Low-rank EM-based imaging for large-scale switched interferometric arrays

Jianhua Wang*, Mohammed Nabil El Korso‡, Lucien Bacharach*, Pascal Larzabal*

***Abstract*—Interferences and computational cost pose significant challenges in large-scale interferometric sensing, impacting the accuracy and numerical efficiency of imaging algorithms. In this paper, we introduce an imaging algorithm using antenna switching based on expectation-maximization (EM) to address both challenges. By leveraging the low-rank noise model, our approach effectively captures interferences in interferometric data. Additionally, the proposed switching strategy between different sub-arrays reduces significantly the computational complexity during image restoration. Through extensive experiments on simulated datasets, we demonstrate the superiority of the low-rank noise model over the Gaussian noise model in the presence of interferences. Furthermore, we show that the proposed switching approach yields similar imaging performance with fewer antennas compared to the full array configuration, thereby reducing computational complexity, while outperforming non-switching configurations with the same number of antennas.**

***Index Terms*—Antenna array processing, Barankin Bound, Cramér-Rao bound, EM algorithm, interferometric array.**

I. Introduction

The advent of large-scale big-data interferometers [1–3] has seen a dramatic increase in both the size and number of antennas. This expansion aims to enhance the sensitivity and resolution of interferometry. However, the concomitant increase in antenna numbers and data dimensionality poses significant challenges, making the computational imaging process more intricate and computationally demanding [4, 5]. As one of the methods to address this challenge, the concept of antenna sub-arrays selection has been introduced to mitigate the computational complexity inherent to the imaging process. This strategy involves partitioning antennas into sub-arrays and utilizing only a subset of antennas at each time [6, 7].

Various antenna selection techniques based on estimation lower bounds have been proposed to optimize the sub-array selection, including the Cramér-Rao bound (CRB) [8–10] and the Barankin-type bound (BTB) [11]. These techniques leverage the statistical performance of the interferometric imaging estimators. Recognizing the merits of both approaches, we propose a strategy that combines these two methods, instead of favoring one over the other based on the application scenarios. Our proposed strategy involves dynamically switching between sub-arrays over time, capitalizing on the correlation between sub-arrays.

Moreover, a significant challenge in high-sensitivity interferometric imaging is the presence of interferences, such as Radio Frequency Interference (RFI) [12–14]. RFI comprises unwanted signals originating from various sources, such as communication systems or high voltage transmission lines. Traditional methods often rely on identifying and flagging strong outliers in the raw data using techniques like the generalized likelihood ratio test or manual inspection [15]. Various methods to perform spatial filtering on the correlation matrix have been proposed [14, 16]. While effective for strong interferences, such methods struggle with weaker signals, increasingly prevalent with the growing number of sensors in large-scale big-data interferometers.

To address this gap, building upon the low-rank perturbation model's success in antenna array calibration [17], we propose its application in the presence of interference sources, leveraging its inherent low-rank structure to improve the imaging quality compared to traditional additive Gaussian noise model [18, 19]. Furthermore, we introduce an antenna sub-array switching strategy based on the low-rank perturbation model for interferometric imaging. The proposed strategy is formulated as an expectation-maximization (EM) algorithm, exploiting the correlation between antenna sub-arrays [20], leading to estimate an augmented array visibilities vector. Through simulations based on the existing Very Large Array (VLA) radio interferometer, we showcase the effectiveness of the low-rank based model in terms of normalized mean square error (NMSE). Our results demonstrate that the proposed antenna switching approach enhances imaging quality while reducing computational complexity. In summary, our key contributions are as follows:

- We extend the low-rank perturbation model recently used in the literature [17, 21], to address the challenge of interference. Our method significantly enhances imaging quality compared to traditional approaches relying on additive Gaussian noise models.
- Our work introduces a novel antenna sub-array switching strategy. This strategy utilizes an EM algorithm to exploit the correlations between sub-arrays. The switching mechanism reduces the computational cost while preserving high quality image.
- The simulations, based on the VLA telescope configuration, demonstrate the practical benefits of our approach, which is promising for large-scale radio telescopes.

II. Interferometric data model

Interferometry involves measuring the spatial coherence of the electric field between pairs of antennas within the antenna array. These measurements, termed visibilities, provide information about the electromagnetic radiation emitted by celestial sources under observation. The main idea of this

* *SATIE,* ‡ *L2S, Université Paris-Saclay, 91190 Gif-sur-Yvette, France*

paper is to dynamically switch the antenna sub-arrays during the observation process, a formulation expressed as

$$\mathbf{y}_k = \mathbf{M}_k \mathbf{r}_k,\ k = 1, \cdots, K \tag{1}$$

where $\mathbf{M}_k \in \mathbb{R}^{p_k \times N_k}$ represents the switching matrix, constructed by selecting $p_k < N$ rows from the identity matrix $\mathbf{I}_N$. The vector $\mathbf{r}_k \in \mathbb{C}^N$ denotes the complete visibility vector, while $\mathbf{y}_k \in \mathbb{C}^{p_k}$ represents the observed visibilities in the k-th observation block with K denoting the total number of observation blocks.

For each complete visibility vector $\mathbf{r}_k$, the interferometric imaging problem can be formulated as

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{x} + \mathbf{W}_k \boldsymbol{\gamma}_k + \mathbf{n}_k \tag{2}$$

where $\mathbf{H}_k \in \mathbb{C}^{N \times m}$ is the forward operator, corresponding to the spatial Fourier transform, symbol $\mathbf{x} \in \mathbb{R}^m$ represents the image vector of size m (number of image pixels) to be restored. The additional noise is modeled as a combination of thermal noise $\mathbf{n}_k$, which follows a Gaussian distribution $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, and interferences modeled by a low-rank model $\mathbf{W}_k \boldsymbol{\gamma}_k$, where $\boldsymbol{\gamma}_k \sim \mathcal{CN}(\boldsymbol{\ell}, \mathbf{I}_n)$, $\mathbf{W}_k \in \mathbb{C}^{N \times n}$, and $\boldsymbol{\ell} \in \mathbb{C}^n$. Typically, $n \ll N$.

