{u}_k\|, \alpha = \alpha_k^{i_0}$ in the proposition, we have $|\alpha_k^{i_0}| \leq 1$ for all $k \in I_{N_0}^{i_0}$. We see that this implies

8

that $\{\bar{x}_k^{i_0}\}$ is bounded. But this is a contradiction to $i_0 \in U$, that is, $\bigcap_{N=1}^{\infty} I_N^{i_0} \neq \emptyset$. Hence, $\{\bar{x}_k\}$ is bounded.

We now show that for $i = 1, \ldots, n$, if $\{\bar{x}_{j_k}^i\}$ is a convergent subsequence of $\{\bar{x}_k^i\}$ with limit point $\bar{x}_i^{**}$, which is not equal to $\bar{x}_i^*$, then $\lim_{k \to \infty} |\alpha_k^i| = 1$. We have

$$\bar{x}_{j_k}^i - \bar{x}_i^* = \Pi_{0 \leq j \leq j_k - 1} \alpha_j^i (\bar{x}_0^i - \bar{x}_i^*).$$

Hence,

$$\log \left| \frac{\bar{x}_{j_k}^i - \bar{x}_i^*}{\bar{x}_0^i - \bar{x}_i^*} \right| = \sum_{0 \leq j \leq j_k - 1} \log |\alpha_j^i|,$$

which holds since $\alpha_j^i \neq \infty$, due to Assumption 4.1. Since $\{\bar{x}_{j_k}^i\}$ converges to $\bar{x}_i^{**}$, which is not equal to $\bar{x}_i^*$, this implies that $\left\{ \sum_{0 \leq j \leq j_k - 1} \log |\alpha_j^i| \right\}$ converges, which further implies that $\lim_{k \to \infty} \log |\alpha_k^i| = 0$. That is, $\lim_{k \to \infty} |\alpha_k^i| = 1$. $\qquad \square$

Lemma 4.5 *Given* $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_n)^T, \bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T, \bar{\bar{x}} = (\bar{\bar{x}}_1, \ldots, \bar{\bar{x}}_n)^T \in \Re^n$ *and* $0 < \epsilon < \min\{\beta, 1/\beta\}$. *Suppose* $\langle \bar{u}, \bar{\bar{x}} - \bar{x} \rangle \geq \|\bar{u}\|^2$ *and*

$$\langle \bar{\bar{x}} - \bar{x}, \bar{u} \rangle \leq \frac{1 - \beta}{1 + \beta} \|\bar{u}\|^2 + \sum_{i \in I} \left[\left(\frac{4\beta}{2 + \beta + \epsilon} \right) \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i) + \frac{2(1 + \beta)(\epsilon - \beta)}{(2 + \beta + \epsilon)^2} (\bar{\bar{x}}_i - \bar{x}_i)^2 \right], \quad (34)$$

where $I \subseteq \{1, \ldots, n\}$. *Then* $\bar{u} = 0$, *and* $\bar{x}_i = \bar{\bar{x}}_i$ *for* $i \in I$.

Proof: From (34), we have

$$
\begin{aligned}
0 \quad \leq \quad & -\sum_{i \notin I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i) + \frac{1 - \beta}{1 + \beta} \sum_{i \notin I} \bar{u}_i^2 + \frac{1 - \beta}{1 + \beta} \sum_{i \in I} \bar{u}_i^2 + \frac{3\beta - \epsilon - 2}{2 + \beta + \epsilon} \sum_{i \in I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i) \\
& + \frac{2(1 + \beta)(\epsilon - \beta)}{(2 + \beta + \epsilon)^2} \sum_{i \in I} (\bar{\bar{x}}_i - \bar{x}_i)^2.
\end{aligned}
\quad (35)
$$

From $\langle \bar{u}, \bar{\bar{x}} - \bar{x} \rangle \geq \|\bar{u}\|^2$, we have

$$0 \leq \sum_{i \in I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i) - \sum_{i \in I} \bar{u}_i^2 + \sum_{i \notin I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i) - \sum_{i \notin I} \bar{u}_i^2. \quad (36)$$

The lemma follows from Proposition 2.3 by letting $x = \sum_{i \in I} \bar{u}_i^2, y = \sum_{i \in I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i), z = \sum_{i \in I} (\bar{\bar{x}}_i - \bar{x}_i)^2, x_1 = \sum_{i \notin I} \bar{u}_i^2$ and $y_1 = \sum_{i \notin I} \bar{u}_i (\bar{\bar{x}}_i - \bar{x}_i)$ in the proposition. Note that the condition $y^2 - xz \leq 0$ in the proposition holds in this case by the Cauchy-Schwartz inequality. $\square$

We have the following theorem:

Theorem 4.6 *For* $\theta = 2 + \beta + \epsilon$, *where* $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, *let* x^{**} *be an accumulation point of* $\{x_k\}$, *where the sequence* $\{x_k\}$ *is generated by the relaxed PR splitting method (2) for a given initial iterate* $x_0 \in \Re^n$. *Then* $J_A(x^{**})$ *solves the monotone inclusion problem (1)*.

Proof: We only need to prove that for $\theta = 2 + \beta + \epsilon$, where $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, any convergent subsequence $\{\bar{x}_{j_k}\}$ of $\{\bar{x}_k\}$, the latter generated using (22) when $\bar{x}_0 = \frac{1}{1+\beta} x_0$, has its limit point $\bar{x}^{**}$ that satisfies (23).

WLOG, assume that $\{\bar{x}_{j_k-1}\}$ converges as well, and let the sequence converges to $\bar{x}^{***}$, for some $\bar{x}^{***} \in \Re^n$. We have $\{\bar{u}_{j_k-1}\}$, where $\bar{u}_{j_k-1} = J_{\bar{A}_0}(\bar{x}_{j_k-1})$, also converges, and let $\bar{u}^{***}$ be its limit point. By Lemma 4.4, for $i = 1, \ldots, n$, we either have $\lim_{k\to\infty} \bar{x}^i_{j_k-1} = \bar{x}^*_i$ or $\lim_{k\to\infty} |\alpha^i_k| = 1$. Let $I := \{i \; ; \; 1 \le i \le n, \lim_{k\to\infty} \alpha^i_{j_k-1} = -1\}$ and $I_1 := \{i \; ; \; 1 \le i \le n, \lim_{k\to\infty} \alpha^i_{j_k-1} = 1\}$. Observe that for $i \in I_1$, $\bar{x}^{**}_i = \bar{x}^{***}_i$, and for $i \notin I \cup I_1$, we have $\lim_{k\to\infty} \bar{x}^i_{j_k-1} = \bar{x}^*_i$. Since $i \notin I \cup I_1$, we must also have $\lim_{k\to\infty} \bar{x}^i_{j_k} = \bar{x}^*_i$, which we can show by contradiction. Hence, for $i \notin I \cup I_1$, $\bar{x}^{**}_i = \bar{x}^{***}_i = \bar{x}^*_i$.

