Supplementary materials: Algorithms derivation for "Detection Methods based on Structured Covariance Matrices for Multivariate SAR Images Processing"

R. Ben Abdallah, A. Mian, A. Breloy, A. Taylor, M. N. El Korso, D. Lautru

Abstract

This document is dedicated to the derivation of Majorization-Minimization algorithms to evaluate the value of a GLRT detector proposed in [1]. Section I presents the Majorization-Minimization framework and derive the required surrogates. Section II develops algorithms for the detector referred to as $t_{\mathrm{glr}}^{\mathrm{lrP}}$ (GLRT for LR CM proportionality testing).

I. Block Majorization-Minimization framework

A. Block Majorization-Minimization algorithm

To solve further-coming optimization problems, we adopt the block majorization-minimization (MM) algorithm framework, which is briefly stated below. For more complete information, we refer the reader to [2]. Consider the following problem:

$$\begin{array}{ll} \underset{\boldsymbol{\theta}}{\operatorname{minimize}} & f(\boldsymbol{\theta}) \\ \text{subject to} & \boldsymbol{\theta} \in \boldsymbol{\Theta}, \end{array} \tag{1}$$

where the optimization variable $\boldsymbol{\theta}$ can be partitioned into m blocks as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$, with each n_i-dimensional block $\boldsymbol{\theta}_i \in \boldsymbol{\Theta}_i$ and $\boldsymbol{\Theta} = \prod_{i=1}^m \boldsymbol{\Theta}_i$. At the $(t+1)$-th iteration, the i-th block $\boldsymbol{\theta}_i$ is updated by solving the following problem:

$$\begin{array}{ll} \underset{\boldsymbol{\theta}_i}{\operatorname{minimize}} & g_i\left(\boldsymbol{\theta}_i | \boldsymbol{\theta}^{(t)}\right) \\ \text{subject to} & \boldsymbol{\theta}_i \in \boldsymbol{\Theta}_i, \end{array} \tag{2}$$

with $i = (t \bmod m) + 1$ (so blocks are updated in cyclic order) and the continuous surrogate function $g_i\left(\boldsymbol{\theta}_i | \boldsymbol{\theta}^{(t)}\right)$ satisfying the following properties:

$$\begin{array}{rcl} f\left(\boldsymbol{\theta}^{(t)}\right) & = & g_i\left(\boldsymbol{\theta}_i^{(t)} | \boldsymbol{\theta}^{(t)}\right), \\ f\left(\boldsymbol{\theta}_1^{(t)}, \ldots, \boldsymbol{\theta}_i, \ldots, \boldsymbol{\theta}_m^{(t)}\right) & \leq & g_i\left(\mathbf{x}_i | \boldsymbol{\theta}^{(t)}\right) \ \forall \boldsymbol{\theta}_i \in \boldsymbol{\Theta}_i, \\ f'\left(\boldsymbol{\theta}^{(t)}; \mathbf{d}_i^0\right) & = & g_i'\left(\boldsymbol{\theta}_i^{(t)}; \mathbf{d}_i | \boldsymbol{\theta}^{(t)}\right) \\ & & \forall \boldsymbol{\theta}_i^{(t)} + \mathbf{d}_i \in \boldsymbol{\Theta}_i, \\ & & \mathbf{d}_i^0 \triangleq (\mathbf{0}; \ldots; \mathbf{d}_i; \ldots; \mathbf{0}), \end{array}$$

where $f'(\boldsymbol{\theta}; \mathbf{d})$ stands for the directional derivative at $\boldsymbol{\theta}$ along $\mathbf{d}$. In short, at each iteration, the block MM algorithm updates the variables in one block by minimizing a tight upperbound of the function while keeping the other blocks fixed.

B. Surrogates and updates

In the following, we derive two propositions needed for the algorithms design. These propositions are generic, but their formulation mimics the notations of [1] for convenience. Specifically, the first proposition will be useful to derive updates w.r.t. positive scaling factors and eigenvalues of the structured covariance matrices. The second proposition will be useful for the update of eigenvectors.