Depending on the application, the form of $\boldsymbol{\ell}$ varies. For instance, in radar [22, 23], $\boldsymbol{\ell} = \mathbf{0}$, while in radio astronomical interferometry [17], $\boldsymbol{\ell} = \mathrm{vec}(\mathbf{I})$. In the following, we assume $\boldsymbol{\ell} = \mathrm{vec}(\mathbf{I})$ and the case $\boldsymbol{\ell} = \mathbf{0}$ can be easily deduced.

Let $\boldsymbol{\theta} = \{\mathbf{x}, \mathbf{W}_k, \sigma^2\}$ denote the set of unknown parameters. The log-likelihood function of the observed visibilities can be expressed as

$$L(\mathbf{y}_1, \cdots, \mathbf{y}_K | \boldsymbol{\theta}) \propto -\sum_{k=1}^{K} \left(\log |\boldsymbol{\Sigma}_k| + \|\mathbf{e}_k\|^2_{\boldsymbol{\Sigma}_k^{-1}} \right) \tag{3}$$

where the weighted ℓ_2-norm is defined as $\|\mathbf{z}\|^2_{\mathbf{A}} = \mathbf{z}^H \mathbf{A} \mathbf{z}$ and

$$\mathbf{e}_k = \mathbf{y}_k - \mathbf{M}_k \mathbf{H}_k \mathbf{x} - \mathbf{M}_k \mathbf{W}_k \,\mathrm{vec}(\mathbf{I}) \tag{4}$$

$$\boldsymbol{\Sigma}_k = \sigma^2 \mathbf{I}_{p_k} + \mathbf{M}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{M}_k^H \tag{5}$$

The resulting regularized maximum likelihood problem for the imaging problem is formulated as

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} L(\mathbf{y}_1, \cdots, \mathbf{y}_K | \boldsymbol{\theta}) + \mathcal{R}(\mathbf{x}) \tag{6}$$

where $\mathcal{R}(\mathbf{x})$ denotes the regularization function that promotes sparsity within the image vector $\mathbf{x}$.

III. EM-BASED IMAGING ALGORITHM

The regularized maximum likelihood problem is quite difficult to tackle especially in the presence of missing data which is due to a switching interferometric array. We thus propose the utilization of an EM algorithm variant, namely, the Expectation-conditional maximization either (ECME) algorithm [24]. The term "either" refers to the fact that the updates of σ^2 and $\mathbf{W}_k$ are conducted by maximizing the conditional expectation of the complete-data log-likelihood function, and the update of $\mathbf{x}$ is performed by maximizing directly the actual (incomplete data) log-likelihood function. In our case, we consider $\{\mathbf{r}_k, \boldsymbol{\gamma}_k\}$ with $k = 1, \cdots, K$ as the complete data space and $\mathbf{y}_k = \mathbf{M}_k \mathbf{r}_k$ represents the observed part of $\mathbf{r}_k$.

A. E-step

The expectation of the log-likelihood function of the complete data given the observed data and the current estimate of the parameters of interest is given by

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) &= \mathbb{E}_{\boldsymbol{\gamma},\mathbf{r}|\mathbf{y},\boldsymbol{\theta}^{(m)}} \left[\log p(\mathbf{y}, \mathbf{r}, \boldsymbol{\gamma}|\boldsymbol{\theta})\right] \\ &= \mathbb{E}_{\mathbf{r}|\mathbf{y},\boldsymbol{\theta}^{(m)}} \left[\mathbb{E}_{\boldsymbol{\gamma}|\mathbf{r},\mathbf{y},\boldsymbol{\theta}^{(m)}} \left[\sum_{k=1}^{K} L(\mathbf{y}_k, \mathbf{r}_k, \boldsymbol{\gamma}_k|\boldsymbol{\theta}) \right] \right] \end{aligned} \tag{7}$$

where the vector $\mathbf{r}$ comprises all $\mathbf{r}_k$, $k = 1, 2, \ldots, K$, similarly for $\mathbf{y}$. After applying the inner expectation with respect to $\boldsymbol{\gamma}_k$, the surrogate function reads

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) \propto &- \sum_{k=1}^{K} \left(N_k \log(\sigma^2) + \frac{\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} \left[\|\mathbf{r}_k - \mathbf{H}_k \mathbf{x}\|^2 \right]}{\sigma^2} \right) \\ &- \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathrm{Tr} \left(\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} \left[\mathbf{W}_k \mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} \left[\boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^H \right] \mathbf{W}_k^H \right] \right) \\ &+ \frac{2}{\sigma^2} \sum_{k=1}^{K} \mathrm{Tr} \left(\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} \left[(\mathbf{r}_k - \mathbf{H}_k \mathbf{x}) (\mathbf{W}_k \mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} [\boldsymbol{\gamma}_k])^H \right] \right) \end{aligned} \tag{8}$$

where

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} \left[\boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^H \right] &= \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}} \\ &+ \mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} [\boldsymbol{\gamma}_k] \mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} [\boldsymbol{\gamma}_k]^H \end{aligned} \tag{9}$$

$$\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k,\mathbf{y}_k,\boldsymbol{\theta}^{(m)}} [\boldsymbol{\gamma}_k] = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}} \,\mathrm{vec}(\mathbf{I}) + \frac{1}{\sigma^{2(m)}} \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}} \mathbf{W}_k^{(m)H} (\mathbf{r}_k - \mathbf{H}_k \mathbf{x}) \tag{10}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}} = \left(\mathbf{I}_n + \frac{1}{\sigma^{2(m)}} \mathbf{W}_k^{(m)H} \mathbf{W}_k^{(m)} \right)^{-1} \tag{11}$$

To complete the calculation of the Q function, we now need to handle the outer expectation with respect to $\mathbf{r}_k$, which is given by the following Proposition.

