SUPPLEMENTARY MATERIALS: ANDERSON ACCELERATED DOUGLAS–RACHFORD SPLITTING*

ANQI FU[†], JUNZI ZHANG[‡], AND STEPHEN BOYD[†]

In this supplementary material, we provide the proofs for the theorems in the main text.

SM1. Preliminaries. We begin with the following lemma, which establishes the connection between residuals of the DRS fixed-point mapping and the primal/dual residuals of the original problem (1.2).

LEMMA SM1.1. *Suppose that* $\liminf_{j\to\infty} \|v^j - F_{\mathrm{DRS}}(v^j)\|_2 \le \epsilon$ *for some* $\epsilon \ge 0$. *Then*

$$\liminf_{j\to\infty} \|r^j_{\mathrm{prim}}\|_2 \le \|A\|_2\epsilon, \quad \liminf_{j\to\infty} \|r^j_{\mathrm{dual}}\|_2 \le \frac{1}{t}\epsilon. \tag{SM1.1}$$

Proof. By expanding F_{DRS}, and in particular line 6 of Algorithm 2.1, we see that

$$\liminf_{j\to\infty} \|x^{j+1/2} - x^{j+1}\|_2 = \liminf_{j\to\infty} \|v^j - v^{j+1}_{\mathrm{DRS}}\|_2 \le \epsilon.$$

Since $Ax^{j+1} = b$ by the projection step in F_{DRS}, we have

$$r^j_{\mathrm{prim}} = Ax^{j+1/2} - b = A(x^{j+1/2} - x^{j+1}),$$

which implies that

$$\liminf_{j\to\infty} \|r^j_{\mathrm{prim}}\|_2 \le \|A\|_2 \liminf_{j\to\infty} \|x^{j+1/2} - x^{j+1}\|_2 \le \|A\|_2\epsilon,$$

and hence $\liminf_{j\to\infty} \|r^j_{\mathrm{prim}}\|_2 \le \|A\|_2\epsilon$.

On the other hand, the optimality conditions from lines 3 and 5 of Algorithm 2.1 give us

$$\frac{1}{t}(x^{j+1/2} - v^j) + g^j = 0, \quad x^{j+1} = v^{j+1/2} - A^T\tilde{\lambda}^j,$$

for some $g^j \in \partial f(x^{j+1/2})$ and $\tilde{\lambda}^j = (AA^T)^\dagger(Av^{j+1/2} - b)$. Thus,

$$\begin{aligned} g^j &= \frac{1}{t}(v^j - x^{j+1/2}) \\ &= \frac{1}{t}(v^{j+1/2} - x^{j+1}) + \frac{1}{t}(v^j - v^{j+1/2}) + \frac{1}{t}(x^{j+1} - x^{j+1/2}) \\ &= \frac{1}{t}A^T\tilde{\lambda}^j + 2\frac{1}{t}(v^j - x^{j+1/2}) + \frac{1}{t}(x^{j+1} - x^{j+1/2}) \\ &= \frac{1}{t}A^T\tilde{\lambda}^j + 2g^j + \frac{1}{t}(x^{j+1} - x^{j+1/2}), \end{aligned} \tag{SM1.2}$$

where we have used line 4 of Algorithm 2.1 in the third equality. Rearranging terms yields $g^j = A^T(-\frac{1}{t}\tilde{\lambda}^j) + \frac{1}{t}(x^{j+1/2} - x^{j+1})$.

*Supplementary material for SISC MS#M129009.
https://doi.org/10.1137/19M1290097

[†]Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (anqif@stanford.edu, boyd@stanford.edu).
[‡]ICME, Stanford University, Palo Alto, CA 94304 USA (junziz@stanford.edu).