Proposition 1 *Let $\tau_i \in \mathbb{R}^+$, and sets of real parameters $\{\lambda_r\}$ with $\lambda_r \geq 0$, $\forall r \in [\![1, R]\!]$, $\{s_{k,i}^r\}$ with $s_{k,i}^r \geq 0$, $\forall i \in [\![1, I]\!]$, $\forall r \in [\![1, R]\!]$, $\forall k \in [\![1, K_i]\!]$, and $\sigma^2 > 0$. The function*

$$\mathcal{L}_i(\tau_i) = \sum_{r=1}^{R} \left[K_i \ln\left(\tau_i \lambda_r + \sigma^2\right) - \sum_{k=1}^{K_i} \frac{\tau_i \ \lambda_r \ s_{k,i}^r}{\tau_i \lambda_r + \sigma^2} \right] \tag{3}$$

is upperbounded at the point $\tau_i^{(t)}$ as

$$\mathcal{L}_i(\tau_i | \tau_i^{(t)}) \leq A_i \ln\left(B_i \tau_i + C_i\right) - D_i \ln\left(\tau_i\right) \tag{4}$$

where

$$\begin{cases} \gamma_i^r = \left(K_i + \sum_{k=1}^{K_i} s_{k,i}^r \frac{\tau_i^{(t)} \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2} \right) & \beta_i^r = \sum_{k=1}^{K_i} s_{k,i}^r \frac{\tau_i^{(t)} \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2} & A_i = \sum_{r=1}^{R} \gamma_i^r \\ B_i = \frac{\sum_{r=1}^{R} \frac{\gamma_i^r \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2}}{\sum_{r=1}^{R} \gamma_i^r} & C_i = \frac{\sum_{r=1}^{R} \frac{\gamma_i^r \sigma^2}{\tau_i^{(t)} \lambda_r + \sigma^2}}{\sum_{r=1}^{R} \gamma_i^r} & D_i = \sum_{r=1}^{R} \beta_i^r \end{cases} \tag{5}$$

with equality at $\tau_i = \tau_i^{(t)}$. The above surrogate function has a unique non-negative minimum that leads to the following MM update:

$$\tau_i^{(t+1)} = \frac{D_i C_i}{(A_i - D_i) B_i} \tag{6}$$

Proof: This generalize Propositions 1 and 2 from [3], which cover the case $K_i = 1$. The proof is in Appendix A. ■

Proposition 2 *Let $\mathbf{V} \in \mathbb{C}^{M \times R}$ be a unitary matrix, $\{\mathbf{B}_i\}$, with $\mathbf{B}_i \in \mathcal{H}_R^{++},\ \forall i \in [\![1, I]\!]$ and $\{\mathbf{A}_i\}$, with $\mathbf{A}_i \in \mathcal{H}_M^{++},\ \forall i \in [\![1, I]\!]$. The function*

$$f(\mathbf{V}) = -\sum_{i=1}^{I} \mathrm{Tr}\{\mathbf{V}^H \mathbf{A}_i \mathbf{V} \mathbf{B}_i\} \tag{7}$$

is upperbounded at point $\mathbf{V}^{(t)}$ as

$$f(\mathbf{V}|\mathbf{V}^{(t)}) \leq -\sum_{i=1}^{I} \left[\mathrm{Tr}\{\mathbf{V}^H \mathbf{A}_i \mathbf{V}^{(t)} \mathbf{B}_i\} + \mathrm{Tr}\{\mathbf{V}^{(t)H} \mathbf{A}_i \mathbf{V} \mathbf{B}_i\} \right] + \mathrm{const.} \tag{8}$$

with equality at $\mathbf{V}^{(t)} = \mathbf{V}$. The above surrogate function has a unique minimum on the set of unitary matrices, that leads to the following MM update

$$\mathbf{V}^{(t+1)} = \mathbf{U}_{\text{left}} \mathbf{U}_{\text{right}}^H \tag{9}$$

where $\mathbf{U}_{\text{left}}$ and $\mathbf{U}_{\text{right}}^H$ are the left and right eigenvectors of the thin-SVD

$$\sum_{i=1}^{I} (\mathbf{A}_i \mathbf{V}^{(t)} \mathbf{B}_i) \overset{\text{TSVD}}{=} \mathbf{U}_{\text{left}} \mathbf{D} \mathbf{U}_{\text{right}}^H \tag{10}$$