Consider the following expression, which has a similar expression in (33):

$$-\left(1 + \frac{2\beta(\alpha^i_{j_k-1} - 1)}{2 + \beta + \epsilon}\right) \bar{u}^i_{j_k-1}(\bar{x}^i_{j_k-1} - \bar{x}^*_i)$$
$$-\left(\frac{(1+\beta)(\alpha^i_{j_k-1} - 1)}{2 + \beta + \epsilon}\right)\left(1 + \frac{(1+\beta)(\alpha^i_{j_k-1} - 1)}{2 + \beta + \epsilon}\right)(\bar{x}^i_{j_k-1} - \bar{x}^*_i)^2. \tag{37}$$

For $i \in I$, as $k \to \infty$, the expression (37) converges to

$$\left(-1 + \frac{4\beta}{2 + \beta + \epsilon}\right)\bar{u}^{***}_i(\bar{x}^{***}_i - \bar{x}^*_i) - \frac{2(1+\beta)(\beta - \epsilon)}{(2 + \beta + \epsilon)^2}(\bar{x}^{***}_i - \bar{x}^*_i)^2. \tag{38}$$

For $i \in I_1$, as $k \to \infty$, it converges to

$$\bar{u}^{***}_i(\bar{x}^*_i - \bar{x}^{***}_i). \tag{39}$$

Finally, for $i \notin I \cup I_1$, it converges to 0.

Hence, (33) leads to the following:

$$\sum_{i \in I}\left[\left(-1 + \frac{4\beta}{2 + \beta + \epsilon}\right)\bar{u}^{***}_i(\bar{x}^{***}_i - \bar{x}^*_i) - \frac{2(1+\beta)(\beta - \epsilon)}{(2 + \beta + \epsilon)^2}(\bar{x}^{***}_i - \bar{x}^*_i)^2\right]$$
$$+ \sum_{i \in I_1} \bar{u}^{***}_i(\bar{x}^* - \bar{x}^{***}) + \frac{1 - \beta}{1 + \beta}\|\bar{u}^{***}\|^2 \; \ge \; 0.$$

That is,

$$\langle \bar{u}^{***}, \bar{x}^{***} - \bar{x}^* \rangle \; \le \; \frac{1 - \beta}{1 + \beta}\|\bar{u}^{***}\|^2 + \sum_{i \in I}\left[\frac{4\beta}{2 + \beta + \epsilon}\bar{u}^{***}_i(\bar{x}^{***}_i - \bar{x}^*_i)\right.$$
$$\left. + \frac{2(1+\beta)(\epsilon - \beta)}{(2 + \beta + \epsilon)^2}(\bar{x}^{***}_i - \bar{x}^*_i)^2\right]. \tag{40}$$

Furthermore, by (29), the inequality $\|\bar{u}_{j_k-1}\|^2 \le \langle \bar{x}_{j_k-1} - \bar{x}^*, \bar{u}_{j_k-1} \rangle$ holds, and it tends to

$$\|\bar{u}^{***}\|^2 \le \langle \bar{u}^{***}, \bar{x}^{***} - \bar{x}^* \rangle, \tag{41}$$

as $k \to \infty$. With (40), (41), using Lemma 4.5, where we let $\bar{x} = \bar{x}^*, \bar{\bar{x}} = \bar{x}^{***}, \bar{u} = \bar{u}^{***}$ in the lemma, we have $\bar{u}^{***} = 0$ and for $i \in I$, $\bar{x}^{***}_i = \bar{x}^*_i$. Hence, for $i \in I$, we have $\bar{x}^{***}_i = \bar{x}^{**}_i$.

10

In summary, for $i \in I$, $\bar{x}_i^{***} = \bar{x}_i^{**} = \bar{x}_i^*$, for $i \in I_1$, $\bar{x}_i^{***} = \bar{x}_i^{**}$, while for $i \notin I \cup I_1$, $\bar{x}_i^{***} = \bar{x}_i^{**} = \bar{x}_i^*$. Furthermore, $J_{\bar{A}_0}(\bar{x}^{***}) = \bar{u}^{***} = 0$. By taking limit in (22), where k in each subscript in (22) is replaced by j_k, we hence have $J_{\bar{B}_0}(-\bar{x}^{***}) = 0$. Since $J_{\bar{A}_0}(\bar{x}^{***}) = J_{\bar{B}_0}(-\bar{x}^{***}) = 0$, (23) is satisfied by $\bar{x}^{***}$, and therefore satisfied by $\bar{x}^{**}$. $\qquad\square$

The following corollary follows immediately from the above theorem:

Corollary 4.7 *Suppose A or B is single-valued, then for $\theta = 2 + \beta + \epsilon$, where $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, the sequence $\{x_k\}$ generated by the relaxed PR splitting method (2) for a given initial iterate $x_0 \in \Re^n$ is a convergent sequence with limit point x^*, and $J_A(x^*)$ solves the monotone inclusion problem (1).*

Proof: We only need to prove that for $\theta = 2 + \beta + \epsilon$, where $0 < \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$, $\{\bar{x}_k\}$ converges to $\bar{x}^*$. Since, by Lemma 4.4, we know that $\{\bar{x}_k\}$ is bounded, this is achieved by showing that any convergent subsequence $\{\bar{x}_{j_k}\}$ of $\{\bar{x}_k\}$ converges to $\bar{x}^*$. By Theorem 4.6, we see that the limit point of $\{\bar{x}_{j_k}\}$ satisfies (23). Since A or B is single-valued, which implies that either $\bar{A}_0$ or $\bar{B}_0$ is single-valued, therefore, $\bar{x}^*$ is the only solution to (23). Hence, the limit point of $\{\bar{x}_{j_k}\}$ is $\bar{x}^*$. $\qquad\square$

Remark 4.8 *We remark that when $n = 1$, for convergence of $\{x_k\}$ given any initial iterate $x_0 \in \Re$, we do not require Assumption 4.1 and only need A, B to be maximal β-strongly monotone, when $\theta = 2 + \beta + \epsilon$, where $0 < \epsilon < \min\{\beta, 1/\beta\}, \beta > 0$ [23]. Hence, Assumption 4.1 is only required when n is greater than or equal to 2 to show convergence of $\{x_k\}$. We believe that this assumption is always satisfied[2], although it is an open problem to show this for all practical problems. Our approach to show convergence arises from the idea to solve a higher dimensional problem by "reducing" it to a smaller dimensional problem ($n = 1$), and then applying our solution method when $n = 1$ [23], where we have convergence. In the process of doing this, we realize that we need Assumption 4.1 in order to show convergence in higher dimensions. In fact, there may exist other approaches that do not require this assumption to show convergence, and that convergence of iterates occurs over the considered range of θ for A and B maximal β-strongly monotone, and either of them single-valued.*