Proposition 1. *Denoting $Q_1(\sigma^2|\boldsymbol{\theta}^{(m)})$ the part of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ relative to σ^2, and $Q_2(\mathbf{W}_k|\boldsymbol{\theta}^{(m)})$ the part of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ relative to $\mathbf{W}_k$, the E-step reads*

$$\begin{aligned} Q_1(\sigma^2|\boldsymbol{\theta}^{(m)}) \propto &- \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathrm{Tr} \left(\mathbf{W}_k^{(m)} \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}}^{(m)} \mathbf{W}_k^{(m)H} \right) \\ &- \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathrm{Tr} \left(\boldsymbol{\Psi}_k \left(\boldsymbol{\xi}_k \boldsymbol{\xi}_k^H + \boldsymbol{\Omega}_k \right) \boldsymbol{\Psi}_k^H \right) \\ &- \sum_{k=1}^{K} N \log(\sigma^2) \end{aligned} \tag{12}$$

in which,

$$\boldsymbol{\Psi}_k = \left(\mathbf{I} - \frac{1}{\sigma^{2(m)}} \mathbf{W}_k^{(m)} \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}} (\mathbf{W}_k^{(m)})^H \right) \tag{13}$$

$$\boldsymbol{\xi}_k = \boldsymbol{\Sigma}_{\mathbf{r}}^{(m)} \mathbf{M}_k^H \boldsymbol{\Sigma}_k^{-1} \mathbf{e}_k^{(m)} \tag{14}$$

$$\boldsymbol{\Omega}_k = \boldsymbol{\Sigma}_{\mathbf{r}}^{(m)} - \boldsymbol{\Sigma}_{\mathbf{r}}^{(m)} \mathbf{M}_k^H \boldsymbol{\Sigma}_k^{-1} \mathbf{M}_k \boldsymbol{\Sigma}_{\mathbf{r}}^{(m)} \tag{15}$$

and

$$Q_2(\mathbf{W}_k|\boldsymbol{\theta}^{(m)}) \propto \frac{1}{\sigma^2} \mathrm{Tr} \left(2 \boldsymbol{\zeta}_k \mathbf{W}_k^H - \mathbf{W}_k \widetilde{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)} \mathbf{W}_k^H \right) \tag{16}$$

where

$$\begin{aligned}\boldsymbol{\zeta}_k = &\left(\mathbf{W}_k^{(m)}\,\mathrm{vec}(\mathbf{I}) + \boldsymbol{\xi}_k\right)(\widetilde{\boldsymbol{\gamma}}_k^{(m)})^H + \mathbf{W}_k^{(m)} \\ &- \boldsymbol{\Sigma}_{\mathbf{r}}^{(m)}\mathbf{M}_k^H\boldsymbol{\Sigma}_k^{-1}\mathbf{M}_k\mathbf{W}_k^{(m)}\end{aligned} \tag{17}$$

$$\widetilde{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)} = \boldsymbol{\Sigma}_{\mathbf{M}} + \widetilde{\boldsymbol{\gamma}}_k^{(m)}\left(\widetilde{\boldsymbol{\gamma}}_k^{(m)}\right)^H \tag{18}$$

$$\widetilde{\boldsymbol{\gamma}}_k^{(m)} = \mathrm{vec}(\mathbf{I}) + (\mathbf{M}_k\mathbf{W}_k^{(m)})^H\boldsymbol{\Sigma}_k^{-1}\mathbf{e}_k \tag{19}$$

$$\boldsymbol{\Sigma}_{\mathbf{M}} = \left(\mathbf{I}_n + \frac{1}{\sigma^{2(m)}}\mathbf{W}_k^{(m)H}\mathbf{M}_k^H\mathbf{M}_k\mathbf{W}_k^{(m)}\right)^{-1} \tag{20}$$

The proof can be found in the supplementary material.

B. M-step

Updating σ^2 and $\mathbf{W}_k$ involves respectively

$$\begin{aligned}(\sigma^2)^{(m+1)} &= \arg\max_{\sigma^2} Q_1(\sigma^2|\boldsymbol{\theta}^{(m)}) \\ &= \frac{1}{\sum_{k=1}^K N_k}\sum_{k=1}^K \mathrm{Tr}\left(\boldsymbol{\Psi}_k^H\boldsymbol{\Psi}_k\left(\boldsymbol{\xi}_k\boldsymbol{\xi}_k^H + \boldsymbol{\Omega}_k\right)\right)\end{aligned} \tag{21}$$

$$\mathbf{W}_k^{(m+1)} = \arg\max_{\mathbf{W}_k} Q_2(\mathbf{W}_k|\boldsymbol{\theta}^{(m)}) = \boldsymbol{\zeta}_k\left(\widetilde{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)}\right)^{-1} \tag{22}$$

Updating $\mathbf{x}$ involves maximizing directly the observed data log-likelihood function (3) given the current estimate of $\mathbf{W}_k^{(m+1)}$ and $(\sigma^2)^{(m+1)}$, which is given by

$$\mathbf{x}^{(m+1)} = \arg\min_{\mathbf{x}}\sum_{k=1}^K\left(\left\|\mathbf{e}_k^{(m)}\right\|^2_{\boldsymbol{\Sigma}_k^{-1(m)}}\right) + \mathcal{R}(\mathbf{x}) \tag{23}$$

where $\boldsymbol{\Sigma}_k^{-1(m)}$ denotes the inverse of matrix $\boldsymbol{\Sigma}_k^{(m)}$.

