A Unified Algorithmic Framework for Distributed Adaptive Signal and Feature Fusion Problems — Part II: Convergence Properties: Supplementary Material

Cem Ates Musluoglu, Charles Hovine, and Alexander Bertrand, *Senior Member, IEEE*

This supplementary material details the steps we follow to prove Lemma 3 from [1], i.e., to show that $\text{rank}(\mathbf{H})=KJ-J$ which implies that the equation $\mathbf{H}\cdot\boldsymbol{\lambda}_{\mathcal{K}}=0$ (equation (82) in [1]) can only have solutions $\boldsymbol{\lambda}_{\mathcal{K}}=[\boldsymbol{\lambda}^T(1),\dots,\boldsymbol{\lambda}^T(K)]^T$ of the form $\boldsymbol{\lambda}(1)=\dots=\boldsymbol{\lambda}(K)$ when Condition 1b (repeated below) is satisfied.

Condition 1b. *For a fixed point $\overline{X}$ of Algorithm 1, the elements of the set $\{D_{j,q}(\overline{X})\}_{j\in\mathcal{J}}$ are linearly independent for any q, where*

$$D_{j,q}(\overline{X})=\begin{bmatrix}\overline{X}_q^T\nabla_{X_q}h_j(\overline{X})\\ \sum\limits_{k\in\mathcal{B}_{n_1q}}\overline{X}_k^T\nabla_{X_k}h_j(\overline{X})\\ \vdots\\ \sum\limits_{k\in\mathcal{B}_{n_{|\mathcal{N}_q|}q}}\overline{X}_k^T\nabla_{X_k}h_j(\overline{X})\end{bmatrix},\tag{99}$$

which is a block-matrix containing $(1+|\mathcal{N}_q|)$ blocks of $Q\times Q$ matrices.

The proof will be accompanied by an example network topology, given in Figure 3, to visualize the structure of some large matrices in the proof, yet we keep the proof itself generic.

At a fixed point $\overline{X}=X^{i+1}=X^i$ of the DASF algorithm, we have shown in Appendix B in [1] that at each node $q\in\mathcal{K}$ the local stationarity conditions can be written as

$$\overline{X}_q^T\nabla_{X_q}f(\overline{X})=-\sum_{j\in\mathcal{J}}\lambda_j(q)\overline{X}_q^T\nabla_{X_q}h_j(\overline{X}),\tag{100}$$

$$\sum_{k\in\mathcal{B}_{nq}}\overline{X}_k^T\nabla_{X_k}f(\overline{X})=-\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{B}_{nq}}\lambda_j(q)\overline{X}_k^T\nabla_{X_k}h_j(\overline{X}),\tag{101}$$

$\forall n\in\mathcal{N}_q$, leading to the equations

$$\begin{aligned}&\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{B}_{nq}}\lambda_j(k)\overline{X}_k^T\nabla_{X_k}h_j(\overline{X})\\&=\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{B}_{nq}}\lambda_j(q)\overline{X}_k^T\nabla_{X_k}h_j(\overline{X}),\ \forall n\in\mathcal{N}_q.\end{aligned}\tag{102}$$

For clarity, we vectorize the the matrices $\overline{X}_k^T\nabla_{X_k}h_j(\overline{X})$ such that $\mathbf{h}_{j,k}=\text{vec}\left(\overline{X}_k^T\nabla_{X_k}h_j(\overline{X})\right)\in\mathbb{R}^{Q^2}$ and create the matrix $H_k=[\mathbf{h}_{1,k},\dots,\mathbf{h}_{J,k}]\in\mathbb{R}^{Q^2\times J}\ \forall k\in\mathcal{K}$. Note that the linear independence condition of $\{D_{j,q}\}_j$ for any node q is then equivalent to

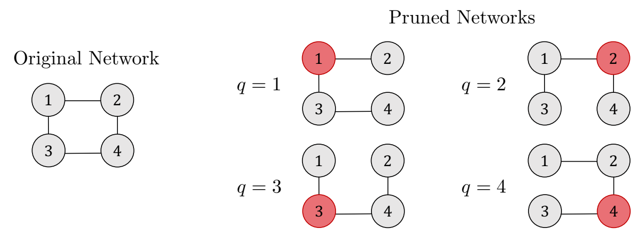

Fig. 3. The example 4−node network we will use to illustrate the equations.

