Technical Report
Analysis of the Convergence Rate of SoPro

Xuyang Wu, Zhihai Qu and Jie Lu

In this report, we discuss the impact of the optimization problem and the network topology on the convergence rate of SoPro [1]. Please see [1] for all the assumptions and notations.

Theorem 1 in [1] says that the linear convergence rate of SoPro is given by

$$\|\mathbf{z}^{k+1}-\mathbf{z}^\star\|_Q^2 \leq (1-\delta)\|\mathbf{z}^k-\mathbf{z}^\star\|_Q^2,$$

where

$$\delta = \sup_{c_1,c_2>0} \min\left\{\frac{\rho\lambda_W\kappa_{\beta,\eta}}{(1+c_1)\|\Lambda_M+D\|^2}, \frac{1-\eta}{(1+1/c_1)(1+c_2)}, \frac{2\eta m_\rho-\beta}{\lambda_{\max}(R+(1+1/c_1)(1+1/c_2)\Lambda_M^2/(\rho\lambda_W))}\right\}. \quad (1)$$

Here, λ_W is the smallest nonzero eigenvalue of the weight matrix W and $\beta \in (0, 2\eta m_\rho)$ is such that $\kappa_{\beta,\eta} := \lambda_{\min}(R - \frac{\Lambda_M}{2(1-\eta)} - \frac{(\Lambda_M-\Lambda_m)^2}{4\beta} - \Lambda_M + \Lambda_m - \rho W) > 0$.

Below, we analyze the convergence rate under the following setting:

1) Each f_i, $i \in \mathcal{V}$ is globally strongly convex with convexity parameter $m_i > 0$.
2) Without loss of generality, we set $m_i = m\ \forall i \in \mathcal{V}$ and $M_i = M\ \forall i \in \mathcal{V}$ for some $0 < m < M$. Also, let $m_\rho = m$.
3) For simplicity, we choose $W = I_{Nd} - A \otimes I_d$, where $A \in \mathbb{R}^{n\times n}$ is a symmetric doubly stochastic matrix so that $\|A\| \leq 1$ and thus $\|W\| \leq 2$. To satisfy the condition on W in [1], we require $A_{ij} > 0$ if $i = j$ or $\{i,j\} \in \mathcal{E}$ and $A_{ij} = 0$ otherwise.
4) Let $\eta = \frac{1}{2}$ and arbitrarily pick $\rho, \epsilon > 0$.
5) Let $D = \rho W + (\frac{M}{2(1-\eta)} + \frac{(M-m)^2}{4\eta m} + \frac{M-3m}{2} + \epsilon)I_{Nd} = \rho W + (\frac{(M-m)^2}{2m} + \frac{3(M-m)}{2} + \epsilon)I_{Nd}$, which satisfies the condition on D in Lemma 2.
6) Let $\beta = m_\rho\eta = \frac{m}{2}$ so that $\kappa_{\beta,\eta} = \epsilon > 0$.

X. Wu, Z. Qu and J. Lu are with the School of Information Science and Technology, ShanghaiTech University, 201210 Shanghai, China. Email: {wuxy, quzhh1, lujie}@shanghaitech.edu.cn.

This work has been supported by the National Natural Science Foundation of China under grant 61603254.

With the above setting, note that $\lambda_{\max}(D) \leq \frac{(M-m)^2}{2m} + \frac{3(M-m)}{2} + \epsilon + 2\rho =: c$. Thus, for any $c_1, c_2 > 0$,

$$\frac{\rho \lambda_W \kappa_{\beta,\eta}}{(1+c_1)\|\Lambda_M + D\|^2} \geq \frac{\rho \epsilon \lambda_W}{(1+c_1)(M+c)^2} =: Q_1(c_1, c_2),$$
$$\frac{1-\eta}{(1+1/c_1)(1+c_2)} = \frac{1}{2(1+1/c_1)(1+c_2)} =: Q_2(c_1, c_2),$$
$$\frac{2\eta m_\rho - \beta}{\lambda_{\max}(R + (1+1/c_1)(1+1/c_2)\Lambda_M^2/(\rho\lambda_W))} \geq \frac{m}{2c + M + m + 2(1+1/c_1)(1+1/c_2)M^2/(\rho\lambda_W)}$$
$$:= Q_3(c_1, c_2).$$

Therefore,

$$\|\mathbf{z}^{k+1} - \mathbf{z}^\star\|_Q^2 \leq (1-\hat{\delta})\|\mathbf{z}^k - \mathbf{z}^\star\|_Q^2,$$

where

$$\hat{\delta} = \sup_{c_1, c_2 > 0} \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\} \leq \delta.$$

Note that for any $c_1, c_2 > 0$, $\min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\} \in (0, 1)$. Hence, $0 < \hat{\delta} < 1$ and we can use $\hat{\delta}$ to bound the convergence rate of SoPro.