To promote sparsity in the image, a wide variety of techniques are available, including, for example, ℓ_p priors with $0 \le p \le 1$ and Sparse Bayesian Learning (SBL) techniques [25–27]. However, as is usually done successfully in radio astronomy [28, 29], an ℓ_1-norm will be used as the regularization function in the sequel, i.e., $\mathcal{R}(\mathbf{x}) = \alpha||\mathbf{x}||_1$. We employ the Iterative Soft-Thresholding Algorithm (ISTA) [30] to solve this optimization problem. The resulting solution is expressed as

$$\mathbf{x}^{(m+1)} = \underset{\alpha||\mathbf{x}||_1}{\mathrm{prox}}\left(\mathbf{x}^{(m)} + \frac{1}{K}\sum_{k=1}^K(\mathbf{M}_k\mathbf{H}_k)^H\boldsymbol{\Sigma}_k^{-1(m)}\mathbf{e}_k^{(m)}\right) \tag{24}$$

Here, the proximal operator is defined as

$$\underset{\alpha||\mathbf{x}||_1}{\mathrm{prox}}\,(\mathbf{u})_i = \mathrm{sgn}(\mathbf{u}_i)\left(|\mathbf{u}_i| - \alpha\right)_+ \tag{25}$$

with $(|\mathbf{u}_i| - \alpha)_+ = \max(0, |\mathbf{u}_i| - \alpha)$. This procedure efficiently updates $\mathbf{x}$ in an iterative manner, incorporating sparsity-inducing regularization with parameter α.

The low-rank ECME-based array switching imaging algorithm is summarized in Algorithm 1.

Algorithm 1: ECME-based antenna switching imager

Input: $\mathbf{H}_k$, α, and $\mathbf{y}_k$, $k = 1, \ldots, K$
Output: Estimates of $\widehat{\mathbf{x}}$, $\widehat{\mathbf{W}_k}$, and $\widehat{\sigma^2}$
Initialisation: $\mathbf{x}^{(m)} \leftarrow \mathbf{x}^{(0)}$, $\mathbf{W}_k^{(m)} \leftarrow \mathbf{W}_k^{(0)}$, and $(\sigma^2)^{(m)} \leftarrow (\sigma^2)^{(0)}$

```
while stop criterion not met do
    E-step: for k = 1, ..., K do
        calculate Ψ_k from (13)
        compute ζ_k from (17)
        compute γγ^H_k^(m) from (18)
    M-step: for k = 1, ..., K do
        update W_k^(m+1) using (22)
    obtain (σ^2)^(m+1) from (21)
    update x^(m+1) using (24)
```

IV. Numerical Evaluation

In this section, we employ the proposed algorithm and validate its effectiveness by conducting simulations in the field of radio astronomical imaging. The simulations are based on the existing radio-astronomical telescope Very Large Array (VLA) [31], which is composed of three branches with 9 antennas on each branch, resulting into a total of 27 antennas. The M31 galaxy serves as the true image, which is based on the H2 region in the M31 galaxy [32] and has been used for investigating radio-astronomy imaging algorithms [28].

The true image is discretized into 64×64 pixels, and then vectorized, resulting in $\mathbf{x} \in \mathbb{R}^{4096}$ in the system model (2). Then the true visibility is calculated from the discretized true image $\mathbf{x}$. To simulate thermal noise, additive white Gaussian noise with zero mean and a variance of σ^2 is incorporated into the visibilities. To simulate RFI noise, we model 100 RFI events as additional point sources with same power into the scene. Only 20% of the visibilities are affected by the RFI, reflecting the selective impact on sensors. The power of each RFI event is quantified relatively to the source power, as defined by $P_{\text{RFIdB}} = 10\log_{10}(P_{\text{RFI}}/P_0)$, where $P_0 = \|\mathbf{x}\|^2$ is the source power. The number of observation blocks K is set to 22. The performance of different algorithms is assessed using the structural similarity index (SSIM) [33] and the normalized mean square error (NMSE) defined as $\|\widehat{\mathbf{x}} - \mathbf{x}\|^2/\|\mathbf{x}\|^2$, where $\widehat{\mathbf{x}}$ denotes the reconstructed image.

In the following part of simulation, firstly, we demonstrate that the low-rank imaging algorithm without switching strategy outperforms the Gaussian-based imaging algorithm in the presence of RFI noise. Then, we evaluate the imaging performance of the proposed antenna switching algorithm, and we compare it with non-switching procedures.

Due to space constraints, additional simulation results can be accessed in our open-source folder [34]. These include simulations concerning imaging restoration with varying levels of low-rank noise and RFI noise power.

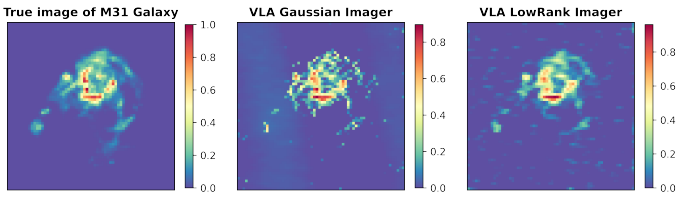

Fig. 1: Reconstructed images of the three methods with RFI noise power -1 dB.

A. Simulation 1: Low-rank based imager versus Gaussian based imager with RFI noise

Figure 1 showcases the reconstructed images under -1 dB RFI noise. The results indicate that the imaging algorithm based on the low-rank noise model consistently outperforms that based on a standard Gaussian noise assumption in terms of NMSE, with reconstructed images more closely resembling the true image.