Proof: The proof is in Appendix B. ■

II. GLRT for LR CM proportionality testing (Section IV.B of [1])

A. General model and likelihood function

We recall that, in [1], $\mathbf{z}_k^i$ denotes a sample, in which the superscript $i \in [\![0, I]\!]$ refers to the index of a set of i.i.d. variables, and $k \in [\![1, K_i]\!]$ to the index of the sample in this set. We have $\mathbf{z}_k^i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_i)$, so for a given sample set $\{\mathbf{z}_k^i\}$, the likelihood of the dataset reads

$$\mathcal{L}(\{\mathbf{z}_k^i\}|\boldsymbol{\theta}) = \prod_{i=0}^{I} \frac{\mathrm{etr}\{-\mathbf{S}_i \mathbf{\Sigma}_i^{-1}(\boldsymbol{\theta})\}}{|\ \mathbf{\Sigma}_i(\boldsymbol{\theta})\ |^{K_i}}, \tag{11}$$

with $\mathbf{S}_i = \sum_{k=1}^{K_i} \mathbf{z}_k^i \mathbf{z}_k^{iH}$ and where $\boldsymbol{\theta}$ denotes the set of unknown parameters defining the $\mathbf{\Sigma}_i$'s (specified below).

B. Expression of the GLRT

We recall that the hypothesis test derived in section IV.B of [1] for the LR CM proportionality testing is expressed as

$$\begin{cases} \mathcal{H}_0 : \left| \mathbf{\Sigma}_i = \tau_i \mathbf{V}_{\mathcal{H}_0} \mathbf{\Lambda}_{\mathcal{H}_0} (\mathbf{V}_{\mathcal{H}_0})^H + \sigma^2 \mathbf{I},\ \forall i \in [\![0, I]\!] \right. \\ \mathcal{H}_1 : \left| \begin{array}{l} \mathbf{\Sigma}_0 = \ \mathbf{V}_{\mathcal{H}_1}^0 \mathbf{\Lambda}_{\mathcal{H}_1}^0 (\mathbf{V}_{\mathcal{H}_1}^0)^H + \sigma^2 \mathbf{I} \\ \mathbf{\Sigma}_i = \tau_i \mathbf{V}_{\mathcal{H}_1}^\star \mathbf{\Lambda}_{\mathcal{H}_1}^\star (\mathbf{V}_{\mathcal{H}_1}^\star)^H + \sigma^2 \mathbf{I},\ \forall i \in [\![1, I]\!] \end{array} \right. \end{cases} \tag{12}$$

Consequently, its corresponding GLRT expression, denoted $t_{\text{glr}}^{\text{lrP}}$ is given by:

$$\frac{\max_{\boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}}} \mathcal{L}\left(\{\mathbf{z}_k^i\}|\mathcal{H}_1, \boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}}\right)}{\max_{\boldsymbol{\theta}_{\mathcal{H}_0}^{\text{lrP}}} \mathcal{L}(\{\mathbf{z}_k^i\}|\mathcal{H}_0, \boldsymbol{\theta}_{\mathcal{H}_0}^{\text{lrP}})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta_{\text{glr}}^{\text{lrP}}, \tag{13}$$

with sets of parameters

$$\begin{aligned} \boldsymbol{\theta}_{\mathcal{H}_0}^{\text{lrP}} &= \left\{ \{\tau_i\}_{i \in [\![0, I]\!]}, \mathbf{V}_{\mathcal{H}_0}, \mathbf{\Lambda}_{\mathcal{H}_0} \right\}, \\ \boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}} &= \left\{ \{\tau_i\}_{i \in [\![1, I]\!]}, \mathbf{V}_{\mathcal{H}_1}^\star, \mathbf{\Lambda}_{\mathcal{H}_1}^\star, \mathbf{V}_{\mathcal{H}_1}^0, \mathbf{A}_{\mathcal{H}_1}^0 \right\}. \end{aligned} \tag{14}$$

Algorithm 1 MM algorithm to compute $\hat{\boldsymbol{\theta}}^{\mathrm{lrP}}_{\mathcal{H}_0}$ for $t^{\mathrm{lrP}}_{\mathrm{glr}}$ (Section IV.B of [1])