$$\mathbf{D_q}=\begin{bmatrix}H_q\\ \sum_{k\in\mathcal{B}_{n_1q}}H_k\\ \vdots\\ \sum_{k\in\mathcal{B}_{n_{|\mathcal{N}_q|}q}}H_k\end{bmatrix}\tag{103}$$

being full rank, i.e., $\text{rank}(\mathbf{D_q})=J$. Then, (102) can be rewritten as

$$\sum_{k\in\mathcal{B}_{nq}}H_k\boldsymbol{\lambda}(k)=\left(\sum_{k\in\mathcal{B}_{nq}}H_k\right)\boldsymbol{\lambda}(q),\ \forall n\in\mathcal{N}_q,\tag{104}$$

where $\boldsymbol{\lambda}(k)=[\lambda_1(k),\dots,\lambda_J(k)]^T$. For example, in the example network from Figure 3, we have for $q=1$:

$$H_2\boldsymbol{\lambda}(2)=H_2\boldsymbol{\lambda}(1),\tag{105}$$

$$H_3\boldsymbol{\lambda}(3)+H_4\boldsymbol{\lambda}(4)=(H_3+H_4)\boldsymbol{\lambda}(1).\tag{106}$$

Equation (104) corresponds to the linear system given as

$$\mathbf{H}_{nq}\cdot\boldsymbol{\lambda}_{\mathcal{K}}=0,\tag{107}$$

with $\mathbf{H}_{nq}$ a block-row matrix, where each block corresponds to a node $l\in\mathcal{K}$:

$$\mathbf{H}_{nq}(l)=\begin{cases}-\sum_{k\in\mathcal{B}_{nq}}H_k & \text{if } l=q\\ H_l & \text{if } l\in\mathcal{B}_{nq}\\ 0 & \text{otherwise.}\end{cases}\tag{108}$$

We can then stack vertically every $\mathbf{H}_{nq}$ for every neighbor $n\in\mathcal{N}_q$ of q, and for every node q to obtain

$$\mathbf{H}\cdot\boldsymbol{\lambda}_{\mathcal{K}}=0.\tag{109}$$

In the example network of Figure 3, we have the matrix given in (110), where we separated by horizontal lines the blocks corresponding to each different node for clarity.

This paper is a supplementary material to [1].

$$
\mathbf{H}=\begin{bmatrix}\mathbf{H}_{2,1}\\\mathbf{H}_{3,1}\\\hline\mathbf{H}_{1,2}\\\mathbf{H}_{4,2}\\\hline\mathbf{H}_{1,3}\\\mathbf{H}_{4,3}\\\hline\mathbf{H}_{2,4}\\\mathbf{H}_{3,4}\end{bmatrix}=\left[\begin{array}{c|c|c|c}-H_2 & H_2 & 0 & 0\\ -H_3-H_4 & 0 & H_3 & H_4\\\hline H_1 & -H_1-H_3 & H_3 & 0\\ 0 & -H_4 & 0 & H_4\\\hline H_1 & 0 & -H_1 & 0\\ 0 & H_2 & -H_2-H_4 & H_4\\\hline H_1 & H_2 & 0 & -H_1-H_2\\ 0 & 0 & H_3 & -H_3\end{array}\right] \tag{110}
$$

$$
\widetilde{\mathbf{H}}=\begin{bmatrix}\mathbf{H}_{2,1}\\\mathbf{H}_{3,1}\\\mathbf{H}_{\Sigma 4}\\\hline\mathbf{H}_{1,2}\\\mathbf{H}_{4,2}\\\mathbf{H}_{\Sigma 4}\\\hline\mathbf{H}_{1,3}\\\mathbf{H}_{4,3}\\\mathbf{H}_{\Sigma 4}\end{bmatrix}=\left[\begin{array}{c|c|c|c}-H_2 & H_2 & 0 & 0\\ -H_3-H_4 & 0 & H_3 & H_4\\ H_1 & H_2 & H_3 & -H_1-H_2-H_3\\\hline H_1 & -H_1-H_3 & H_3 & 0\\ 0 & -H_4 & 0 & H_4\\ H_1 & H_2 & H_3 & -H_1-H_2-H_3\\\hline H_1 & 0 & -H_1 & 0\\ 0 & H_2 & -H_2-H_4 & H_4\\ H_1 & H_2 & H_3 & -H_1-H_2-H_3\end{array}\right] \tag{116}
$$

Note that $\mathbf{H}$ is a $Q^2\sum_{k\in\mathcal{K}}|\mathcal{N}_k|\times KJ$ matrix and since $\forall q\in\mathcal{K}$, $n\in\mathcal{N}_q$,

$$
\sum_{k\in\mathcal{K}\setminus\{l\}}\mathbf{H}_{nq}(k)=-\mathbf{H}_{nq}(l) \tag{111}
$$