B. Simulation 2: Low-rank based switching imager

To alleviate the computational burden of the imaging algorithm, we introduce a switching strategy to alternate between different sub-arrays. To avoid poor selections and/or potential ambiguities in sub-array selection arising from random switching, we advocate for a method that toggles between two types of sub-arrays based on the minimization of mean square error (MSE) lower bounds, namely, the Cramér-Rao Bound (CRB) and Barankin-Type Bound (BTB), building upon our previous work [7]. The CRB-based antenna selection method yields a beampattern with a thin mainlobe, while the BTB-based method provides a relevant trade-off between mainlobe width and sidelobe level [35, 36]. The switching strategy aims to leverage the benefits of both methods.

The sub-arrays are derived from the full VLA configuration comprising 27 antennas. These antenna sub-arrays (AS), designated as AS(CRB) and AS(BTB) based on the CRB and BTB criteria respectively, along with the proposed switching strategy, are visually represented in Fig. 3. Detailed information regarding the number of antennas and visibilities for each method is provided in Table I. Computational efficiency was also assessed. It can be shown that the gain of the computational complexity (for a large N) is β^2 in which $\beta = p_k/N$ is ratio of visibility selection defined as number of selected visibilities divided by full visibilities. Furthermore, the performance evaluation of the proposed switching strategy, in conjunction with both CRB-based and BTB-based sub-arrays, as well as the full array VLA, is depicted in Fig. 2, illustrating the NMSE of the reconstructed images against the power of the RFI noise with 100 Monte Carlo simulations. The reconstructed images under 1 dB RFI noise power are showcased in Fig. 3.

The proposed switching strategy outperforms both CRB- and BTB-based sub-arrays in terms of NMSE. Additionally, the reconstructed image using the proposed switching strategy closely resembles the result obtained with the full array VLA. In summary, the proposed switching strategy maintains

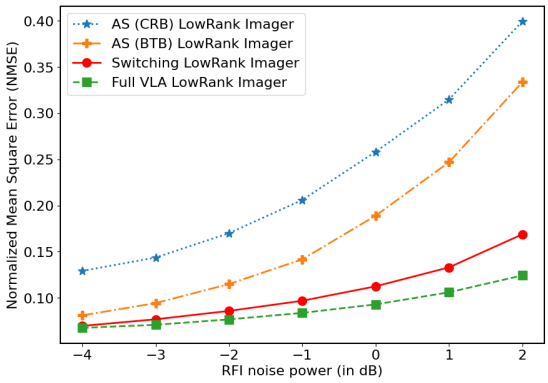

Fig. 2: Evolution of NMSE for each method with varying RFI noise power.

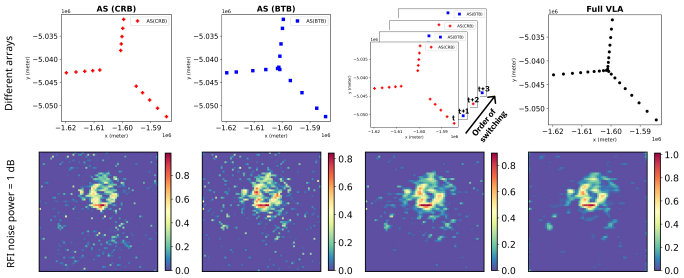

Fig. 3: Reconstructed image of the four methods with RFI noise power 1 dB.

Table I: Number of antennas/visibilities required by each method, NMSE and SSIM of each method for RFI 1dB.

		Ant. No.	Vis. No.	NMSE	SSIM
Non-switching	Array-Full	27	7722	0.1060	0.8121
	Array-CRB	15	2310	0.3149	0.5172
	Array-BTB	15	2310	0.2469	0.5356
Switching method		**15**	**2310**	**0.1330**	**0.7303**

imaging algorithm performance with a reduced number of antennas and visibilities compared to the full array, thereby reducing computational burden.

V. Conclusions

This paper presents an application of the EM algorithm to the radio interferometric imaging problem, leveraging a low-rank noise model to account for interference effects. Additionally, we propose a switching strategy based on pre-selected sub-arrays to reduce computational requirements for image restoration. The simulations show that the low-rank noise model outperforms the Gaussian noise model in the presence of RFI and the proposed switching strategy yields comparable imaging performance with fewer antennas and visibilities compared to the full array. By leveraging low-rank noise modeling and a novel switching strategy, we enhance the robustness and efficiency of imaging algorithms, paving the way for low-cost and high-quality imaging in large-scale radio telescopes.

REFERENCES

[1] A. Schutz, A. Ferrari, D. Mary, É. Thiébaut, and F. Soulez, "Large scale 3D image reconstruction in optical interferometry," in *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 474–478.

[2] S. Vijay Kartik, R. E. Carrillo, J.-P. Thiran, and Y. Wiaux, "A Fourier dimensionality reduction model for big data interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 468, no. 2, pp. 2382–2400, 2017.

[3] D. H. T. Minh and Y.-N. Ngo, "ComSAR: A new algorithm for processing Big Data SAR Interferometry," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 820–823.

[4] X. Wang, Y. Guo, F. Wen, J. He, and T.-K. Truong, "EMVS-MIMO radar with sparse Rx geometry: Tensor modeling and 2D direction finding," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[5] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2020.

[6] M. Juhlin and A. Jakobsson, "Optimal sensor placement for localizing structured signal sources," *Signal Processing*, p. 108679, 2022.

[7] J. Wang, L. Bacharach, P. Larzabal, and M. N. El Korso, "A comparison of antenna placement criteria based on the cramér-rao and barankin bounds for radio interferometer arrays," *Signal Processing*, p. 109404, 2024.