1: **Input**: $\{\mathbf{S}_i, K_i\}$ for $i \in [\![0, I]\!]$, R, and σ^2.
2: **repeat**
3: $\quad t \leftarrow t+1$
4: $\quad$ Update $\{\tau_i^{(t)}\}$, $\forall i \in [\![0, I]\!]$ with (6)
5: $\quad$ Update $\{\lambda_r^{(t)}\}$, $\forall r \in [\![1, R]\!]$ with (21)
6: $\quad$ Update $\hat{\mathbf{V}}^{(t)}_{\mathcal{H}_0}$ with (24)
7: **until** Some convergence criterion is met.
8: **Output**: $\hat{\boldsymbol{\theta}}^{\mathrm{lrP}}_{\mathcal{H}_0} = \left\{ \{\hat{\tau}_i\}_{i \in [\![0,I]\!]},\ \hat{\mathbf{V}}_{\mathcal{H}_0},\ \hat{\boldsymbol{\Lambda}}_{\mathcal{H}_0} \right\}$

C. MM algorithm to optimize the likelihood under $\mathcal{H}_0$ w.r.t. $\boldsymbol{\theta}^{\mathrm{lrP}}_{\mathcal{H}_0}$

Under $\mathcal{H}_0$, we have the following optimization problem:

$$\begin{array}{ll} \underset{\boldsymbol{\theta}^{\mathrm{lrP}}_{\mathcal{H}_0}}{\text{maximize}} & \mathcal{L}(\{\mathbf{z}_k^i\}|\mathcal{H}_0, \boldsymbol{\theta}^{\mathrm{lrP}}_{\mathcal{H}_0}) \\ \text{subject to} & \tau_i \geq 0,\ \forall i \in [\![0, I]\!] \\ & [\boldsymbol{\Lambda}_{\mathcal{H}_0}]_{r,r} \geq 0,\ \forall r \in [\![1, R]\!] \\ & (\mathbf{V}_{\mathcal{H}_0})^H \mathbf{V}_{\mathcal{H}_0} = \mathbf{I}. \end{array} \tag{15}$$

This problem is equivalent to maximizing the negative log-likelihood as

$$\begin{array}{ll} \underset{\boldsymbol{\theta}^{\mathrm{lrP}}_{\mathcal{H}_0}}{\text{minimize}} & \sum_{i=0}^{I} \left[K_i \ln(|\boldsymbol{\Sigma}_i|) + \mathrm{Tr}\left\{ \mathbf{S}_i \boldsymbol{\Sigma_i}^{-1} \right\} \right] \\ \text{subject to} & \boldsymbol{\Sigma}_i = \tau_i \mathbf{V}_{\mathcal{H}_0} \boldsymbol{\Lambda}_{\mathcal{H}_0} (\mathbf{V}_{\mathcal{H}_0})^H + \sigma^2 \mathbf{I},\ \forall i \in [\![0, I]\!] \\ & \tau_i \geq 0,\ \forall i \in [\![0, I]\!] \\ & [\boldsymbol{\Lambda}_{\mathcal{H}_0}]_{r,r} \geq 0,\ \forall r \in [\![1, R]\!] \\ & (\mathbf{V}_{\mathcal{H}_0})^H \mathbf{V}_{\mathcal{H}_0} = \mathbf{I}, \end{array} \tag{16}$$

for which we derive MM updates in the following. The corresponding algorithm is summed up in the box Algorithm 1.

1) Update $\{\tau_i\}$: First, remark that $\boldsymbol{\Sigma}_i^{-1}$ can be expressed thanks to the matrix inversion lemma as:

$$\boldsymbol{\Sigma}_i^{-1} = (\tau_i \mathbf{V}_{\mathcal{H}_0} \boldsymbol{\Lambda}_{\mathcal{H}_0} (\mathbf{V}_{\mathcal{H}_0})^H + \sigma^2 \mathbf{I})^{-1} = \sigma^{-2}\mathbf{I} - \mathbf{V}_{\mathcal{H}_0} \boldsymbol{\Gamma}_i (\mathbf{V}_{\mathcal{H}_0})^H,\ \forall i \in [\![0, I]\!] \tag{17}$$

with

$$\begin{aligned} \boldsymbol{\Gamma}_i &= \mathrm{diag}\left(\alpha_{i,1},\ \ldots,\ \alpha_{i,R} \right) \\ \alpha_{i,r} &= \frac{\tau_i \lambda_r}{\sigma^2(\tau_i \lambda_r + \sigma^2)},\ \forall r \in [\![1, R]\!],\ \forall i \in [\![0, I]\!] \\ \lambda_r &= [\boldsymbol{\Lambda}_{\mathcal{H}_0}]_{r,r},\ \forall r \in [\![1, R]\!] \end{aligned} \tag{18}$$