5 Results on Convergence Rates

We start the section by relating $\bar{u}_k$ and $\bar{v}_k$ in the following way:

$$\bar{v}_k = -\gamma_k \bar{u}_k + \tau_k \bar{z}_k, \text{ where } \|\bar{z}_k\| = 1 \text{ and } \langle \bar{u}_k, \bar{z}_k \rangle = 0. \tag{42}$$

Recall that $\bar{u}_k = J_{\bar{A}_0}(\bar{x}_k)$ and $\bar{v}_k = J_{\bar{B}_0}\left(\frac{2}{1+\beta}\bar{u}_k - \bar{x}_k\right)$. Note that there always exist γ_k, τ_k and $\bar{z}_k$ such that $\bar{v}_k = -\gamma_k \bar{u}_k + \tau_k \bar{z}_k$, $\|\bar{z}_k\| = 1$ and $\langle \bar{u}_k, \bar{z}_k \rangle = 0$ in (42). Also, γ_k can be computed from $\bar{u}_k, \bar{v}_k$, and is equal to $-\langle \bar{u}_k, \bar{v}_k \rangle / \|\bar{u}_k\|^2$. We have the following definition:

$$\bar{\gamma} := \limsup_{k \to \infty} \gamma_k. \tag{43}$$

Furthermore, let $\bar{\gamma}_{\max}$ be the maximum of $\bar{\gamma}$ over all $\bar{x}_0 \in \Re^n$. Note that $\bar{\gamma}_{\max}$ is only dependent on problem data and parameters. It does not depend on the initial iterate and subsequent iterates generated.

---

[2]Preliminary numerical evidence that assumption always holds is provided in Section 6.

To arrive at our convergence rates results, we observe from (29), (30) that

$$\langle \bar{u}_{k-1}, \bar{x}_{k-1} - \bar{u}_{k-1} - \bar{x}^* \rangle + \left\langle \bar{v}_{k-1}, \frac{2}{1+\beta} \bar{u}_{k-1} - \bar{x}_{k-1} - \bar{v}_{k-1} + \bar{x}^* \right\rangle \geq 0. \tag{44}$$

The following lemma enables us to arrive at Lemma 5.2, which is key to proving our convergence rates results in Theorem 5.4.

Lemma 5.1 *For $k \geq 1$, we have*

$$(1-\theta)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 \leq \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2)$$
$$+ (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - \beta(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2).(45)$$

Proof: We have

$$\frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2) + (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle$$
$$- \beta(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) - (1-\theta)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2$$
$$= \frac{(1+\beta)^2}{\theta}\left(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \left\|(\bar{x}_{k-1} - \bar{x}^*) + \frac{\theta}{1+\beta}(\bar{v}_{k-1} - \bar{u}_{k-1})\right\|^2\right)$$
$$+ (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - \beta(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) - (1-\theta)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2$$
$$= \frac{(1+\beta)^2}{\theta}\left(-\frac{2\theta}{1+\beta}\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - \frac{\theta^2}{(1+\beta)^2}\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2\right)$$
$$+ (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - \beta(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) - (1-\theta)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2$$
$$= -(1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - \beta(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) - \|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2$$
$$= (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{u}_{k-1} \rangle - (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} \rangle - (1+\beta)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2)$$
$$+ 2\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle$$
$$\geq 0,$$

where the first equality follows from Proposition 3.1, and the inequality follows from (44). □

The following lemma follows from the above lemma.

Lemma 5.2 *For $k \geq 1$, we have*

$$4\beta\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 \leq \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2), \tag{46}$$

where $\theta = 2 + \beta + \epsilon$, $0 \leq \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$.

Proof: By expanding the left-hand side of (45) in Lemma 5.1, using $\theta = 2 + \beta + \epsilon$, and upon algebraic manipulations, we obtain the following:

$$-(1+\epsilon)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) \leq \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2)$$
$$+ (1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - 2(1+\beta+\epsilon)\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle. \tag{47}$$

Observe that

$$(1+\beta)\langle \bar{x}_{k-1} - \bar{x}^*, \bar{v}_{k-1} - \bar{u}_{k-1} \rangle - 2(1+\beta+\epsilon)\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle$$

$$\leq \quad -2(\beta + \epsilon)\langle \bar{u}_{k-1} \, \bar{v}_{k-1}\rangle - (1+\beta)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2),$$

where the inequality follows from (29) and (30). Substituting the above inequality into (47) and upon algebraic manipulations, we have

$$2(\beta + \epsilon)\langle \bar{u}_{k-1} \, \bar{v}_{k-1}\rangle + (\beta - \epsilon)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) \leq \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2).$$

The lemma then follows from the above by observing that $\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2 = \|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 + 2\langle \bar{u}_{k-1}, \bar{v}_{k-1}\rangle$. $\qquad\square$

We need the following lemma to prove results on the R-linear convergence rate of $\|x_k - x^*\|$ in Theorem 5.4:

Lemma 5.3 *Suppose A or B is Lipschitz continuous with Lipschitz constant $L(\geq 0)$, then for all $k \geq 0$,*

$$\|\bar{x}_k - \bar{x}^*\| \leq \bar{L}(\|\bar{u}_k\| + \|\bar{v}_k\|), \tag{48}$$

where $\bar{L} := \max\left\{\frac{L}{1+\beta} + 1, \frac{2}{1+\beta}\right\}$.

Proof: Since A or B is Lipschitz continuous with Lipschitz continuous $L(\geq 0)$, it is clear that either $\bar{A}_0$ or $\bar{B}_0$ is also Lipschitz continuous with Lipschitz constant $\frac{L}{1+\beta}$.

Suppose $\bar{A}_0$ is Lipschitz continuous. We have from (23) and (24) that

$$\|\bar{x}_k - \bar{u}_k - \bar{x}^*\| \leq \frac{L}{1+\beta}\|\bar{u}_k\|.$$

It then follows that (48) holds for all $k \geq 0$ from the above by the triangle inequality.