It can be seen that $\hat{\delta}$ only depends on m, M, and λ_W. Below, we discuss how these factors affect the value of $\hat{\delta}$. To this end, consider the following lemma:

Lemma 1. *Suppose* $\sup_{c_1, c_2 > 0} \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\} = \min\{Q_1(c_1^\star, c_2^\star), Q_2(c_1^\star, c_2^\star), Q_3(c_1^\star, c_2^\star)\}$ *for some* $c_1^\star, c_2^\star > 0$. *Then,* $Q_1(c_1^\star, c_2^\star) = Q_2(c_1^\star, c_2^\star) = Q_3(c_1^\star, c_2^\star)$.

Proof. See Appendix A. □

Based on Lemma 1, we derive the lemma below, which provides a sufficient condition on the strict increase of $\hat{\delta}$.

Lemma 2. *Suppose* $Q_i(c_1, c_2)$ $\forall i = 1, 2, 3$ *are associated with problem* $\mathcal{P}$ *and network* $\mathcal{G}$. *Let* $Q_i'(c_1, c_2)$ $\forall i = 1, 2, 3$ *be defined in the same way as* $Q_i(c_1, c_2)$ $\forall i = 1, 2, 3$ *but correspond to a different problem* $\mathcal{P}'$ *and a different network* $\mathcal{G}'$. *If (i)* $Q_i'(c_1, c_2) \geq Q_i(c_1, c_2)$ $\forall i = 1, 2, 3$ *and (ii) there exists* $j \in \{1, 2, 3\}$ *such that* $Q_j'(c_1, c_2) > Q_j(c_1, c_2)$, *then*

$$\sup_{c_1, c_2 > 0} \min\{Q_1'(c_1, c_2), Q_2'(c_1, c_2), Q_3'(c_1, c_2)\} > \sup_{c_1, c_2 > 0} \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}. \quad (2)$$

Proof. See Appendix B. □

For any $c_1, c_2 > 0$, observe that

1) $Q_1(c_1, c_2)$ and $Q_3(c_1, c_2)$ increase if m increases, M decreases, or λ_W increases.

2) $Q_2(c_1, c_2)$ is independent of m, M, and λ_W.

From Lemma 2, we thereby conclude that **larger m, smaller M, or larger λ_W leads to larger $\hat{\delta}$, which implies faster convergence of SoPro.** Here, m and M are the problem characteristics and λ_W indicates how well the network is connected (which plays a similar role as the algebraic connectivity).

Appendix

A. *Proof of Lemma 1*

Equivalently, we show that given $c_1, c_2 > 0$, if $Q_i(c_1, c_2)$ $\forall i = 1, 2, 3$ are nonidentical, then there exist $c_1', c_2' > 0$ such that

$$\min\{Q_1(c_1', c_2'), Q_2(c_1', c_2'), Q_3(c_1', c_2')\} > \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}. \quad (3)$$

To this end, *suppose* $Q_i(c_1, c_2)$ $\forall i = 1, 2, 3$ *are nonidentical* and consider the following three mutually exclusive and exhaustive cases:

Case 1: $Q_1(c_1, c_2) = \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\}$, where $Q_2(c_1, c_2) \neq Q_3(c_1, c_2)$.

We execute the following two steps to find $c_1', c_2' > 0$ such that (3) holds:

Step 1: Find $c_2' > 0$ such that $\min\{Q_2(c_1, c_2'), Q_3(c_1, c_2')\} > \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\}$. To do so, consider two subcases as follows:

- If $Q_2(c_1, c_2) < Q_3(c_1, c_2)$, then decrease c_2 to some $c_2' \in (0, c_2)$ satisfying $Q_2(c_1, c_2') = Q_3(c_1, c_2')$.
- If $Q_2(c_1, c_2) > Q_3(c_1, c_2)$, then increase c_2 to some $c_2' \in (c_2, \infty)$ satisfying $Q_2(c_1, c_2') = Q_3(c_1, c_2')$.

Because $Q_2(c_1, c_2)$ increases with the decrease of c_2 and goes to 0 as $c_2 \to \infty$ and because $Q_3(c_1, c_2)$ increases with the increase of c_2 and goes to 0 as $c_2 \to 0$, c_2' in Step 1 exists. Also, since Q_1 is independent of c_2, $Q_1(c_1, c_2') = Q_1(c_1, c_2) = \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\} < \min\{Q_2(c_1, c_2'), Q_3(c_1, c_2')\}$.

Step 2: Find $c_1' > 0$ such that $Q_1(c_1, c_2') < Q_1(c_1', c_2') = \min\{Q_2(c_1', c_2'), Q_3(c_1', c_2')\}$. To do so, go to Case 2 (with c_2 replaced by c_2').