Finally, since we compute $r_{\text{dual}}^j = g^j + A^T\lambda^j$ using $\lambda^j \in \text{argmin}_\lambda \|g^j + A^T\lambda\|_2$ (c.f. residuals and dual variables in §2),

$$\liminf_{j\to\infty} \|r_{\text{dual}}^j\|_2 \le \liminf_{j\to\infty} \|g^j + A^T\bar{\lambda}^j\|_2 = \frac{1}{t}\liminf_{j\to\infty} \|x^{j+1/2} - x^{j+1}\|_2 \le \frac{1}{t}\epsilon,$$

where $\bar{\lambda}^j = \frac{1}{t}\tilde{\lambda}^j$. This completes our proof. ∎

Remark SM1.2. When $\epsilon = 0$, Lemma SM1.1 implies that

$$\liminf_{j\to\infty} \|r_{\text{prim}}^j\|_2 = \liminf_{j\to\infty} \|r_{\text{dual}}^j\|_2 = 0.$$

Furthermore, notice that we could have calculated r_{dual}^j using

$$\lambda^j = \bar{\lambda}^j = \frac{1}{t}(AA^T)^\dagger(Av^{k+1/2} - b),$$

and the results would still hold.

SM2. Proof of Theorems 4.3 and 4.5. We now prove the convergence results in the error-free setting. Define the infimal displacement vector of F_{DRS} as $\delta v^\star = \Pi_{\overline{\mathbf{range}(I - F_{\text{DRS}})}}(0)$. It follows directly that $\|\delta v^\star\|_2 = \inf_{v\in\mathbf{R}^n} \|v - F_{\text{DRS}}(v)\|_2$. We will later show that in A2DR, $\lim_{k\to\infty} v^k - v^{k+1} = \delta v^\star$. In particular, Theorem 4.5 gives us $\delta v = \delta v^\star$.

We begin by showing that $\delta v^\star = 0$ if and only if problem (1.2) is solvable. To see this, first notice that by [SM1, Corollary 6.5],

$$\delta v^\star = \text{argmin}_{z\in\mathcal{Z}} \|z\|_2,$$

where

$$\mathcal{Z} = \overline{\mathbf{dom}\, f - \mathbf{dom}\, g} \cap t(\overline{\mathbf{dom}\, f^* + \mathbf{dom}\, g^*}), \quad g(x) = \mathcal{I}_{\{v\,:\,Av=b\}}(x).$$

Since $\mathbf{dom}\, g = \{x : Ax = b\}$ and $\mathbf{dom}\, g^* = \mathbf{range}(A^T) = -\mathbf{range}(A^T)$, the problem is solvable if and only if

$$\mathbf{dist}(\mathbf{dom}\, f, \mathbf{dom}\, g) = \mathbf{dist}(\mathbf{dom}\, f^*, -\mathbf{dom}\, g^*) = 0,$$

which holds if and only if $0 \in \overline{\mathbf{dom}\, f - \mathbf{dom}\, g}$ and $0 \in \overline{\mathbf{dom}\, f^* + \mathbf{dom}\, g^*}$, *i.e.*, $\delta v^\star = 0$.

Below we denote the initial iteration counts for accepting AA candidates as k_i (*i.e.*, when $I_{\text{safeguard}}$ is `True` or $R_{\text{AA}} \ge R$, and the check in Algorithm 3.1, line 14 passes), and the iteration counts for accepting DRS candidates as l_i. Notice that for each iteration k, either $k = k_i + K$ for some i and $0 \le K \le R - 1$, or $k = l_i$ for some i.

- **Case (i) [Theorem 4.3,** (4.2)]
  First, suppose that problem (1.2) is solvable. Then, $\delta v^\star = 0$. By Lemma SM1.1, to prove (4.2), it suffices to prove that $\liminf_{k\to\infty} \|g^k\|_2 = 0$. If the set of k_i is infinite, *i.e.*, the AA candidate is adopted an infinite number of times, then

  $$0 \le \liminf_{k\to\infty} \|g^k\|_2 \le \liminf_{i\to\infty} \|g^{k_i}\|_2 \le D\|g^0\|_2 \lim_{i\to\infty} (i+1)^{-(1+\epsilon)} = 0.$$

  Here we used the fact that $n_{AA}/R = i$ in iteration k_i.