$\forall l\in\mathcal{K}$, all $\boldsymbol{\lambda}_{\mathcal{K}}\in\mathbb{R}^{KJ}$ such that $\boldsymbol{\lambda}(1)=\cdots=\boldsymbol{\lambda}(K)$ is a solution of (109). A basis for this solution space is given by

$$
\mathcal{E}=\left\{\begin{bmatrix}\mathbf{e}_1\\\vdots\\\mathbf{e}_1\end{bmatrix},\begin{bmatrix}\mathbf{e}_2\\\vdots\\\mathbf{e}_2\end{bmatrix},\dots\begin{bmatrix}\mathbf{e}_J\\\vdots\\\mathbf{e}_J\end{bmatrix}\right\}\subset\mathbb{R}^{KJ}, \tag{112}
$$

where $\mathbf{e}_k$'s represent the standard basis for $\mathbb{R}^J$. Since $\mathcal{E}$ spans a $J-$dimensional space, we have

$$
\text{rank}(\mathbf{H})\leq KJ-J. \tag{113}
$$

To show that the solution of (109) all satisfy $\boldsymbol{\lambda}(1)=\cdots=\boldsymbol{\lambda}(K)$, we need to show that all solutions are in $\text{span}(\mathcal{E})$, i.e., $\text{rank}(\mathbf{H})=KJ-J$, which is stated in Lemma 3 from [1], repeated below and followed by a proof. By the dimensions of $\mathbf{H}$, this is only possible if we have $Q^2\sum_{k\in\mathcal{K}}|\mathcal{N}_k|\geq KJ-J$ or equivalently $J\leq\frac{Q^2}{K-1}\sum_{k\in\mathcal{K}}|\mathcal{N}_k|$.

Lemma 3. *If Condition 1b holds, then rank*$(\mathbf{H})=KJ-J$.

Proof. Let us take the bottom part of the matrix $\mathbf{H}$ corresponding to the neighbors of the last node K, and take the sum over its block-rows $\mathbf{H}_{nK}$, $n\in\mathcal{N}_K$, defined in (108), which results in a new summed block row

$$
\mathbf{H}_{\Sigma K}(l)=\begin{cases}-\sum_{k\in\mathcal{K}\setminus\{K\}}H_k & \text{if } l=K\\ H_l & \text{if } l\neq K.\end{cases} \tag{114}
$$

In the example network, we have $K=4$, therefore

$$
\mathbf{H}_{\Sigma 4}=\left[H_1\ \middle|\ H_2\ \middle|\ H_3\ \middle|\ -H_1-H_2-H_3\ \right]. \tag{115}
$$

We then insert vertically the matrices $\mathbf{H}_{\Sigma K}$ after each block $\mathbf{H}_{n_{|\mathcal{N}_q|}q}$, $q\neq K$, of $\mathbf{H}$. Note that this insertion cannot change the rank of the matrix, as all the inserted rows are sums of existing rows in $\mathbf{H}$. We also remove the block-rows $[\mathbf{H}_{n_1K}^T,\dots,\mathbf{H}_{n_{|\mathcal{N}_K|}K}^T]^T$, i.e., the block-rows corresponding to node K, to obtain a new matrix $\widetilde{\mathbf{H}}$. For the example in Figure 3, this results in the matrix given in (116). Note that we have

$$
\text{rank}(\mathbf{H})\geq\text{rank}(\widetilde{\mathbf{H}}) \tag{117}
$$

since the removal of rows can only reduce the rank.

We will first look at the rank of $\widetilde{\mathbf{H}}$ and derive from it the rank of $\mathbf{H}$. For this, we will apply the Gaussian elimination method using elementary row operations (EROs) which are known to not change the rank. Referring to the block decomposition of example (116), $\widetilde{\mathbf{H}}$ has K block-columns, and each of these block-columns of $\widetilde{\mathbf{H}}$ will be referred to as **the block-column at position** $k\in\mathcal{K}$. We refer to the matrix $[\mathbf{H}_{n_1q}^T,\dots,\mathbf{H}_{n_{|\mathcal{N}_q|}q}^T,\mathbf{H}_{\Sigma K}]^T$, $q\neq K$, and the resulting matrices obtained by applying EROs to it as **the submatrix corresponding to node** q. For example, in (116) the block-column at position 3 of the submatrix corresponding to node 2 is equal to $[H_3^T,0,H_3^T]^T$.