Suppose $\bar{B}_0$ is Lipschitz continuous. Then, from (23) and (25), using (26), the following holds:

$$\left\|\frac{2}{1+\beta}\bar{u}_k - \bar{v}_k - (\bar{x}_k - \bar{x}^*)\right\| \leq \frac{L}{1+\beta}\|\bar{v}_k\|. \tag{49}$$

The result then follows by the triangle inequality. $\qquad\square$

Theorem 5.4 *Let $\{x_k\}$ be generated by the relaxed PR splitting method (2) given an initial iterate $x_0 \in \Re^n$, where $\theta = 2 + \beta + \epsilon$, $0 \leq \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$. For*

$$\bar{\gamma}_{\max} < \frac{\sqrt{\beta} - \sqrt{\epsilon}}{\sqrt{\beta} + \sqrt{\epsilon}} \tag{50}$$

there exists $i \leq k$ and $i \geq k_0$, where $k \geq 2k_0$, and

$$M := \min\left\{M_1(\beta - \epsilon), \frac{f(\bar{\gamma}_{\max})(1 - M_1)(\beta - \epsilon)}{8(\beta + \epsilon)\max\{\bar{\gamma}, 1\}}\right\} > 0, \tag{51}$$

where $f(\bar{\gamma}_{\max}) = -2(\beta + \epsilon)\bar{\gamma}_{\max} + (\beta - \epsilon)(1 + \bar{\gamma}_{\max}^2) > 0$ and $0 < M_1 < 1$, such that

$$\|x_i - x_{i-1}\| \leq \left(\frac{2\sqrt{\theta}\|x_{k_0-1} - x^*\|}{\sqrt{M}}\right)\frac{1}{\sqrt{k}}.$$

13

Here, $k_0 \geq 1$ depends only on problem data and parameters. Furthermore, if A or B is Lipschitz continuous with Lipschitz constant $L(\geq 0)$, then for all $k \geq k_0$,

$$\|x_k - x^*\| \leq \left[\sqrt{1 - \frac{M\theta}{2(1+\beta)^2 \bar{L}^2}} \right] \|x_{k-1} - x^*\|,$$

where $\bar{L}$ is defined in Lemma 5.3.

Proof: Given that $\bar{v}_{k-1} = -\gamma_{k-1}\bar{u}_{k-1} + \tau_{k-1}\bar{z}_{k-1}$, where $\|\bar{z}_{k-1}\| = 1$, $\langle \bar{u}_{k-1}, \bar{z}_{k-1} \rangle = 0$, we have

$$
\begin{aligned}
& \frac{4\beta\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle}{\|\bar{u}_{k-1}\|^2} + \frac{(\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2}{\|\bar{u}_{k-1}\|^2} \\
&= \frac{2(\beta + \epsilon)\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle}{\|\bar{u}_{k-1}\|^2} + \frac{(\beta - \epsilon)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2)}{\|\bar{u}_{k-1}\|^2} \\
&= -2(\beta + \epsilon)\gamma_{k-1} + (\beta - \epsilon)\left[1 + \gamma_{k-1}^2 + \frac{\tau_{k-1}^2}{\|\bar{u}_{k-1}\|^2} \right].
\end{aligned}
\tag{52}
$$

Let $M_1 \in (0, 1)$ be given. Consider

$$I := \left\{ k \geq 1 \; ; \; -2(\beta + \epsilon)\gamma_{k-1} + (1 - M_1)(\beta - \epsilon)\left[1 + \gamma_{k-1}^2 + \frac{\tau_{k-1}^2}{\|\bar{u}_{k-1}\|^2} \right] \geq 0 \right\}.$$

Therefore, from (52), we see that for $k \in I$, we have

$$
\begin{aligned}
M_1(\beta - \epsilon)(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) &= M_1(\beta - \epsilon)((1 + \gamma_{k-1}^2)\|\bar{u}_{k-1}\|^2 + \tau_{k-1}^2) \\
&\leq 4\beta\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2.
\end{aligned}
\tag{53}
$$

Let us now consider $k \notin I$. From (52), we have

$$\frac{4\beta\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle}{\|\bar{u}_{k-1}\|^2} + \frac{(\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2}{\|\bar{u}_{k-1}\|^2} \geq -2(\beta + \epsilon)\gamma_{k-1} + (\beta - \epsilon)(1 + \gamma_{k-1}^2). \tag{54}$$

Let

$$f(\gamma) := -2(\beta + \epsilon)\gamma + (\beta - \epsilon)(1 + \gamma^2).$$

We have $f(\bar{\gamma}_{\max}) > 0$ if $\bar{\gamma}_{\max}$ satisfies (50). We see that since $\limsup_{k \to \infty} \gamma_k = \bar{\gamma} \leq \bar{\gamma}_{\max}$, for k sufficiently large, say $k \geq k_0$, we have $-2(\beta + \epsilon)\gamma_{k-1} + (\beta - \epsilon)(1 + \gamma_{k-1}^2) \geq f(\bar{\gamma}_{\max})/2 > 0$, and hence from (54), we have

$$4\beta\langle \bar{u}_{k-1}, \bar{v}_{k-1} \rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 \geq \frac{f(\bar{\gamma}_{\max})}{2}\|\bar{u}_{k-1}\|^2. \tag{55}$$

For $k \notin I$ and $k \geq k_0$, by definition of I, we have that

$$\frac{\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2}{\|\bar{u}_{k-1}\|^2} = 1 + \gamma_{k-1}^2 + \frac{\tau_{k-1}^2}{\|\bar{u}_{k-1}\|^2} \leq \frac{2(\beta + \epsilon)\gamma_{k-1}}{(1 - M_1)(\beta - \epsilon)} \leq \frac{4(\beta + \epsilon)\max\{\bar{\gamma}, 1\}}{(1 - M_1)(\beta - \epsilon)},$$

Therefore, from (55) which holds when $k \geq k_0$, for $k \notin I$, we have

$$\frac{f(\bar{\gamma}_{\max})(1 - M_1)(\beta - \epsilon)}{8(\beta + \epsilon)\max\{\bar{\gamma}, 1\}}(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2)$$

14

$$\leq \quad \frac{f(\bar{\gamma}_{\max})}{2}\|\bar{u}_{k-1}\|^2 \leq 4\beta\langle\bar{u}_{k-1}, \bar{v}_{k-1}\rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2. \tag{56}$$

Combining (53) and (56), where the latter holds when $k \geq k_0$, then

$$M(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) \leq 4\beta\langle\bar{u}_{k-1}, \bar{v}_{k-1}\rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2, \tag{57}$$

where M is defined by (51). Hence, from (46) in Lemma 5.2, we have, for $k \geq k_0$,

$$M(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) \leq \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2). \tag{58}$$

Therefore, using $\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 \leq 2(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2)$, Proposition 3.1 and the definitions of $\bar{x}_k, \bar{x}^*$, we have from (58)

$$(k - k_0) \inf_{k_0 \leq i \leq k} \|x_i - x_{i-1}\|^2 \leq \theta^2 \sum_{i=k_0}^{k} \|\bar{u}_{i-1} - \bar{v}_{i-1}\|^2 \quad \leq \quad \frac{2\theta(1+\beta)^2}{M}(\|\bar{x}_{k_0-1} - \bar{x}^*\|^2)$$

$$= \quad \frac{2\theta\|x_{k_0-1} - x^*\|^2}{M}.$$

Hence, for $k \geq 2k_0$,

$$\inf_{k_0 \leq i \leq k} \|x_i - x_{i-1}\|^2 \leq \frac{2\theta\|x_{k_0-1} - x^*\|^2}{M(k - k_0)} \leq \frac{4\theta\|x_{k_0-1} - x^*\|^2}{Mk}.$$

The first result in the theorem then follows from the above.