On the other hand, if the set of k_i is finite, Algorithm 3.1 reduces to the vanilla DRS algorithm after a finite number of iterations. By [SM3, Theorem 2], this means that $\lim_{k\to\infty} g^k = \lim_{k\to\infty} v^k - v^{k+1} = \delta v^\star = 0$. Thus, we always have $\liminf_{k\to\infty} \|g^k\|_2 = 0$, and this fact coupled with Lemma SM1.1 immediately gives us (4.2).
Notice that the case of finite k_i's cannot actually happen. Otherwise, since $\lim_{k\to\infty} \|g^k\|_2 = 0$ and n_{AA} is upper bounded (because AA candidates are rejected after some point), the check on line 14 of Algorithm 3.1 must pass eventually. This means that an AA candidate is accepted one more time, which is a contradiction. Hence it must be that AA candidates are adopted an infinite number of times.

- **Case (ii) [Theorem 4.3, iteration convergence]**
Now suppose that $F_{\rm DRS}$ has a fixed point. As $G_{\rm DRS}$ is non-expansive, if the AA candidate is adopted in iteration k,

$$\begin{aligned}\|g^{k+1}\|_2 = \|G_{\rm DRS}(v^{k+1})\|_2 &\le \|G_{\rm DRS}(v^{k+1}) - G_{\rm DRS}(v^k)\|_2 + \|G_{\rm DRS}(v^k)\|_2\\ &\le (\|H_k\|_2 + 1)\|g^k\|_2 \le 2(1+1/\eta)\|g^k\|_2,\end{aligned}$$

where we have used Lemma 4.2 to bound $\|H_k\|_2$. This immediately implies that for any $0 \le K \le R-1$,

$$\|g^{k_i+K}\|_2 \le (2+2/\eta)^K \|g^{k_i}\|_2 \le D\|g^0\|_2 (2+2/\eta)^K (i+1)^{-(1+\epsilon)}, \tag{SM2.1}$$

and so we have $\lim_{i\to\infty} \|g^{k_i+K}\|_2 = 0$.
In addition, since AA candidates are accepted in all iterations $k_i + K$, again by Lemma 4.2, we have that for any $w \in \mathbf{R}^n$,

$$\begin{aligned}\|v^{k_i+K+1} - w\|_2 &\le \|v^{k_i+K} - w\|_2 + (1+2/\eta)\|g^{k_i+K}\|_2\\ &\le \cdots \le \|v^{k_i} - w\|_2 + (1+2/\eta)\sum_{j=0}^{K} \|g^{k_i+j}\|_2\\ &\le \|v^{k_i} - w\|_2 + (1+2/\eta)\|g^{k_i}\|_2 \sum_{j=0}^{K} (2+2/\eta)^j\\ &\le \|v^{k_i} - w\|_2 + (1+2/\eta) C_R D \|g^0\|_2 (i+1)^{-(1+\epsilon)},\end{aligned} \tag{SM2.2}$$

where $C_R = \sum_{j=0}^{R-1} (2+2/\eta)^j$ is a constant.
Now let $v^\star$ be a fixed point of $F_{\rm DRS}$. Since $F_{\rm DRS}$ is 1/2-averaged, by inequality (5) in [SM4],

$$\|v^{l_i+1} - v^\star\|_2^2 \le \|v^{l_i} - v^\star\|_2^2 - \|g^{l_i}\|_2^2 \le \|v^{l_i} - v^\star\|_2^2 \tag{SM2.3}$$

for any $i \ge 0$. Hence for any $k \ge 0$,

$$\|v^k - v^\star\|_2 \le \|v^0 - v^\star\|_2 + (1+2/\eta) C_R D \|g^0\|_2 \sum_{i=0}^{\infty} (i+1)^{-(1+\epsilon)} = E < \infty,$$

