

RADIO-ASTRONOMY IMAGING AND INTERFERENCE EXCISION USING TENSOR DECOMPOSITION AND CANONICAL CORRELATION ANALYSIS

Mikael Sørensen and Nicholas D. Sidiropoulos

University of Virginia, Charlottesville, USA

ABSTRACT

Antenna arrays with a large number of sensors are becoming increasingly common in radio astronomy. This has motivated the development of array signal processing tools for high-resolution imaging that exploit source signal properties such as sparsity and spectral or temporal variability. We propose a new multi-frequency covariance matrix model for radio astronomical imaging that exploits spectral variability of the astronomical sources. We show that tensor decomposition methods can be used to compute high-resolution images of astronomical scenes that comprise Q point sources. In this context, tensor decomposition can reduce the problem to simpler single-point source imaging problems. We also explain how canonical correlation analysis can be used to mitigate or even altogether remove the effect of (unknown) narrowband interference sources, which is a key challenge in this context.

Index Terms— Radio astronomy, interference excision, tensor decomposition, canonical correlation analysis.

1. INTRODUCTION

Traditional radio-astronomical imaging is non-parametric, using spatial matched filtering to generate a blurred (“dirty”) image of the sky which is subsequently processed by deconvolution algorithms to mitigate severe blurring and noise effects. More modern array signal processing tools have also been used to obtain higher-resolution radio astronomical images of the sky [1]. In particular, covariance matrix based methods have been considered, which exploit that the sources are (approximately) *uncorrelated* and that the astronomical scene is often *sparse* so that it can be modeled as a sum of Q point-sources [2, 3, 4]. More recently, the time variation of radio astronomical signals due to the earth’s rotation has been exploited as a source of signal diversity [5]. In this paper we will consider source signals that exhibit spectral variation, and examine how this property can be exploited. Specifically, we show that by exploiting spectral variability of the sources, it is possible to reduce a Q -point source model into a set of decoupled simpler single-point source models. We also show that the spectral variability enables us, in a relatively simple way, to mitigate or even altogether remove the effect of narrowband interference signals.

The contributions of the paper can be summarized as follows. First, a new multi-frequency based covariance matrix based imaging model that can exploit spectral profile variability of the sources is proposed. Second, a constrained tensor decomposition based framework for computing radio astronomical images that can reduce a Q -point source model into a set of decoupled single-point source models is developed. Third, we explain that Generalized Canonical Correlation Analysis (GCCA) can be used to mitigate or even remove the effect of narrowband interference sources.

(ms8tz,nikos)@virginia.edu. Supported in part by NSF AST-2132700.

2. RADIO-ASTRONOMY IMAGING DATA MODEL

2.1. Review of classical covariance matrix model

The starting point of our work is the covariance matrix model for radio-astronomical imaging discussed in [2, 3, 4] and reviewed next. Consider a radio telescope comprising I sensors, observing Q point-sources (stars). Let $s_q(t, f)$ denote the signal from the q -th source at time-instant t and frequency f . Under the standard narrowband and far-field assumptions, the output of the i -th sensor at frequency f at the n -th time sample can be expressed as $x_i(n, f) = \sum_{q=1}^Q a_{i,q}(n, f) s_q(n, f)$, where $a_{i,q}(n, f) = e^{-i2\pi f \mathbf{p}_i^T \mathbf{z}_q / c}$, in which $i = \sqrt{-1}$, $\mathbf{p}_i = [x_i, y_i]^T$ denotes the location of the i -th sensor, $\mathbf{z}_q = [\cos(\theta_{q,n}) \sin(\phi_{q,n}), \sin(\theta_{q,n}) \sin(\phi_{q,n})]$ is the directional vector in which $\theta_{q,n}$ and $\phi_{q,n}$ denote the azimuth and elevation angles associated with the q -th point-source, and c is the speed of propagation; see [2, 3, 4] for details. Stacking yields

$$\mathbf{x}(n, f) = \mathbf{A}(n, f) \mathbf{s}(n, f), \quad (1)$$

where $\mathbf{x}(n, f) = [x_1(n, f), \dots, x_I(n, f)]^T \in \mathbb{C}^I$, $\mathbf{s}(n, f) = [s_1(n, f), \dots, s_Q(n, f)]^T \in \mathbb{C}^Q$, and antenna response matrix $\mathbf{A}(n, f) = [\mathbf{a}_1(n, f), \dots, \mathbf{a}_Q(n, f)] \in \mathbb{C}^{I \times Q}$ with $\mathbf{a}_q(n, f) = [a_{1,q}(n, f), \dots, a_{I,q}(n, f)]^T \in \mathbb{C}^I$. We assume that $\mathbf{x}(n, f)$ is short-term stationary over a period of N samples. Note that due to earth’s rotation, the directional vector is slowly changing over time. For simplicity, we also assume that the directional vector $\mathbf{z}_q$ is evolving sufficiently slow so that $\mathbf{z}_q := \mathbf{z}_{q,1} = \dots = \mathbf{z}_{q,N}$ and $\mathbf{A}(f) := \mathbf{A}(1, f) = \dots = \mathbf{A}(N, f)$ is a reasonable assumption. Assume also that the sources have zero mean and are uncorrelated, i.e., $\mathbb{E}[\mathbf{s}(n, f) \mathbf{s}(n, f)^H]$ is diagonal, where $\mathbb{E}[\cdot]$ denotes expectation and H denotes conjugate-transpose. Under the mentioned assumptions, the covariance matrix of the observed signals is given by