[8] H. Gazzah and J. P. Delmas, "CRB-based design of linear antenna arrays for near-field source localization," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 4, pp. 1965–1974, 2014.

[9] J. P. Delmas, M. N. El Korso, H. Gazzah, and M. Castella, "CRB analysis of planar antenna arrays for optimizing near-field source localization," *Signal Processing*, vol. 127, pp. 117–134, 2016.

[10] M. N. El Korso, A. Renaux, R. Boyer, and S. Marcos, "Deterministic performance bounds on the mean square error for near field source localization," *IEEE transactions on signal processing*, vol. 61, no. 4, pp. 871–877, 2012.

[11] J. Tabrikian, O. Isaacs, and I. Bilik, "Cognitive antenna selection for automotive radar using bobrovsky-zakai bound," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 892–903, 2021.

[12] W. A. Baan, "Implementing rfi mitigation in radio science," *Journal of Astronomical Instrumentation*, vol. 8, no. 01, p. 1940010, 2019.

[13] S. Wei, H. Zhang, X. Zeng, Z. Zhou, J. Shi, and X. Zhang, "CARNet: An effective method for SAR image interference suppression," *International Journal of Applied Earth Observation and Geoinformation*, vol. 114, p. 103019, 2022.

[14] G. Sharma and A. A. B. Raj, "Low complexity interference mitigation technique in IRCI-free SAR imaging algorithms," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[15] A. Offringa, A. De Bruyn, S. Zaroubi, and M. Biehl, "A LOFAR RFI detection pipeline and its first results," *arXiv preprint arXiv:1007.2089*, 2010.

[16] A. Leshem, A.-J. van der Veen, and A.-J. Boonstra, "Multichannel interference mitigation techniques in radio astronomy," *The Astrophysical Journal Supplement Series*, vol. 131, no. 1, p. 355, 2000.

[17] Y. Mhiri, M. N. El Korso, A. Breloy, and P. Larzabal, "Multifrequency array calibration in presence of radio frequency interferences," *Signal Processing*, vol. 199, p. 108613, 2022.

[18] C. Tasse, "Nonlinear Kalman filters for calibration in radio interferometry," *Astronomy & Astrophysics*, vol. 566, p. A127, 2014.

[19] S. Salvini and S. J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *Astronomy & Astrophysics*, vol. 571, p. A97, 2014.

[20] A. Aubry, A. De Maio, S. Marano, and M. Rosamilia, "Structured covariance matrix estimation with missing-(complex) data for radar applications via expectation-maximization," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5920–5934, 2021.

[21] M. An, J. Yin, T. Liu, Z. Wu, and Y. Li, "Radio Frequency Interference Mitigation Based on Low-Rank Sparse Decomposition for Polarimetric Weather Radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–19, 2024.

[22] J. Fang, Z. Xu, B. Zhang, W. Hong, and Y. Wu, "Fast compressed sensing SAR imaging based on approximated observation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 352–363, 2013.

[23] X. Dong and Y. Zhang, "A novel compressive sensing algorithm for SAR imaging," *IEEE Journal of selected topics in applied earth observations and remote sensing*, vol. 7, no. 2, pp. 708–720, 2013.

[24] C. Liu and D. B. Rubin, "The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence," *Biometrika*, vol. 81, no. 4, pp. 633–648, 1994.

[25] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[26] J. Palmer, B. Rao, and D. Wipf, "Perspectives on sparse Bayesian learning," *Advances in neural information processing systems*, vol. 16, 2003.

[27] R. Giri and B. Rao, "Type I and type II Bayesian methods for sparse signal recovery using scale mixtures," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3418–3428, 2016.

[28] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "PURIFY: a new approach to radio-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 439, no. 4, pp. 3591–3604, 2014.

[29] X. Cai, M. Pereyra, and J. D. McEwen, "Uncertainty quantification for radio interferometric imaging–I. Proximal MCMC methods," *Monthly Notices of the Royal Astronomical Society*, vol. 480, no. 3, pp. 4154–4169, 2018.

[30] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[31] P. J. Napier, A. R. Thompson, and R. D. Ekers, "The very large array: Design and performance of a modern synthesis radio telescope," *Proceedings of the IEEE*, vol. 71, no. 11, pp. 1295–1320, 1983.

[32] A. Dabbech, C. Ferrari, D. Mary, E. Slezak, O. Smirnov, and J. S. Kenyon, "Moresane: Model reconstruction by synthesis-analysis estimators-a sparse deconvolution algorithm for radio interferometric imaging," *Astronomy & Astrophysics*, vol. 576, p. A7, 2015.

[33] R. Dosselmann and X. D. Yang, "A comprehensive assessment of the structural similarity index," *Signal, Image and Video Processing*, vol. 5, pp. 81–91, 2011.

[34] https://github.com/jianhua-WANG-ENS/Low-rank-EM-switching.

[35] R. McAulay and E. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, no. 6, pp. 669–676, 1971.

[36] M. Morelande and B. Ristic, "Signal-to-noise ratio threshold effect in track before detect," *IET radar, sonar & navigation*, vol. 3, no. 6, pp. 601–608, 2009.