implying that $\|v^k - v^\star\|_2$ is bounded.
As a result, by squaring both sides of (SM2.2) and combining with (SM2.3), we get that

$$\sum_{i=0}^{\infty} \|g^{l_i}\|_2^2 \le \|v^0 - v^\star\|_2^2 + \text{const},$$

where

$$\text{const} = \left((1+2/\eta)C_R D\|g^0\|_2\right)^2 \sum_{i=0}^{\infty}(i+1)^{-(2+2\epsilon)} + (2+4/\eta)C_R DE\|g^0\|_2 \sum_{i=0}^{\infty}(i+1)^{-(1+\epsilon)} < \infty.$$

Thus, $\lim_{i\to\infty}\|g^{l_i}\|_2 = 0$. Together with the fact that $\lim_{i\to\infty}\|g^{k_i+K}\|_2 = 0$ for $0 \le K \le R-1$, we immediately obtain $\lim_{k\to\infty}\|g^k\|_2 = 0$, and an application of Lemma SM1.1 yields (4.2).

Notice that in our derivation, we implicitly assumed both index sets are infinite. The set of k_i is always infinite by the same logic as in case (i). Moreover, if the set of l_i is finite, the arguments above involving l_i can be ignored, as eventually $k = k_i + K$ for all i above some threshold.

It still remains to be shown that v^k converges to a fixed-point of F_{DRS}. To do this, we first show that $\|v^k - v^\star\|_2$ is quasi-Fejérian. Squaring both sides of the first inequality in (SM2.2) and combining it with (SM2.1) and (SM2.3), we get that for any $k \ge 0$,

$$\|v^{k+1} - v^\star\|_2^2 \le \|v^k - v^\star\|_2^2 + \epsilon^k, \tag{SM2.4}$$

where $\epsilon^{l_i} = 0$ and

$$\epsilon^{k_i+K} = 2DE\|g^0\|_2(1+2/\eta)(2+2/\eta)^K(i+1)^{-(1+\epsilon)} + \left(D\|g^0\|_2(1+2/\eta)\right)^2(2+2/\eta)^{2K}(i+1)^{-(2+2\epsilon)}$$

for $0 \le K \le R-1$. Hence $\epsilon^k \ge 0$ and $\sum_{k=0}^{\infty}\epsilon^k < \infty$. In other words, $\|v^k - v^\star\|_2$ is quasi-Fejérian.

Since $\lim_{k\to\infty}\|g^k\|_2 = 0$ and inequality (SM2.4) holds, we can invoke [SM2, Theorem 3.8] to conclude that $\lim_{k\to\infty}\|v^k - v^\star\|_2$ exists and v^k converges to some fixed-point of F_{DRS} (not necessarily $v^\star$). The convergence of $x^{k+1/2}$ to a solution of (1.2) follows directly from the continuity of the proximal operators.

- **Case (iii) [Theorem 4.5]**

  Now suppose that problem (1.2) is pathological, then $\delta v^\star \ne 0$. Since

  $$\|\delta v^\star\|_2 = \inf_{v\in\mathbf{R}^n}\|v - F_{\text{DRS}}(v)\|_2,$$

  the safeguard will always be invoked for sufficiently large iteration k because $\|g^k\|_2 \ge \|\delta v^\star\|_2 > 0$. Hence the algorithm reduces to vanilla DRS in the end. We can thus prove the result in case (iii) by appealing to previous work on vanilla DRS [SM3, SM1, SM5].