Suppose A or B is Lipschitz continuous with Lipschitz constant L, then when $k \geq k_0$, we have

$$\begin{aligned} \frac{M}{2\bar{L}^2}\|\bar{x}_{k-1} - \bar{x}^*\|^2 &\leq \quad \frac{M}{2}(\|\bar{u}_{k-1}\| + \|\bar{v}_{k-1}\|)^2 \leq M(\|\bar{u}_{k-1}\|^2 + \|\bar{v}_{k-1}\|^2) \\ &\leq \quad 4\beta\langle\bar{u}_{k-1}, \bar{v}_{k-1}\rangle + (\beta - \epsilon)\|\bar{u}_{k-1} - \bar{v}_{k-1}\|^2 \\ &\leq \quad \frac{(1+\beta)^2}{\theta}(\|\bar{x}_{k-1} - \bar{x}^*\|^2 - \|\bar{x}_k - \bar{x}^*\|^2), \end{aligned}$$

where the first inequality follows from (48) in Lemma 5.3, the third inequality follows from (57), which holds when $k \geq k_0$, and the last inequality follows from (46) in Lemma 5.2. The second result in the theorem then follows from the definitions of $\bar{x}_k, \bar{x}^*$, and the above. $\qquad\square$

Corollary 5.5 *Let $\{x_k\}$ be generated by the relaxed PR splitting method (2) given an initial iterate $x_0 \in \Re^n$, where $\theta = 2 + \beta + \epsilon$, $0 \leq \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$. Suppose A is Lipschitz continuous with Lipschitz constant $L(> 0)$ such that*

$$L < 2\sqrt{2(1+\beta)}\sqrt{\frac{\sqrt{\beta} - \sqrt{\epsilon}}{\sqrt{\beta} + \sqrt{\epsilon}}} - (1+\beta). \tag{59}$$

In particular, when we choose ϵ to be zero (that is, $\theta = 2 + \beta$), we require

$$L < 2\sqrt{2(1+\beta)} - (1+\beta),$$

where we need $\beta < 7$ for $2\sqrt{2(1+\beta)} - (1+\beta)$ to be positive. Then there exists $i \leq k$ and $i \geq k_0$, where $k \geq 2k_0$, and M given by (51), such that

$$\|x_i - x_{i-1}\| \leq \left(\frac{2\sqrt{\theta}\|x_{k_0-1} - x^*\|}{\sqrt{M}}\right)\frac{1}{\sqrt{k}},$$

and for all $k \geq k_0$,

$$\|x_k - x^*\| \leq \left[\sqrt{1 - \frac{M\theta}{2(1 + \beta)^2 \bar{L}^2}}\right] \|x_{k-1} - x^*\|,$$

where $\bar{L}$ is defined in Lemma 5.3. Here, $k_0 \geq 1$ depends only on problem data and parameters.

Proof: Given that A is Lipschitz continuous with Lipschitz constant L that satisfies (59). The corollary is proved by showing that $\bar{\gamma}_{\max}$ satisfies (50). Since A is Lipschitz continuous with Lipschitz continuous L, then $\bar{A}_0$ is also Lipschitz continuous with Lipschitz constant $\frac{L}{1+\beta}$. From (23) and (24), we then have

$$\|\bar{x}_k - \bar{u}_k - \bar{x}^*\| \leq \frac{L}{1 + \beta}\|\bar{u}_k\|. \tag{60}$$

Hence,

$$\|\bar{x}_k - \bar{x}^*\|^2 - 2\langle\bar{x}_k - \bar{x}^*, \bar{u}_k\rangle + \|\bar{u}_k\|^2 \leq \frac{L^2}{(1 + \beta)^2}\|\bar{u}_k\|^2.$$

Therefore,

$$\left(\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}\right)^2 \leq \left(\frac{L^2}{(1 + \beta)^2} - 1\right) + 2\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}. \tag{61}$$

From (30), we have

$$-\frac{2\langle\bar{u}_k, \bar{v}_k\rangle}{(1 + \beta)\|\bar{u}_k\|^2} \leq -\left(\frac{\|\bar{v}_k\|}{\|\bar{u}_k\|}\right)^2 + \left(\frac{\|\bar{v}_k\|}{\|\bar{u}_k\|}\right)\left(\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}\right). \tag{62}$$

With (61), by Proposition A.1 applied to the expression on right-hand side of the above inequality, where $x = \frac{\|\bar{v}_k\|}{\|\bar{u}_k\|}$ and $y = \frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}$ in the propostion, we have from (62) that for all $k \geq 0$,

$$\gamma_k = -\frac{\langle\bar{u}_k, \bar{v}_k\rangle}{\|\bar{u}_k\|^2} \leq \frac{1 + \beta}{8}\left(1 + \frac{L}{1 + \beta}\right)^2.$$

Therefore,

$$\bar{\gamma}_{\max} \leq \frac{1 + \beta}{8}\left(1 + \frac{L}{1 + \beta}\right)^2.$$

Hence, since L satisfies (59), $\bar{\gamma}_{\max}$ satisfies (50) in Theorem 5.4, and the corollary then follows from the theorem. □

Corollary 5.6 Let $\{x_k\}$ be generated by the relaxed PR splitting method (2) given an initial iterate $x_0 \in \Re^n$, where $\theta = 2 + \beta + \epsilon$, $0 \leq \epsilon < \min\{\beta, 1/\beta\}$, $\beta > 0$. Suppose B is Lipschitz continuous with Lipschitz constant $L(> 0)$ such that

$$L < 1 + \beta.$$

Then there exists $i \leq k$ and $i \geq k_0$, where $k \geq 2k_0$, and M given by (51), such that

$$\|x_i - x_{i-1}\| \leq \left(\frac{2\sqrt{\theta}\|x_{k_0-1} - x^*\|}{\sqrt{M}}\right)\frac{1}{\sqrt{k}},$$

16

and for all $k \geq k_0$,

$$\|x_k - x^*\| \leq \left[\sqrt{1 - \frac{M\theta}{2(1+\beta)^2 \bar{L}^2}}\right] \|x_{k-1} - x^*\|,$$

where $\bar{L}$ is defined in Lemma 5.3. Here, $k_0 \geq 1$ depends only on problem data and parameters.