For other variables fixed, the problem in (16) is separable for each τ_i. After some direct calculus, the objective of (16) w.r.t. the variable τ_i only can be obtained as

$$\mathcal{L}_i(\tau_i) = \sum_{r=1}^{R} \left[K_i \ln\left(\tau_i \lambda_r + \sigma^2 \right) - \sum_{k=1}^{K_i} \frac{\tau_i\ \lambda_r\ s_{k,i}^r}{\tau_i \lambda_r + \sigma^2} \right] + \text{const.} \tag{19}$$

with $s_{k,i}^r = \mathbf{v}_r^H \mathbf{z}_k^i \mathbf{z}_k^{iH} \mathbf{v}_r$ and where $\mathbf{V}_{\mathcal{H}_0} = [\mathbf{v}_1,\ \ldots,\ \mathbf{v}_R]$. With this formulation of the objective over τ_i, we can directly apply Proposition 1 to obtain the update of all τ_i's as in (6).

2) Update $\boldsymbol{\Lambda}_{\mathcal{H}_0}$: First recall that $\boldsymbol{\Lambda}_{\mathcal{H}_0}$ is diagonal and that we have the notation $\lambda_r = [\boldsymbol{\Lambda}_{\mathcal{H}_0}]_{r,r}$, $\forall r \in [\![1, R]\!]$. Following the derivations of the previous section, the objective of (16) is separable in λ_r and reads for each λ_r as:

$$\mathcal{L}_r(\lambda_r) = \sum_{i=1}^{I} \left[K_i \ln(\tau_i \lambda_r + \sigma^2) - \sum_{k=1}^{K_i} \frac{\tau_i \lambda_r s_{k,i}^r}{\tau_i \lambda_r + \sigma^2} \right] + \text{const.} \tag{20}$$

Thus, one can note that $\{\tau_i\}$ and $\{\lambda_r\}$ play similar role in the objective. We can therefore obtain a surrogate by adapting the formulation in Proposition 1. The resulting update for these variables is given as:

$$\lambda_r^{(t+1)} = \frac{D_r C_r}{(A_r - D_r) B_r} \tag{21}$$

Algorithm 2 MM algorithm to compute $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}^{\text{lrP}}$ for $t_{\text{glr}}^{\text{lrP}}$ (Section IV.B of [1])

1: **Input**: $\{\mathbf{S}_i, K_i\}$ for $i \in [\![0, I]\!]$, R, and σ^2.
2: Call Algorithm 1 on the restricted set $\{\mathbf{S}_i\}$ for $i \in [\![1, I]\!]$, the output is $\{\hat{\tau}_i\}, \hat{\mathbf{\Lambda}}_{\mathcal{H}_1}^{\star}, \hat{\mathbf{V}}_{\mathcal{H}_1}^{\star}$
3: Compute $\hat{\mathbf{A}}_{\mathcal{H}_1}^{0}$ and $\hat{\mathbf{V}}_{\mathcal{H}_1}^{0}$ with (29)
4: **Output**: $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}^{\text{lrP}} = \left\{\{\hat{\tau}_i\}_{i\in[\![1,I]\!]}, \hat{\mathbf{V}}_{\mathcal{H}_1}^{\star}, \hat{\mathbf{\Lambda}}_{\mathcal{H}_1}^{\star}, \hat{\mathbf{V}}_{\mathcal{H}_1}^{0}, \hat{\mathbf{A}}_{\mathcal{H}_1}^{0}\right\}$