  Recall that $\lim_{k\to\infty} v^k - v^{k+1} = \delta v^\star \ne 0$ [SM3, Theorem 2]. First, we will show that problem (1.2) is dual strongly infeasible if and only if

  $$\lim_{k\to\infty} Ax^{k+1/2} = b.$$

  If the problem is dual strongly infeasible, then by [SM5, Lemma 1], it is primal feasible and has an improving direction $d = -\frac{1}{t}\delta v^\star$ [SM5, Corollary

3]. Along this direction, both f and $g = \mathcal{I}_{\{x\,:\,Ax=b\}}$ remain feasible, and in particular, $A\delta v^\star = 0$. Hence

$$\lim_{k\to\infty} Ax^{k+1/2} - Ax^{k+1} = \lim_{k\to\infty} A(v^k - v^{k+1}) = A\delta v^\star = 0,$$

which implies that $\lim_{k\to\infty} Ax^{k+1/2} = b$ since $Ax^{k+1} = b$ for all $k \geq 0$. Conversely, if $\lim_{k\to\infty} Ax^{k+1/2} = b$, then $\mathbf{dist}(\mathbf{dom}\, f, \mathbf{dom}\, g) = 0$ because $x^{k+1/2} \in \mathbf{dom}\, f$. This implies problem (1.2) is not primal strongly infeasible, so it must be dual strongly infeasible since we assumed the problem is pathological.
Hence if $\lim_{k\to\infty} Ax^{k+1/2} = b$, problem (1.2) is dual strongly infeasible, and by [SM5, Lemma 1 and Corollary 3], it is unbounded and

$$\delta v^\star = t\Pi_{\overline{\mathbf{dom}\, f^* + \mathbf{dom}\, g^*}}(0),$$

which implies that

$$\|\delta v^\star\|_2 = t\,\mathbf{dist}(\mathbf{dom}\, f^*, \mathbf{range}(A^T)).$$

Otherwise, the problem is not dual strongly infeasible and thus must be primal strongly infeasible by our assumption of pathology, so from [SM1, Corollary 6.5],

$$\|\delta v\|_2 \geq \mathbf{dist}(\mathbf{dom}\, f, \{x \,:\, Ax = b\}).$$

When the dual problem is feasible, $\delta v^\star = \Pi_{\overline{\mathbf{dom}\, f - \mathbf{dom}\, g}}(0)$ [SM5, Corollary 5], which implies that

$$\|\delta v^\star\|_2 = \mathbf{dist}(\mathbf{dom}\, f, \{x \,:\, Ax = b\}).$$

SM3. Proof of Theorem 4.4. The proof resembles that of Theorem 4.3 (with identical notation), so here we mainly highlight the differences caused by the computational errors η_1^k, η_2^k. We begin by bounding the difference between the error-corrupted fixed-point mapping, denoted by $\hat{F}_{\text{DRS}}$, and the error-free mapping F_{DRS}. Starting from any $v^k \in \mathbf{R}^n$, we have by definition

$$\|\hat{v}^{k+1/2} - v^{k+1/2}\|_2 = 2\|\hat{x}^{k+1/2} - x^{k+1/2}\|_2 = 2\|\eta_1^k\|_2,$$

$$\|\hat{x}^{k+1} - x^{k+1}\|_2 \leq \|\hat{v}^{k+1/2} - v^{k+1/2}\|_2 + \|\eta_2^k\|_2 = 2\|\eta_1^k\|_2 + \|\eta_2^k\|_2,$$

where the inequality comes from the non-expansiveness of Π. Let $\hat{G}_{\text{DRS}}(v) = v - \hat{F}_{\text{DRS}}(v)$. Since $\|\eta_1^k\|_2 \leq \epsilon'$ and $\|\eta_2^k\|_2 \leq \epsilon'$,

$$\begin{aligned}
\|g^k - G_{\text{DRS}}(v^k)\|_2 &= \|\hat{G}_{\text{DRS}}(v^k) - G_{\text{DRS}}(v^k)\|_2 \\
&= \|\hat{F}_{\text{DRS}}(v^k) - F_{\text{DRS}}(v^k)\|_2 \\
&\leq \|\hat{x}^{k+1} - x^{k+1}\|_2 + \|\hat{x}^{k+1/2} - x^{k+1/2}\|_2 \\
&\leq 3\|\eta_1^k\|_2 + \|\eta_2^k\|_2 \leq 4\epsilon'.
\end{aligned}$$

Thus, by Lemma SM1.1, it suffices to prove that $\liminf_{k\to\infty} \|G_{\text{DRS}}(v^k)\|_2 \leq 4\epsilon' + 4\sqrt{L\epsilon'}$.