For each $q\neq K$, let us sum all block-rows $\mathbf{H}_{nq}$, $n\in\mathcal{N}_q$. The result is then subtracted from the block-row $\mathbf{H}_{\Sigma K}$ leading to the final block-row of the submatrix corresponding to each node q being of the form $[0|\dots|0|\sum_{k\in\mathcal{K}}H_k|0|\dots|0|-\sum_{k\in\mathcal{K}}H_k]$, where the first non-zero matrix is at position $q\neq K$. For our example network, we obtain

$$
\left[\begin{array}{c|c|c|c}-H_2 & H_2 & 0 & 0\\ -H_3-H_4 & 0 & H_3 & H_4\\ \sum_{k=1}^4 H_k & 0 & 0 & -\sum_{k=1}^4 H_k\\\hline H_1 & -H_1-H_3 & H_3 & 0\\ 0 & -H_4 & 0 & H_4\\ 0 & \sum_{k=1}^4 H_k & 0 & -\sum_{k=1}^4 H_k\\\hline H_1 & 0 & -H_1 & 0\\ 0 & H_2 & -H_2-H_4 & H_4\\ 0 & 0 & \sum_{k=1}^4 H_k & -\sum_{k=1}^4 H_k\end{array}\right]. \tag{118}
$$

An important observation is that, for each $q \neq K$, the block-column at position q corresponding to the submatrix of node q can be obtained by applying EROs to $\mathbf{D_q}$ defined in (103), i.e., it is equal to $\mathbf{D_q}$ up to EROs. Since rank$(\mathbf{D_q}) = J$ by Condition 1b, we can apply the necessary EROs to obtain the following reduced echelon form for the submatrix corresponding to node $q = 1$:

$$\left[\begin{array}{c|c|c|c|c} \begin{matrix} I_J \\ 0 \end{matrix} & * & \cdots & * & * \end{array}\right]. \tag{119}$$

We can use the J pivots in the first J columns of (119) to create zeros at all the entries underneath (in the submatrices corresponding to $q \neq 1$) using EROs. As a result, we obtain

$$\left[\begin{array}{c|c|c|c|c} \begin{matrix} I_J \\ 0 \end{matrix} & * & \cdots & * & * \\ \hline 0 & \mathcal{R}\mathbf{D_2} & \cdots & * & * \\ \hline \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & * & \cdots & \mathcal{R}\mathbf{D_{K-1}} & * \end{array}\right], \tag{120}$$

where $\mathcal{R}\mathbf{D_k}$ is a matrix equal to $\mathbf{D_k}$ up to EROs. To see why (120) holds, i.e., that each block-column at position $k \notin \{1, K\}$ of the submatrix corresponding to node k is indeed equal to $\mathbf{D_k}$ up to EROs, we note that, initially, every row above this block is either full of zeros or a row of H_k (see (118) for a visual example). Therefore, the EROs we applied to the full matrix to create zeros at all entries underneath the J pivots in the block-column corresponding to node $q = 1$ do not change the fact that the block-column at position k of the submatrix corresponding to node k is equal to $\mathbf{D_k}$ up to EROs. Since EROs do not change the rank of a matrix, the submatrix $\mathcal{R}\mathbf{D_2}$ should again have rank J and so we can again create J pivots to create zeros underneath. Repeating this process for $2 \leq q \leq K - 1$, we obtain

$$\left[\begin{array}{c|c|c|c|c} \begin{matrix} I_J \\ 0 \end{matrix} & * & \cdots & * & * \\ \hline 0 & \begin{matrix} I_J \\ 0 \end{matrix} & \cdots & * & * \\ \hline \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & \cdots & \begin{matrix} I_J \\ 0 \end{matrix} & * \end{array}\right]. \tag{121}$$

Since there are at least $K - 1$ block-columns containing J pivots, rank$(\widetilde{\mathbf{H}}) \geq KJ - J$. We previously established in (117) that rank$(\mathbf{H}) \geq$ rank$(\widetilde{\mathbf{H}})$, hence rank$(\mathbf{H}) \geq KJ - J$. We also already established in (113) that rank$(\mathbf{H}) \leq KJ - J$, and therefore it should hold that rank$(\mathbf{H}) = KJ - J$, which is what had to be proven. $\square$

REFERENCES

[1] C. A. Musluoglu, C. Hovine, and A. Bertrand, “A unified algorithmic framework for distributed adaptive signal and feature fusion problems — Part II: Convergence properties,” 2022.