with

$$\begin{cases} \beta_i^r = \sum_{k=1}^{K_i} s_{k,i}^r \frac{\tau_i \lambda_r^{(t)}}{\tau_i \lambda_r^{(t)} + \sigma^2} & \gamma_i^r = K_i + \sum_{k=1}^{K_i} s_{k,i}^r \frac{\tau_i \lambda_r^{(t)}}{\tau_i \lambda_r^{(t)} + \sigma^2} & A_r = \sum_{i=1}^{I} \gamma_i^r \\ B_r = \frac{\sum_{i=1}^{I} \frac{\gamma_i^r \tau_i}{\tau_i \lambda_r^{(t)} + \sigma^2}}{\sum_{i=1}^{I} \gamma_i^r} & C_r = \frac{\sum_{i=1}^{I} \frac{\gamma_i^r \sigma_n^2}{\tau_i \lambda_r^{(t)} + \sigma^2}}{\sum_{i=1}^{I} \gamma_i^r} & D_r = \sum_{i=1}^{I} \beta_i^r \end{cases} \tag{22}$$

3) Update $\mathbf{V}_{\mathcal{H}_0}$: Using (17), the objective in (16) w.r.t. $\mathbf{V}_{\mathcal{H}_0}$ fixing remaining variables is expressed as

$$\mathcal{L}_v(\mathbf{V}_{\mathcal{H}_0}) = -\sum_{i=0}^{I} \text{Tr}\{(\mathbf{V}_{\mathcal{H}_0})^H \mathbf{S}_i \mathbf{V}_{\mathcal{H}_0} \mathbf{\Gamma}_i\} + \text{const.} \tag{23}$$

with $\mathbf{\Gamma}_i$ in (18). Thus, we can directly apply Proposition 2 to obtain the update

$$\mathbf{V}_{\mathcal{H}_0}^{(t+1)} = \mathbf{U}_{\text{left}} \mathbf{U}_{\text{right}}^H, \quad \text{with} \quad \sum_{i=0}^{I} \left(\mathbf{S}_i \mathbf{V}_{\mathcal{H}_0}^{(t)} \mathbf{\Gamma}_i\right) \overset{\text{TSVD}}{=} \mathbf{U}_{\text{left}} \mathbf{D} \mathbf{U}_{\text{right}}^H \tag{24}$$

D. MM algorithm to optimize the likelihood under $\mathcal{H}_1$ w.r.t. $\boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}}$

Under $\mathcal{H}_1$, we have the following optimization problem:

$$\begin{array}{ll} \underset{\boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}}}{\text{maximize}} & \mathcal{L}(\{\mathbf{z}_k^i\}|\mathcal{H}_1, \boldsymbol{\theta}_{\mathcal{H}_1}^{\text{lrP}}) \\ \text{subject to} & \tau_i \geq 0, \ \forall i \in [\![1, I]\!] \\ & \left[\mathbf{\Lambda}_{\mathcal{H}_1}^{\star}\right]_{r,r} \geq 0, \ \forall r \in [\![1, R]\!] \\ & \left[\mathbf{A}_{\mathcal{H}_1}^{0}\right]_{r,r} \geq 0, \ \forall r \in [\![1, R]\!] \\ & (\mathbf{V}_{\mathcal{H}_1}^{\star})^H \mathbf{V}_{\mathcal{H}_1}^{\star} = \mathbf{I} \\ & (\mathbf{V}_{\mathcal{H}_1}^{0})^H \mathbf{V}_{\mathcal{H}_1}^{0} = \mathbf{I} \end{array} \tag{25}$$

This problem is equivalent to two separate ones in respectively $\{\{\tau_i\}_{i\in[\![1,I]\!]}, \ \mathbf{\Lambda}_{\mathcal{H}_1}^{\star}, \ \mathbf{V}_{\mathcal{H}_1}^{\star}\}$ and $\{\hat{\mathbf{A}}_{\mathcal{H}_1}^{0}, \ \hat{\mathbf{V}}_{\mathcal{H}_1}^{0}\}$, for which we derive appropriate solutions in the following. The corresponding global algorithm is summed up in the box Algorithm 2.