Proof: Since B is Lipschitz continuous with Lipschitz constant $L(> 0)$, we have $\bar{B}_0$ is Lipschitz continuous with Lipschitz constant $\frac{L}{1+\beta}$, and the following holds:

$$\left\|\frac{2}{1+\beta}\bar{u}_k - \bar{v}_k - (\bar{x}_k - \bar{x}^*)\right\| \leq \frac{L}{1+\beta}\|\bar{v}_k\|.$$

Squaring both sides of the above inequality, upon algebraic manipulations and applying Cauchy-Schwartz inequality, we obtain:

$$
\begin{aligned}
-\frac{4\langle \bar{u}_k, \bar{v}_k \rangle}{(1+\beta)\|\bar{u}_k\|^2} \quad \leq \quad & -\left(\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}\right)^2 + \frac{4}{1+\beta}\left(\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}\right) + 2\left(\frac{\|\bar{v}_k\|}{\|\bar{u}_k\|}\right)\left(\frac{\|\bar{x}_k - \bar{x}^*\|}{\|\bar{u}_k\|}\right) \\
& + \left[\frac{L^2}{(1+\beta)^2} - 1\right]\left(\frac{\|\bar{v}_k\|}{\|\bar{u}_k\|}\right)^2 - \frac{4}{(1+\beta)^2}.
\end{aligned}
$$

By Proposition A.2, when $L < 1 + \beta$, the expression on the right-hand side of the above inequality is less than or equal to zero. This implies that for all $k \geq 0$, $\gamma_k \leq 0$, which further implies that $\bar{\gamma}_{\max} \leq 0$. Condition (50) in Theorem 5.4 is hence satisfied, and the corollary follows. $\qquad\square$

Note that in the above corollaries, we do not attempt to optimize the bounds on the Lipschitz constants for A and B for the results to hold, and finding the optimized bounds for these is left to interested readers.

Remark 5.7 *Note that for $\theta = 2 + \beta, \beta > 0$, by Lemma 4.5 in [22], $\{\|x_k - x_{k-1}\|\}$ is nonincreasing. Hence, the above theorem and corollaries imply that when $\theta = 2 + \beta$, for $\bar{\gamma}_{\max} < 1$ or for A Lipschitz continuous with Lipschitz constant $L < 2$ or for B Lipschitz continuous with Lipschitz constant $L < 1 + \beta$, the pointwise convergence rate of $\|x_k - x_{k-1}\|$ for iterates $\{x_k\}$ generated using (2) on (1) is $\mathcal{O}(1/\sqrt{k})$. This is a partial extension of the pointwise convergence rate result in [22] from $\theta \in (0, 2 + \beta)$ to $\theta = 2 + \beta$. It is still an open problem whether a full extension is possible.*

Remark 5.8 *Observe from the proof of Theorem 5.4 that when $\theta \in (0, 2 + \beta)$ (that is, $\epsilon < 0$), we have pointwise convergence rate for $\|x_k - x_{k-1}\|$ of $\mathcal{O}(1/\sqrt{k})$ and R-linear convergence for $\|x_k - x^*\|$ without the need for (50) to hold. This is in line with what is known in the literature, as found for example in [1, 9, 13, 22].*

6 Numerical Study

In this section, we consider the weighted Lasso minimization problem that is studied in the literature, such as [6, 17, 22]. The problem is given by

$$\min_{u \in \Re^n} f(u) + g(u), \tag{63}$$

where $f(u) := \frac{1}{2}\|Cu - b\|^2$ and $g(u) := \|Wu\|_1$ for every $u \in \Re^n$. Here, $C \in \Re^{m \times n}$ and $b \in \Re^m$. C is a sparse data matrix with nonzero entries to zero entries in an average ratio of $1 : 20$ per row. Each component of b and each nonzero element of C is drawn from the Gaussian distribution with zero mean and unit variance, while $W \in \Re^{n \times n}$ is a diagonal matrix with positive diagonal elements drawn from the uniform distribution on the interval $[0, 1]$. This setup is inspired by [17, 22]. It can be seen easily that f defined above is a α-strongly convex function on $\Re^n$, where $\alpha = \lambda_{\min}(C^T C)$ is the minimum eigenvalue of $C^T C$. Hence, ∇f is maximal α-strongly monotone. Furthermore, f is differentiable and its gradient is κ-Lipschitz continuous on $\Re^n$, where $\kappa = \lambda_{\max}(C^T C)$, which is the maximum eigenvalue of $C^T C$. Clearly, g defined above is a convex function on $\Re^n$.

We consider solving (63) with f and g defined in the previous paragraph using the relaxed PR splitting method (2) on the monotone inclusion problem (1) with

$$A = \nabla f - \frac{\alpha}{2}I, \quad B = \partial g + \frac{\alpha}{2}I. \tag{64}$$

We have that A and B are maximal β-strongly monotone, where $\beta = \alpha/2$.

We first verify the convergence of iterates generated using (2) with A and B given by (64) for $\theta \in (2 + \beta, 2 + \beta + \min\{\beta, 1/\beta\})$. Convergence of iterates in this range of θ is predicted by our theory, in particular, Corollary 4.7. We consider a random instance of (C, W, b) with $m = 300$ and $n = 200$, and set the initial iterate x_0 to be $(1, \ldots, 1)^T$. We run the algorithm for θ varying from 1 to $2 + \beta + \min\{\beta, 1/\beta\}$, with the algorithm terminating at the kth iteration when $\|x_k - x_{k-1}\| \leq 10^{-5}$. Our results are given in Figure 1. For this instance of (C, W, b), β is given by 0.1896. We see from Figure 1 that the number of iterations performed before termination decreases as θ increases from 1, reaches its minimum around $\theta = 2.25$, which is close to $\theta = 2 + \beta$, and increases again. Our results indicate that convergence using (2) occurs as predicted by Corollary 4.7, when θ lies between $2 + \beta$ and $2 + \beta + \min\{\beta, 1/\beta\}$. Indeed, when we set $\theta = 2 + \beta + \min\{\beta, 1/\beta\} + 0.05$, the algorithm still does not terminate when the number of iterations has reached 1000.

Next, we test the validity of Assumption 4.1 on the weighted Lasso minimization problem (63), by applying (2) on (1) with A, B given by (64). In our numerical experiments, we set $n = m$ to be 200 and 400, and x_0 to be $(1, \ldots, 1)^T$ and $(0, -1, \ldots, -1)^T$. For each of the four scenarios, we generate (C, W) randomly 100 times, and run the algorithm for 800 iterations on those instances of (C, W) with $\alpha = \lambda_{\min}(C^T C) > 0$, that is, $\beta > 0$. We called these instances "acceptable". We set θ to be always equal to $2 + \beta + \frac{1}{2}\min\{\beta, 1/\beta\}$.