With the above two steps, $\min\{Q_1(c_1', c_2'), Q_2(c_1', c_2'), Q_3(c_1', c_2')\} = Q_1(c_1', c_2') > Q_1(c_1, c_2') = Q_1(c_1, c_2) = \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}$, i.e., (3) holds.

Case 2: $Q_1(c_1, c_2) < \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\}$. To meet (3), decrease c_1 to some $c_1' \in (0, c_1)$ satisfying $Q_1(c_1', c_2) = \min\{Q_2(c_1', c_2), Q_3(c_1', c_2)\}$. Because $Q_1(c_1, c_2)$ increases with the decrease of c_1 and goes to 0 as $c_1 \to \infty$ and because $Q_2(c_1, c_2)$ and $Q_3(c_1, c_2)$ decrease with the decrease of c_1 and both go to 0 as $c_1 \to 0$, such c_1' exists. Since $\min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\} = Q_1(c_1, c_2) < Q_1(c_1', c_2) = \min\{Q_1(c_1', c_2), Q_2(c_1', c_2), Q_3(c_1', c_2)\}$, (3) holds by letting $c_2' = c_2$.

Case 3: $Q_1(c_1, c_2) > \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\}$. In this case, increase c_1 to some $c_1' \in (c_1, \infty)$ satisfying $Q_1(c_1', c_2) = \min\{Q_2(c_1', c_2), Q_3(c_1', c_2)\}$. Similar to Case 2, such c_1' exists. Since

$\min\{Q_1(c_1', c_2), Q_2(c_1', c_2), Q_3(c_1', c_2)\} = \min\{Q_2(c_1', c_2), Q_3(c_1', c_2)\} > \min\{Q_2(c_1, c_2), Q_3(c_1, c_2)\} = \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}$, (3) holds by letting $c_2' = c_2$.

From the above three cases, for any $c_1, c_2 > 0$, as long as $Q_i(c_1, c_2)\ \forall i = 1, 2, 3$ are not identical, the value of $\min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}$ can be increased by tuning c_1 and c_2. This completes the proof.

B. Proof of Lemma 2

Let $c_1^\star, c_2^\star > 0$ be such that

$$\min\{Q_1(c_1^\star, c_2^\star), Q_2(c_1^\star, c_2^\star), Q_3(c_1^\star, c_2^\star)\} = \sup_{c_1, c_2 > 0} \min\{Q_1(c_1, c_2), Q_2(c_1, c_2), Q_3(c_1, c_2)\}.$$

From Lemma 1, $Q_1(c_1^\star, c_2^\star) = Q_2(c_1^\star, c_2^\star) = Q_3(c_1^\star, c_2^\star)$. Suppose conditions (i) and (ii) in the statement of Lemma 2 hold. Consider the following two mutually exclusive and exhaustive cases:

- $Q_i'(c_1^\star, c_2^\star) > Q_i(c_1^\star, c_2^\star)\ \forall i = 1, 2, 3$. In this case, $\sup_{c_1, c_2 > 0} \min\{Q_1'(c_1, c_2), Q_2'(c_1, c_2), Q_3'(c_1, c_2)\} \geq \min\{Q_1'(c_1^\star, c_2^\star), Q_2'(c_1^\star, c_2^\star), Q_3'(c_1^\star, c_2^\star)\} > \min\{Q_1(c_1^\star, c_2^\star), Q_2(c_1^\star, c_2^\star), Q_3(c_1^\star, c_2^\star)\}$.
- There exists $i \in \{1, 2, 3\}$ such that $Q_i'(c_1^\star, c_2^\star) = Q_i(c_1^\star, c_2^\star)$. In this case, due to condition (ii), there exists $j \in \{1, 2, 3\} \setminus \{i\}$ such that $Q_j'(c_1^\star, c_2^\star) > Q_j(c_1^\star, c_2^\star) = Q_i(c_1^\star, c_2^\star)$. Thus, $Q_j'(c_1^\star, c_2^\star) \neq Q_i'(c_1^\star, c_2^\star)$. It follows from Lemma 1 that

$$\sup_{c_1, c_2 > 0} \min\{Q_1'(c_1, c_2), Q_2'(c_1, c_2), Q_3'(c_1, c_2)\} > \min\{Q_1'(c_1^\star, c_2^\star), Q_2'(c_1^\star, c_2^\star), Q_3'(c_1^\star, c_2^\star)\}.$$

In both of the above cases, (2) holds.

References

[1] X. Wu, Z. Qu, and J. Lu, "A second-order proximal algorithm for consensus optimization," 2019, submitted to *IEEE Transactions on Automatic Control*.