A Partition-Based Relaxed ADMM for Distributed Convex Optimization over Lossy Networks: Technical Proofs

N. Bastianello, R. Carli, L. Schenato, M. Todescato

Appendix

In this paper we describe the technical proofs for the results presented in [1].

A. Proof of Proposition 1

As we showed in Section III-A of the main paper, it is possible to reformulate the partition-based problem (8) so that it conforms to problem

$$\begin{aligned} &\min_{\mathbf{x}} \ \{f(\mathbf{x}) + \iota_{(I-P)}(\mathbf{y})\} \\ &\text{s.t.} \quad A\mathbf{x} + \mathbf{y} = 0 \end{aligned} \tag{A1}$$

to which the R-ADMM can be applied. The three update equations (4), (5) and (6) that characterize the R-ADMM applied to problem (A1) yield

$$\begin{aligned} \mathbf{y}(k+1) = \arg\min_{\mathbf{y}} \{&\mathcal{L}_\rho(\mathbf{x}(k), \mathbf{y}; \mathbf{w}(k)) \\ &+ \rho(2\alpha - 1)\langle \mathbf{y}, (A\mathbf{x}(k) + \mathbf{y}(k))\rangle\} \end{aligned} \tag{A2}$$

$$\begin{aligned} \mathbf{w}(k+1) = \mathbf{w}(k) &- \rho(A\mathbf{x}(k) + \mathbf{y}(k+1)) \\ &- \rho(2\alpha - 1)(A\mathbf{x}(k) + \mathbf{y}(k)) \end{aligned} \tag{A3}$$

$$\mathbf{x}(k+1) = \arg\min_{\mathbf{x}} \mathcal{L}_\rho(\mathbf{x}, \mathbf{y}(k+1); \mathbf{w}(k+1)) \tag{A4}$$

where $\mathbf{w}$ is the vector of Lagrange multipliers and the augmented Lagrangian is

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x}, \mathbf{y}; \mathbf{w}) = f(\mathbf{x}) &+ \iota_{(I-P)}(\mathbf{y}) - \mathbf{w}^\top (A\mathbf{x} + \mathbf{y}) \\ &+ \frac{\rho}{2} \|A\mathbf{x} + \mathbf{y}\|^2 . \end{aligned}$$

However, as shown in [2], the R-ADMM for problem (A1) can be equivalently characterized with the set of four iterates

$$\mathbf{y}(k) = \arg\min_{\mathbf{y} = P\mathbf{y}} \left\{ -\mathbf{z}^\top(k)\mathbf{y} + \frac{\rho}{2}\|\mathbf{y}\|^2 \right\} \tag{A5}$$

$$\mathbf{w}(k) = \mathbf{z}(k) - \rho\mathbf{y}(k) \tag{A6}$$

$$\begin{aligned} \mathbf{x}(k) = \arg\min_{\mathbf{x}} \Big\{ &f(\mathbf{x}) - (2\mathbf{w}(k) - \mathbf{z}(k))^\top A\mathbf{x} \\ &+ \frac{\rho}{2}\|A\mathbf{x}\|^2 \Big\} \end{aligned} \tag{A7}$$

$$\mathbf{z}(k+1) = (1 - 2\alpha)\mathbf{z}(k) + 2\alpha(\mathbf{w}(k) - \rho A\mathbf{x}(k)). \tag{A8}$$

Similarly to what has been done in [3], it is now possible to leverage the distributed nature of problem (A1) in order to simplify Equations (A5)–(A8).

The authors are with the Department of Information Engineering, University of Padova, via Gradenigo 6/b 35131, Padova, Italy. nicola.bastianello.3@studenti.unipd.it, [carlirug|schenato]@dei.unipd.it, mrc.todescato@gmail.com