On the one hand, if the set of k_i (AA candidates) is infinite,

$$\begin{aligned}
\liminf_{k\to\infty} \|G_{\text{DRS}}(v^k)\|_2 &\leq \liminf_{i\to\infty} \|G_{\text{DRS}}(v^{k_i})\|_2 \leq \liminf_{i\to\infty} \|g^{k_i}\|_2 + 4\epsilon' \\
&\leq D\|g^0\|_2 \lim_{i\to\infty} (i+1)^{-(1+\epsilon)} + 4\epsilon' = 4\epsilon'.
\end{aligned}$$

Otherwise, the set of k_i is finite, and the algorithm reduces to vanilla DRS after a finite number of iterations. Without loss of generality, suppose we start running the error-corrupted vanilla DRS algorithm from the first iteration.

Let $v^\star$ be a fixed-point of F_{DRS}. By inequality (5) in [SM4],

$$\begin{aligned} \|v^{k+1} - v^\star\|_2^2 &\leq \left(\|\hat{F}_{\text{DRS}}(v^k) - F_{\text{DRS}}(v^k)\|_2 + \|F_{\text{DRS}}(v^k) - v^\star\|_2\right)^2 \\ &\leq 16(\epsilon')^2 + 8\epsilon'\|v^k - v^\star\|_2 + \|F_{\text{DRS}}(v^k) - v^\star\|_2^2 \\ &\leq 16(\epsilon')^2 + 16L\epsilon' + \|v^k - v^\star\|_2^2 - \|G_{\text{DRS}}(v^k)\|_2^2 \end{aligned} \tag{SM3.1}$$

for all $k \geq 0$, where in the second step, we use the fact that $\|\hat{F}_{\text{DRS}}(v^k) - F_{\text{DRS}}(v^k)\|_2 \leq 4\epsilon'$ and F_{DRS} is non-expansive, and in the third step, we employ $\|v^k\|_2 \leq L$ and $\|v^\star\|_2 \leq L$ along with the triangle inequality. Rearranging terms and telescoping the inequalities,

$$\frac{1}{K}\sum_{k=0}^{K-1} \|G_{\text{DRS}}(v^k)\|_2^2 \leq \frac{1}{K}\|v^0 - v^\star\|_2^2 + 16(\epsilon')^2 + 16L\epsilon',$$

which immediately implies that

$$\liminf_{k\to\infty} \|G_{\text{DRS}}(v^k)\|_2 \leq \sqrt{16(\epsilon')^2 + 16L\epsilon'} \leq 4\epsilon' + 4\sqrt{L\epsilon'}.$$

Together with Lemma SM1.1, this completes the proof.

REFERENCES

[SM1] H. H. Bauschke, W. L. Hare, and W. M. Moursi, *On the range of the Douglas-Rachford operator*, Math. Oper. Res., 41(3) (2016), pp. 884–897.

[SM2] P. L. Combettes, *Quasi-Fejérian analysis of some optimization algorithms*, Stud. Comput. Math., 8 (2001), pp. 115–152.

[SM3] A. Pazy, *Asymptotic behavior of contractions in Hilbert space*, Isr. J. Math., 9(2) (1971), pp. 235–240.

[SM4] E. K. Ryu and S. Boyd, *A primer on monotone operator methods*, Appl. Comput. Math, 15(1) (2016), pp. 3–43.

[SM5] E. K. Ryu, Y. Liu, and W. Yin, *Douglas-Rachford splitting and ADMM for pathological convex optimization*, Comput. Optim. Appl., 74 (2019), pp. 747–778.