SUPPLEMENTARY MATERIAL

A. Proof of Proposition 1

Proof. We will treat each parameter individually.

a) Determining σ^2: Based on the surrogate function (8), the log-likelihood function of the complete data after the inner expectation with respect to $\boldsymbol{\gamma}_k$ can be reformulated as

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k}\left[L(\sigma^2|\mathbf{y}_k,\mathbf{r}_k,\boldsymbol{\gamma}_k)\right] &\propto -\frac{1}{\sigma^2}\operatorname{Tr}\left(\mathbf{W}_k\boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}}^{(m)}\mathbf{W}_k^H\right)\\ &-\frac{1}{\sigma^2}||\mathbf{r}_k-\mathbf{H}_k\mathbf{x}-\mathbf{W}_k\widehat{\boldsymbol{\gamma}}_k^{(m)}||^2-N_k\log(\sigma^2)\end{aligned} \tag{26}$$

where

$$\widehat{\boldsymbol{\gamma}}_k^{(m)} = \operatorname{vec}(\mathbf{I}) + \mathbf{W}_k^H\boldsymbol{\Sigma}_{\mathbf{r}}^{-1}\boldsymbol{\eta}_k \tag{27}$$

Then, inserting $\widehat{\boldsymbol{\gamma}}_k$ into (26) and after some algebra, we obtain

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k}\left[L(\sigma^2|\mathbf{y}_k,\mathbf{r}_k,\boldsymbol{\gamma}_k)\right] &\propto -\frac{1}{\sigma^2}\operatorname{Tr}\left(\mathbf{W}_k\boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}}\mathbf{W}_k^H\right)\\ &-\frac{1}{\sigma^2}\operatorname{Tr}\left(\boldsymbol{\Psi}_k\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H\boldsymbol{\Psi}_k^H\right)-N_k\log(\sigma^2)\end{aligned} \tag{28}$$

Then taking the outer expectation with respect to $\mathbf{r}_k$, the surrogate function depending only on σ^2 becomes

$$Q_1(\sigma^2|\boldsymbol{\theta}^{(m)}) = \sum_{k=1}^{K}\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}}\left[\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k}\left[L(\sigma^2|\mathbf{y}_k,\mathbf{r}_k,\boldsymbol{\gamma}_k)\right]\right] \tag{29}$$

From this, we derive (12). For the details regarding the computation of

$$\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}}\left[\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H\right] = \boldsymbol{\xi}_k\boldsymbol{\xi}_k^H + \boldsymbol{\Omega}_k \tag{30}$$

please refer to *Part B: Conditional distribution of* $\mathbf{r}_k$ *given* $\mathbf{y}_k$.

b) Determining $\mathbf{W}_k$: By denoting $\boldsymbol{\Upsilon} = \mathbf{r}_k - \mathbf{H}_k\mathbf{x}$, the log-likelihood function of the complete data after the inner expectation with respect to $\boldsymbol{\gamma}_k$ can be reformulated as

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k}\left[L(\mathbf{W}_k|\mathbf{y}_k,\mathbf{r}_k,\boldsymbol{\gamma}_k)\right] &\propto -\frac{1}{\sigma^2}\operatorname{Tr}(\mathbf{W}_k\widehat{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)}\mathbf{W}_k^H)\\ &-\frac{1}{\sigma^2}\operatorname{Tr}\left(\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^H-\boldsymbol{\Upsilon}(\mathbf{W}_k\widehat{\boldsymbol{\gamma}}_k^{(m)})^H-\mathbf{W}_k\widehat{\boldsymbol{\gamma}}_k^{(m)}\boldsymbol{\Upsilon}^H\right)\end{aligned} \tag{31}$$

where

$$\widehat{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}|\mathbf{r}}^{(m)} + \widehat{\boldsymbol{\gamma}}_k^{(m)}\left(\widehat{\boldsymbol{\gamma}}_k^{(m)}\right)^H \tag{32}$$

Then, inserting $\widehat{\boldsymbol{\gamma}}_k$ and $\widehat{\boldsymbol{\gamma}\boldsymbol{\gamma}^H}_k^{(m)}$ into (31) and taking the outer expectation with respect to $\mathbf{r}_k$, the surrogate function depending only on $\mathbf{W}_k$ becomes

$$Q_2(\mathbf{W}_k|\boldsymbol{\theta}^{(m)}) = \sum_{k=1}^{K}\mathbb{E}_{\mathbf{r}_k|\mathbf{y}_k,\boldsymbol{\theta}^{(m)}}\left[\mathbb{E}_{\boldsymbol{\gamma}_k|\mathbf{r}_k}\left[L(\mathbf{W}_k|\mathbf{y}_k,\mathbf{r}_k,\boldsymbol{\gamma}_k)\right]\right] \tag{33}$$

From this, after some algebra, we derive (16).

That concludes the proof. □

B. Conditional distribution of $\mathbf{r}_k$ given $\mathbf{y}_k$

Given that $\mathbf{r}_k \sim \mathcal{CN}(\mathbf{v},\boldsymbol{\Sigma})$ with

$$\mathbf{v} = \mathbf{H}_k\mathbf{x} + \mathbf{W}_k\operatorname{vec}(\mathbf{I}) \tag{34}$$

$$\boldsymbol{\Sigma} = \sigma^2\mathbf{I} + \mathbf{W}_k\mathbf{W}_k^H \tag{35}$$

we can deduce

$$\begin{aligned}\mathbf{M}_k\mathbf{r}_k &\sim \mathcal{CN}(\mathbf{M}_k\mathbf{v},\mathbf{M}_k\boldsymbol{\Sigma}\mathbf{M}_k^H)\\ \overline{\mathbf{M}}_k\mathbf{r}_k &\sim \mathcal{CN}(\overline{\mathbf{M}}_k\mathbf{v},\overline{\mathbf{M}}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H)\end{aligned} \tag{36}$$

Moreover, the covariance between $\mathbf{M}_k\mathbf{r}_k$ and $\overline{\mathbf{M}}_k\mathbf{r}_k$ is

$$\operatorname{cov}(\mathbf{M}_k\mathbf{r}_k,\overline{\mathbf{M}}_k\mathbf{r}_k) = \mathbf{M}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H \tag{37}$$

Thus, the joint distribution of $\mathbf{y}_k$ and $\overline{\mathbf{M}}_k\mathbf{r}_k$ is given by

$$\begin{pmatrix}\mathbf{y}_k\\ \overline{\mathbf{M}}_k\mathbf{r}_k\end{pmatrix} \sim \mathcal{CN}\left(\begin{bmatrix}\mathbf{M}_k\mathbf{v}\\ \overline{\mathbf{M}}_k\mathbf{v}\end{bmatrix},\begin{bmatrix}\mathbf{M}_k\boldsymbol{\Sigma}\mathbf{M}_k^H & \mathbf{M}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H\\ \overline{\mathbf{M}}_k\boldsymbol{\Sigma}\mathbf{M}_k^H & \overline{\mathbf{M}}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H\end{bmatrix}\right) \tag{38}$$

where $\mathbf{y}_k = \mathbf{M}_k\mathbf{r}_k$.