In order to test the validity of Assumption 4.1, we need to know the optimal solution to (63). We do this by setting b to be zero. Hence, the optimal solution to (63) is $u^* = 0$. Therefore, $x^* = 0$, since $u^* = J_A(x^*)$, where A is given in (64). Hence, to validate Assumption 4.1, we need to verify that for each iterate, x_k, we have $x_k^i \neq 0$, for $i = 1, \ldots, n$.

Our results are shown in Table 1. These results give preliminary indication that Assumption 4.1 should hold in practice, and is only a technical assumption needed to prove convergence using the relaxed PR splitting method (2) to solve the monotone inclusion problem (1) for θ beyond the range of $(0, 2 + \beta]$, and less than $2 + \beta + \min\{\beta, 1/\beta\}$.
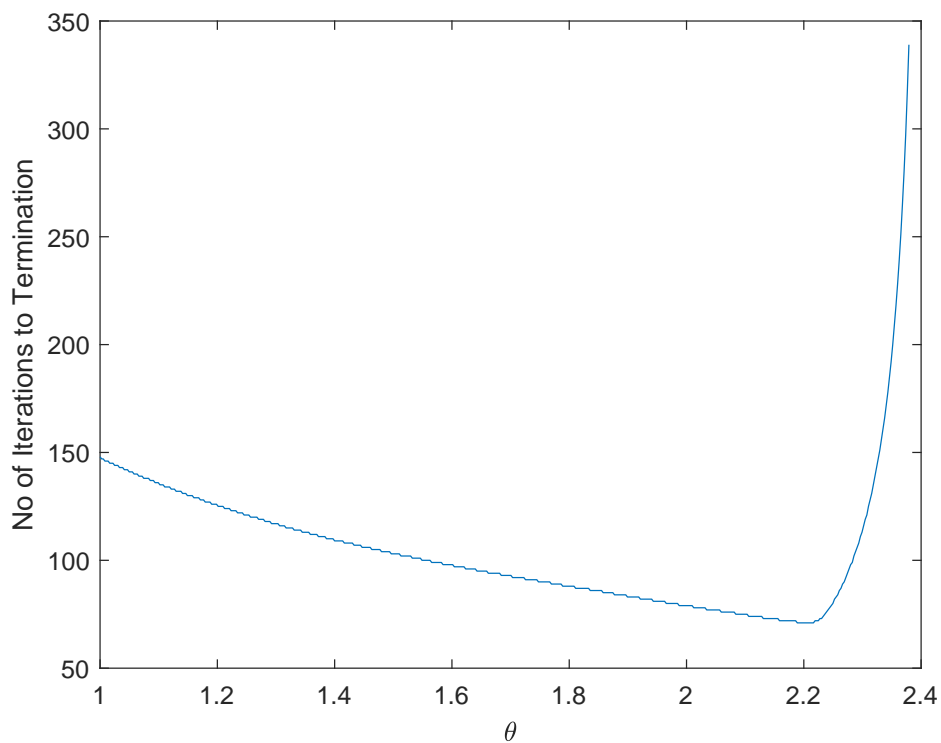
Figure 1: Graph showing how the number of iterations before termination, upon applying (2), varies with θ.

	# "acceptable" instances	# instances with Assumption 4.1 satisfied
$n = m = 200$ $x_0 = (1, \ldots, 1)^T$	98	98
$n = m = 400$ $x_0 = (1, \ldots, 1)^T$	100	100
$n = m = 200$ $x_0 = (0, -1, \ldots, -1)^T$	100	100
$n = m = 400$ $x_0 = (0, -1, \ldots, -1)^T$	100	100

Table 1: For each scenario, 100 instances with $b = 0$ are generated.

7 Conclusion

In this paper, we consider the relaxed PR splitting method (2) to solve the monotone inclusion problem (1). We consider $A, B : \Re^n \rightrightarrows \Re^n$ in (1) to be maximal β-strongly monotone operators, with $\beta > 0$, in this paper. We show for the first time that for $\theta \in (2 + \beta, 2 + \beta + \min\{\beta, 1/\beta\})$, an accumulation point, x^{**}, of iterates $\{x_k\}$ generated using (2) on (1) has $J_A(x^{**})$ that solves (1). As a consequence, if A or B is single-valued, then $\{x_k\}$ converges to a limit point x^*, where $J_A(x^*)$ solves (1). These are shown under a technical assumption. Note that for $n = 1$, not having this technical assumption leads to trivial consideration. Furthermore, for $\theta \in [2 + \beta, 2 + \beta + \min\{\beta, 1/\beta\})$, we provide pointwise convergence rate and R-linear convergence rate results of $\{x_k\}$ in Theorem 5.4 and Corollaries 5.5, 5.6. Through numerical experiments on the weighted Lasso minimization problem, we provide preliminary evidence to show that the technical assumption used in this paper is likely to hold in practice.

Acknowledgement

References

[1] S. Bartz, M.N. Dao & H.M. Phan, *Conical averagedness and convergence analysis of fixed point algorithms*, Journal of Global Optimization, **82**(2022), pp. 351-373.

[2] H.H. Bauschke & P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[3] H.H. Bauschke, J.Y. Bello Cruz, T.T.A. Nghia, H.M. Phan & X. Wang, *The rate of linear convergence of the Douglas-Rachford algorithm for subspaces is the cosine of the Friedrichs angle*, Journal of Approximation Theory, **185**(2014), pp. 63-79.

[4] H.H. Bauschke, J.Y. Bello Cruz, T.T.A. Nghia, H.M. Phan & X. Wang, *Optimal rates of linear convergence of relaxed alternating projections and generalized Douglas-Rachford methods for two subspaces*, Numerical Algorithms, **73**(2016), pp. 33-76.

[5] H.H. Bauschke, D. Noll & H.M. Phan, *Linear and strong convergence of algorithms involving averaged nonexpansive operators*, Journal of Mathematical Analysis and Applications, **421**(2015), pp. 1-20.

[6] E.J. Candes, M.B. Wakin & S. Boyd, *Enhancing sparsity by reweighted l_1 minimization*, Journal of Fourier Analysis and Applications, **14**(2008), pp. 877 - 905.

[7] P.L. Combettes, *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization, **53**(2004), pp. 475-504.

[8] P.L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, Journal of Convex Analysis, **16**(2009), pp. 727-748.

[9] M.N. Dao & H.M. Phan, *Adaptive Douglas-Rachford splitting algorithm for the sum of two operators*, SIAM Journal on Optimization, **29**(2019), pp. 2697-2724.

[10] D. Davis, *Convergence rate analysis of the Forward-Douglas-Rachford splitting scheme*, SIAM Journal on Optimization, **25**(2015), pp. 1760-1786.