1) Solving the MLE for $\{\tau_i\}_{i\in[\![1,I]\!]}$, $\mathbf{\Lambda}_{\mathcal{H}_1}^{\star}$, and $\mathbf{V}_{\mathcal{H}_1}^{\star}$ under $\mathcal{H}_1$: This problem requires solving:

$$\begin{array}{ll} \underset{\{\tau_i\}, \mathbf{\Lambda}_{\mathcal{H}_1}^{\star}, \mathbf{V}_{\mathcal{H}_1}^{\star}}{\text{minimize}} & \sum_{i=1}^{I} \left[K_i \ln(|\mathbf{\Sigma}_i|) + \text{Tr}\left\{\mathbf{S}_i \mathbf{\Sigma}_{\boldsymbol{i}}^{-1}\right\}\right] \\ \text{subject to} & \mathbf{\Sigma}_i = \tau_i \mathbf{V}_{\mathcal{H}_1}^{\star} \mathbf{\Lambda}_{\mathcal{H}_1}^{\star} (\mathbf{V}_{\mathcal{H}_1}^{\star})^H + \sigma^2 \mathbf{I}, \ \forall i \in [\![1, I]\!] \\ & \tau_i \geq 0, \ \forall i \in [\![1, I]\!] \\ & \left[\mathbf{\Lambda}_{\mathcal{H}_1}^{\star}\right]_{r,r} \geq 0, \ \forall r \in [\![1, R]\!] \\ & (\mathbf{V}_{\mathcal{H}_1}^{\star})^H \mathbf{V}_{\mathcal{H}_1}^{\star} = \mathbf{I}, \end{array} \tag{26}$$

which is identical to (16) except that the set $i = 0$ is excluded. Hence, we can directly apply Algorithm 1 to obtain the solutions $\{\hat{\tau}_i\}, \hat{\mathbf{\Lambda}}_{\mathcal{H}_1}^{\star}, \hat{\mathbf{V}}_{\mathcal{H}_1}^{\star}$.

2) Solving the MLE for $\hat{\mathbf{A}}_{\mathcal{H}_1}^{0}$ and $\hat{\mathbf{V}}_{\mathcal{H}_1}^{0}$ under $\mathcal{H}_0$: This problem requires solving

$$\begin{array}{ll} \underset{\mathbf{A}_{\mathcal{H}_1}^{0}, \mathbf{V}_{\mathcal{H}_1}^{0}}{\text{minimize}} & K_0 \ln(|\mathbf{\Sigma}_0|) + \text{Tr}\left\{\mathbf{S}_0 \mathbf{\Sigma}_{\mathbf{0}}^{-1}\right\} \\ \text{subject to} & \mathbf{\Sigma}_0 = \mathbf{V}_{\mathcal{H}_1}^{0} \mathbf{A}_{\mathcal{H}_1}^{0} (\mathbf{V}_{\mathcal{H}_1}^{0})^H + \sigma^2 \mathbf{I} \\ & \left[\mathbf{A}_{\mathcal{H}_1}^{0}\right]_{r,r} \geq 0, \ \forall r \in [\![1, R]\!] \\ & (\mathbf{V}_{\mathcal{H}_1}^{0})^H \mathbf{V}_{\mathcal{H}_1}^{0} = \mathbf{I} \end{array} \tag{27}$$

The solution corresponds to the MLE of LR structured CM in the context of Gaussian data. Denote the EVD of the SCM as follows:

$$\mathbf{S}_0/K_0 \overset{\text{EVD}}{=} \left[\ \mathbf{U}_R\ \mathbf{U}_R^{\perp}\ \right] \begin{bmatrix} \mathbf{D}_R & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{M-R} \end{bmatrix} \left[\ \mathbf{U}_R\ \mathbf{U}_R^{\perp}\ \right]^H \tag{28}$$

The solution is given as (cf. section IV.A. in [1] and [4]).

$$\begin{cases} \left[\hat{\mathbf{A}}^0_{\mathcal{H}_1}\right]_{r,r} = \max([\mathbf{D}_R]_{r,r} - \sigma^2,\ 0),\ \forall r \in [\![1, R]\!] \\ \hat{\mathbf{V}}^0_{\mathcal{H}_1} = \mathbf{U}_R \end{cases} \tag{29}$$