First of all, solving the system of KKT conditions for (A5) yields $\mathbf{y}(k) = (I + P)\mathbf{z}(k)/(2\rho)$, and therefore Equations (A5)–(A8) become

$$\mathbf{y}(k) = (I + P)\mathbf{z}(k)/(2\rho) \tag{A9}$$

$$\mathbf{w}(k) = (I - P)\mathbf{z}(k)/2 \tag{A10}$$

$$\mathbf{x}(k) = \arg\min_{\mathbf{x}} \left\{ f(\mathbf{x}) + (P\mathbf{z}(k))^\top A\mathbf{x} + \frac{\rho}{2}\|A\mathbf{x}\|^2 \right\} \tag{A11}$$

$$\mathbf{z}(k+1) = (1 - \alpha)\mathbf{z}(k) - \alpha P\mathbf{z}(k) - 2\alpha\rho A\mathbf{x}(k). \tag{A12}$$

Since we are interested in the trajectory $k \to \mathbf{x}(k)$ and by the fact that the update (A11) depends only on the vector $\mathbf{z}(k)$, then the R-ADMM for problem (A1) can be described by Equations (A11) and (A12) only.

Notice now that the trajectory $k \to \mathbf{x}(k)$ generated by (A11) is equivalent to that generated by (A4) if the initial condition for $\mathbf{x}$ is the same and if $\mathbf{z}(0) = \mathbf{w}(0) + \rho\mathbf{y}(0)$ since Equation (A6) has to hold at time $k = 0$. Therefore Propositon 1 is proved if we can show that (A11) and (A12) can be rewritten as (11) and (12).

Recall that the permutation matrix P swaps the element $z_i^{(i,j)}$ with the element $z_i^{(j,i)}$ of vector $\mathbf{z}$, and that the row of $A\mathbf{x}$ relative to the auxiliary variable $z_i^{(j,i)}$ is $-x_i^{(i)}$. Therefore it follows that

$$\begin{aligned} (P\mathbf{z})^\top A\mathbf{x} &= \begin{bmatrix} \cdots & z_i^{(j,i)\top} & \cdots & z_i^{(i,j)\top} & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ -x_i^{(i)} \\ \vdots \\ -x_i^{(j)} \\ \vdots \end{bmatrix} \\ &= -\sum_{i=1}^{N} \left\{ \sum_{j \in \mathcal{N}_i} z_i^{(i,j)\top} x_i^{(i)} + \sum_{j \in \mathcal{N}_i} z_j^{(i,j)\top} x_j^{(i)} \right\}. \end{aligned}$$

Moreover, for each node i $x_i^{(i)}$ appears in $|\mathcal{N}_i|$ constraints and $\{x_j^{(i)}\}_{j \in \mathcal{N}_i}$, in one constraint each. Hence we have

$$\|A\mathbf{x}\|^2 = |\mathcal{N}_i| \left\| x_i^{(i)} \right\|^2 + \sum_{j \in \mathcal{N}_i} \left\| x_j^{(i)} \right\|^2 .$$

Therefore Equations (11) and (12) can be derived from (A11) and (A12) using the particular structure of the problem, proving Proposition 1. □

B. Proof of Propositions 2 and 3

As was mentioned above, the partition-based problem can be reformulated as (A1) which can be solved by the application of the R-ADMM. Therefore both the convergence results of Propositions 2 and 3 follow from those of Propositions 2 and 3 of [3].

Indeed the R-ADMM is guaranteed to converge in both the loss-less and lossy scenarios as long as the step-size and penalty parameters are such that $0 < \alpha < 1$ and $\rho > 0$. Moreover, the components of the primal variables vector, which in the partition-based case are the subvectors $\mathbf{x}^{(i)}$, are guaranteed to converge to the optimum value, that is, each variable $x_i^{(i)}$ converges to the optimum x_i^*. $\square$

References

[1] N. Bastianello, R. Carli, L. Schenato, and M. Todescato, "A partition-based implementation of the relaxed admm for distributed convex optimization over lossy networks," in *IEEE 57th Annual Conference on Decision and Control (CDC), 2018 [submitted]*. IEEE, 2018.

[2] D. Davis and W. Yin, "Convergence rate analysis of several splitting schemes," in *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, 2016, pp. 115–163.

[3] N. Bastianello, M. Todescato, R. Carli, and L. Schenato, "Distributed optimization over lossy networks via relaxed peaceman-rachford splitting: a robust admm approach," in *European Control Conference (ECC), 2018*. IEEE, 2018.