[11] D. Davis & W. Yin, *Convergence rate analysis of several splitting schemes*, in "Splitting Methods in Communication, Imaging, Science, and Engineering" (R. Glowinski, S. Osher, W. Yin, eds), pp. 115-163, Scientific Computation book series, Springer, Cham., 2016.

[12] D. Davis & W. Yin, *Faster convergence rates of relaxed Peaceman-Rachford and ADMM under regularity assumptions*, Mathematics of Operations Research, **42**(2017), pp. 783-805.

[13] Y. Dong & A. Fischer, *A family of operator splitting methods revisited*, Nonlinear Analysis, **72**(2010), pp. 4307-4315.

[14] J. Eckstein & D.P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, **55**(1992), pp. 293-318.

[15] F. Facchinei & J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problem*, Volume II, Springer-Verlag, 2003.

[16] P. Giselsson, *Tight global linear convergence rate bounds for Douglas-Rachford splitting*, Journal of Fixed Point Theory and Applications, **19**(2017), pp. 22412270.

[17] P. Giselsson & S. Boyd, *Linear convergence and metric selection for Douglas-Rachford splitting and ADMM*, IEEE Transactions on Automatic Control, **62**(2017), pp. 532 - 544.

[18] B. He & X. Yuan, *On the convergence rate of Douglas-Rachford operator splitting method*, Mathematical Programming, Series A, **153**(2015), pp. 715-722.

[19] F. John, *Extremum problems with inequalities as side conditions*, in "Studies and Essays, Courant Anniversary Volume" (K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker, eds), pp. 187 - 204, Wiley (Interscience), New York, 1948.

[20] P.-L. Lions & B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, **16**(1979), pp. 964-979.

[21] O.L. Mangasarian & S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, Journal of Mathematical Analysis and Applications, **17**(1967), pp. 37 - 47.

[22] R.D.C. Monteiro & C.-K. Sim, *Complexity of the relaxed Peaceman-Rachford splitting method for the sum of two maximal strongly monotone operators,* Computational Optimization and Applications, **70**(2018), pp. 763-790.

[23] C.-K. Sim, *A note on convergence using the relaxed Peaceman-Rachford splitting method on the sum of two maximal strongly monotone operators,* unpublished manuscript, August 2020 (https://drive.google.com/file/d/1Yf5rn21mhof4j3u8qQRlFVzONfbtn6jV/view).

A   Appendix

Proposition A.1 *We have*

$$\max_{x \ge 0, 0 \le y^2 \le L^2/(1+\beta)^2 - 1 + 2y} -x^2 + xy$$

is less than or equal to $\frac{1}{4} \left(1 + \frac{L}{1+\beta} \right)^2$.

Proof: We have an optimal solution (x^*, y^*) to the above maximization problem satisfies:

$$\begin{pmatrix} 2x^* - y^* \\ -x^* \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 2y^* - 2 \end{pmatrix} = 0, \tag{65}$$

with

$$\lambda_1 x^* = 0, \quad \lambda_1 \ge 0, \quad x^* \ge 0, \tag{66}$$

$$\lambda_2 y^* = 0, \quad \lambda_2 \ge 0, \quad y^* \ge 0, \tag{67}$$

$$\lambda_3 \left((y^*)^2 - 2y^* + 1 - \frac{L^2}{(1+\beta)^2} \right) = 0, \quad \lambda_3 \ge 0, \quad (y^*)^2 \le \frac{L^2}{(1+\beta)^2} - 1 + 2y^*. \tag{68}$$

We first observe if $x^* = 0$, then $y^* = 0$, and the proposition is proved. Suppose $x^* > 0$. Then, from (66), we have $\lambda_1 = 0$ and this implies that $y^* = 2x^*$ from the first "row" of (65). Also, observe from the second "row" in (65) with $x^* > 0$ and $\lambda_2 \ge 0$ that we must have $\lambda_3 > 0$. Hence from (68), we get

$$(y^*)^2 - 2y^* + 1 - \frac{L^2}{(1+\beta)^2} = 0.$$

Therefore,

$$y^* = \frac{2 \pm \sqrt{4 - 4 \left(1 - \frac{L^2}{(1+\beta)^2} \right)}}{2} = 1 \pm \frac{L}{1+\beta}.$$

Now, since $y^* = 2x^* > 0$, from (67), we get $\lambda_2 = 0$. Therefore, from the "second" row in (65), we have $x^* = \lambda_3 (2y^* - 2) > 0$. We observe then that $y^* \ne 1 - \frac{L}{1+\beta}$. Hence, $y^* = 1 + \frac{L}{1+\beta}$. And $x^* = \frac{1}{2} y^* = \frac{1}{2} \left(1 + \frac{L}{1+\beta} \right)$. The proposition is hence proved. $\qquad \square$

Proposition A.2 *For $L < 1 + \beta$, we have*

$$\max_{x,y \ge 0} -y^2 + \frac{4}{1+\beta} y + 2xy + \left[\frac{L^2}{(1+\beta)^2} - 1 \right] x^2 - \frac{4}{(1+\beta)^2}$$

is less than or equal to zero.

Proof: We have an optimal solution (x^*, y^*) to the above maximization problem satisfies:

$$\begin{pmatrix} -2y^* - 2 \left(\frac{L^2}{(1+\beta)^2} - 1 \right) x^* \\ 2y^* - 2x^* - \frac{4}{1+\beta} \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0, \tag{69}$$

22

with

$$\lambda_1 x^* = 0, \quad \lambda_1 \geq 0, \quad x^* \geq 0, \tag{70}$$

$$\lambda_2 y^* = 0, \quad \lambda_2 \geq 0, \quad y^* \geq 0. \tag{71}$$

Suppose $x^* > 0$. Then, from (70), we have $\lambda_1 = 0$, and hence from the first "row" in (69), we get

$$y^* = \left(1 - \frac{L^2}{(1+\beta)^2}\right) x^*. \tag{72}$$

Given that $L < 1 + \beta$, from (72), we have $y^* > 0$, which then implies by (71) that $\lambda_2 = 0$. Hence, from the second "row" in (69), we obtain

$$y^* = x^* + \frac{2}{1+\beta}. \tag{73}$$

By (72) and (73), we have

$$x^* = -\frac{2(1+\beta)}{L^2},$$

which is less than zero. This is a contradiction to $x^* > 0$. Hence, $x^* = 0$. Suppose $y^* = 0$. Then the optimal value to the maximization problem is $-4/(1+\beta)^2$. Suppose $y^* > 0$. Then $\lambda_2 = 0$, and we have from the second "row" of (69), $y^* = 2/(1+\beta)$. Together with $x^* = 0$, the objective function value of the problem is then equal to zero, and we are done. $\qquad \square$