Appendix A
Proof of proposition 1

The first part of the proof follows the lines of Proposition 1 in [3]. First, we construct an inequality by the first order Taylor expansion of $\ln$ as:

$$-\sum_{k=1}^{K_i} \frac{\tau_i\ \lambda_r\ s^r_{k,i}}{\tau_i \lambda_r + \sigma^2} \leq -\sum_{k=1}^{K_i} s^r_{k,i} \frac{\tau_i^{(t)} \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2} \left[1 + \ln(\tau_i \lambda_r) - \ln(\tau_i \lambda_r + \sigma^2)\right] + \text{const.} \tag{30}$$

The cost function $\mathcal{L}_i$ is therefore majorized by a first surrogate as:

$$\mathcal{L}_i(\tau_i | \tau_i^{(t)}) \leq \sum_{r=1}^{R} \left[\left(K_i + \sum_{k=1}^{K_i} s^r_{k,i} \frac{\tau_i^{(t)} \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2} \right) \ln\left(\tau_i \lambda_r + \sigma^2\right) - \left(\sum_{k=1}^{K_i} s^r_{k,i} \frac{\tau_i^{(t)} \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2} \right) \ln(\tau_i) \right] \tag{31}$$

that reads

$$\mathcal{L}_i(\tau_i) \leq \sum_{r=1}^{R} \left[\gamma_i^r\ \ \ln\left(\tau_i \lambda_r + \sigma^2\right)\ -\ \beta_i^r\ \ \ln(\tau_i) \right] \tag{32}$$

with γ_i^r and β_i^r in (5). Second, thanks to the Jensen's inequality, we have:

$$\sum_{r=1}^{R} \gamma_i^r\ \ \ln\left(\tau_i \lambda_r + \sigma^2\right) \leq \left(\sum_{r=1}^{R} \gamma_i^r \right) \ln \left(\frac{\sum_{r=1}^{R} \gamma_i^R \frac{\tau_i \lambda_r + \sigma^2}{\tau_i^{(t)} \lambda_r + \sigma^2}}{\sum_{r=1}^{R} \gamma_i^r} \right) + \text{const.} \tag{33}$$

that splits into

$$\sum_{r=1}^{R} \gamma_i^r\ \ \ln\left(\tau_i \lambda_r + \sigma^2\right) \leq \left(\sum_{r=1}^{R} \gamma_i^r \right) \ln \left(\frac{\sum_{r=1}^{R} \frac{\gamma_i^r \lambda_r}{\tau_i^{(t)} \lambda_r + \sigma^2}}{\sum_{r=1}^{R} \gamma_i^r} \tau_i + \frac{\sum_{r=1}^{R} \frac{\gamma_i^r \sigma_n^2}{\tau_i^{(t)} \lambda_r + \sigma^2}}{\sum_{r=1}^{R} \gamma_i^r} \right) + \text{const.} \tag{34}$$

Finally, we have the second surrogate function

$$\mathcal{L}_i(\tau_i) \leq A_i \ln\left(B_i \tau_i + C_i\right) - D_i \ln\left(\tau_i\right) \tag{35}$$

with A_i, B_i, C_i and D_i in (5). The proof concludes by directly applying Proposition 2 of [3], that states that the above majorizer is quasiconvex and has a unique minimizer given in (6).

Appendix B
Proof of proposition 2

The function f in (7) is concave so it can be majorized by its first order Taylor expansion, which is the surrogate given in (8). Minimizing (8) under unitary constraints is equivalent to solve

$$\begin{array}{ll} \underset{\mathbf{V}}{\text{minimize}} & \left\| \left(\sum_{i=1}^{I} \mathbf{A}_i \mathbf{V}^{(t)} \mathbf{B}_i \right) - \mathbf{V} \right\|_F^2 \\ \text{subject to} & \mathbf{V}^H \mathbf{V} = \mathbf{I} \end{array} \tag{36}$$

which is an orthogonal Procrustes problem [3] that has a unique solution given in (9).

References

[1] R. Ben Abdallah, A. Mian, A. Breloy, A. Taylor, M.N. El Korso, and D. Lautru, "Detection methods based on structured covariance matrices for multivariate SAR images processing," .

[2] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, Feb 2017.

[3] Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal, and G. Ginolhac, "Low-complexity algorithms for low rank clutter parameters estimation in radar systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 1986–1998, April 2016.

[4] B. Kang, V. Monga, and M. Rangaswamy, "Rank-constrained maximum likelihood estimation of structured covariance matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 1, pp. 501–515, January 2014.