Consequently, the conditional distribution of $\overline{\mathbf{M}}_k\mathbf{r}_k$ given $\mathbf{y}_k$ is

$$\overline{\mathbf{M}}_k\mathbf{r}_k|\mathbf{y}_k \sim \mathcal{CN}(\overline{\mathbf{M}}_k\mathbf{v}+\boldsymbol{\mu}_k,\mathbf{G}_k) \tag{39}$$

where

$$\begin{aligned}\boldsymbol{\mu}_k &= \overline{\mathbf{M}}_k\boldsymbol{\Sigma}\mathbf{M}_k^H\left(\mathbf{M}_k\boldsymbol{\Sigma}\mathbf{M}_k^H\right)^{-1}\mathbf{e}_k\\ \mathbf{G}_k &= \overline{\mathbf{M}}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H - \overline{\mathbf{M}}_k\boldsymbol{\Sigma}\mathbf{M}_k^H\left(\mathbf{M}_k\boldsymbol{\Sigma}\mathbf{M}_k^H\right)^{-1}\mathbf{M}_k\boldsymbol{\Sigma}\overline{\mathbf{M}}_k^H\end{aligned} \tag{40}$$

Consider the matrix $\mathbf{B}_k$ defined as

$$\mathbf{B}_k = \begin{bmatrix}\mathbf{M}_k\\ \overline{\mathbf{M}}_k\end{bmatrix} = \begin{bmatrix}\mathbf{M}_k^T & \overline{\mathbf{M}}_k^T\end{bmatrix}^T \tag{41}$$

which satisfies

$$\mathbf{B}_k^T\mathbf{B}_k = \mathbf{M}_k^T\mathbf{M}_k + \overline{\mathbf{M}}_k^T\overline{\mathbf{M}}_k = \mathbf{I}_{N_k} \tag{42}$$

and then denoting $\boldsymbol{\epsilon}_k = \mathbf{r}_k - \mathbf{v}$, we have

$$\begin{aligned}&\mathbb{E}_{\overline{\mathbf{M}}_k\mathbf{r}_k|\mathbf{y}_k}\left[\mathbf{B}_k\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^H\mathbf{B}_k^T\right]\\ &= \mathbb{E}_{\overline{\mathbf{M}}_k\mathbf{r}_k|\mathbf{y}_k}\begin{bmatrix}\mathbf{e}_k\mathbf{e}_k^H & \mathbf{e}_k\boldsymbol{\epsilon}_k^H\overline{\mathbf{M}}_k^T\\ \overline{\mathbf{M}}_k\boldsymbol{\epsilon}_k\mathbf{e}_k^H & \overline{\mathbf{M}}_k\mathbf{r}_k\boldsymbol{\epsilon}_k^H\overline{\mathbf{M}}_k^T\end{bmatrix}\end{aligned} \tag{43}$$

where $\mathbf{e}_k = \mathbf{M}_k\boldsymbol{\epsilon}_k = \mathbf{y}_k - \mathbf{M}_k\mathbf{v}$.

Thus, after some algebra, we have

$$\begin{aligned}&\mathbb{E}_{\overline{\mathbf{M}}_k\mathbf{r}_k|\mathbf{y}_k}\left[\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^H\right]\\ &= \mathbf{B}_k^T\mathbb{E}_{\overline{\mathbf{M}}_k\mathbf{r}_k|\mathbf{y}_k}\left[\mathbf{B}_k\boldsymbol{\epsilon}_k\boldsymbol{\epsilon}_k^H\mathbf{B}_k^T\right]\mathbf{B}_k\\ &= \begin{bmatrix}\mathbf{M}_k^T & \overline{\mathbf{M}}_k^T\end{bmatrix}\begin{bmatrix}\mathbf{e}_k\mathbf{e}_k^H & \mathbf{e}_k\boldsymbol{\mu}_k^H\\ \boldsymbol{\mu}_k\mathbf{e}_k^H & \mathbf{G}_k+\boldsymbol{\mu}_k\boldsymbol{\mu}_k^H\end{bmatrix}\begin{bmatrix}\mathbf{M}_k\\ \overline{\mathbf{M}}_k\end{bmatrix}\\ &= \boldsymbol{\xi}_k\boldsymbol{\xi}_k^H + \overline{\mathbf{M}}_k^H\mathbf{G}_k\overline{\mathbf{M}}_k\end{aligned} \tag{44}$$

with

$$\boldsymbol{\xi}_k = \mathbf{M}_k^H\mathbf{e}_k + \overline{\mathbf{M}}_k^H\boldsymbol{\mu}_k \tag{45}$$

Inserting $\boldsymbol{\mu}_k$ and after further simplification, we have

$$\boldsymbol{\xi}_k = \boldsymbol{\Sigma}\mathbf{M}_k^H\left(\mathbf{M}_k\boldsymbol{\Sigma}\mathbf{M}_k^H\right)^{-1}\mathbf{e}_k \tag{46}$$