$$\mathbf{R}(f) := \mathbb{E}[\mathbf{x}(n, f) \mathbf{x}(n, f)^H] = \mathbf{A}(f) \mathbf{B}(f) \mathbf{A}(f)^H, \quad (2)$$

where $\mathbf{B}(f) = \mathbb{E}[\mathbf{s}(n, f) \mathbf{s}(n, f)^H] = \text{diag}(\sigma_1^2(f), \dots, \sigma_Q^2(f))$ is the source covariance matrix with positive diagonal elements. The entries of $\mathbf{R}(f)$ are equal to $(\mathbf{R}(f))_{ij} = \sum_{q=1}^Q I_f(\mathbf{z}_q) e^{-i\mathbf{u}_{ij}^T \mathbf{z}_q}$, where $\mathbf{u}_{ij} = (2\pi f/c)(\mathbf{p}_i - \mathbf{p}_j)$ and

$$I_f(\mathbf{z}) = \sum_{q=1}^Q \sigma_q^2(f) \delta(\mathbf{z} - \mathbf{z}_q) \quad (3)$$

is the brightness image, where $\delta(\cdot)$ denotes the Kronecker delta function. In radio astronomy the goal is to find the brightness image $I_f(\mathbf{z})$. Note that it is parameterized by the source intensities $\{\sigma_q^2(f)\}$ and the directional vectors $\{\mathbf{z}_q\}$.

2.2. A new multi-frequency covariance matrix model

In contrast to the classical covariance matrix model, it is now assumed that we measure the incoming signals at $K \geq 2$ neighbouring

frequencies, $f_1, \dots, f_K$, so that we obtain the following extension of (2):

$$\mathbf{R}(f_k) := \mathbf{A}(f_k)\mathbf{B}(f_k)\mathbf{A}(f_k)^H, \quad k \in \{1, \dots, K\}. \quad (4)$$

Again we assume that the sources are uncorrelated so that $\mathbf{B}(f_k) = \mathbb{E}[\mathbf{s}(n, f_k)\mathbf{s}(n, f_k)^H] = \text{diag}(\sigma_1^2(f_k), \dots, \sigma_Q^2(f_k))$ is a positive diagonal matrix. The key assumption and main idea of this work is to select frequencies so that $|f_{k_1} - f_{k_2}|$ is small relative to the center frequency (e.g., 10 MHz versus several GHz) for all $k_1, k_2 \in \{1, \dots, K\}$. This implies

$$\mathbf{A} := \mathbf{A}(f_1) \approx \dots \approx \mathbf{A}(f_K). \quad (5)$$

The key observation is that by inserting (5) into (4) we obtain the tensor decomposition

$$\mathbf{R}(f_k) = \mathbf{A}\mathbf{B}(f_k)\mathbf{A}^H = \mathbf{A}D_k(\mathbf{B})\mathbf{A}^H, \quad k \in \{1, \dots, K\}, \quad (6)$$

where $\mathbf{B} \in \mathbb{R}^{K \times Q}$ is a nonnegative matrix with entries $(\mathbf{B})_{kq} = \sigma_q^2(f_k)$ and $D_k(\mathbf{B})$ denotes the diagonal matrix that holds the k -th row of $\mathbf{B}$ on its diagonal. When $K \geq 2$, then (6) corresponds to a constrained version of the Canonical Polyadic Decomposition (CPD) [6, 7] that under mild conditions is unique up to column scaling (and irrelevant permutation) ambiguities; see [8] and references therein for details. This implies that the model parameters $\{\theta_q, \phi_q\}$ and $\{\sigma_q^2(f_k)\}$ can be computed via (6), as will be discussed next.

3. CONSTRAINED TENSOR DECOMPOSITION BASED RADIO-ASTRONOMY IMAGING

Based on the proposed multi-frequency covariance matrix model, we will now develop a tensor decomposition based three-step procedure for computing the model parameters needed to construct the brightness image $I_f(\mathbf{z})$ given by (3). The core idea of the tensor decomposition based method is to reduce a Q -point source model into a set of single-point source models. In short, first we compute the factor matrices $\mathbf{A}$ and $\mathbf{B}$ in (6). Next, we extract the directional parameters $\{\theta_q\}$ and $\{\phi_q\}$ via the columns of $\mathbf{A}$. Finally, we compute the source intensities $\{\sigma_q^2(f_k)\}$ from $\{\mathbf{R}(f_k)\}$, given $\{\theta_q\}$ and $\{\phi_q\}$.

3.1. Step 1: Source separation based on constrained CPD

Stacking yields

$$\mathbf{R} = [\text{vec}(\mathbf{R}(f_1)), \dots, \text{vec}(\mathbf{R}(f_K))] = (\mathbf{A}^* \odot \mathbf{A})\mathbf{B}^T, \quad (7)$$

where $\text{vec}(\mathbf{R}(f_k)) \in \mathbb{C}^{I^2}$ denotes the vectorized version of the matrix $\mathbf{R}(f_k) \in \mathbb{C}^{I \times I}$, $*$ denotes conjugation, and $\odot$ denotes the Khatri-Rao (columnwise Kronecker) product. In practice, we compute $\mathbf{A}$ and $\mathbf{B}$ by solving the constrained optimization problem

$$\min_{\mathbf{A} \in \mathbb{C}^{I \times Q}, \mathbf{B} \geq 0} \|\mathbf{R} - (\mathbf{A}^* \odot \mathbf{A})\mathbf{B}^T\|_F^2, \quad (8)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. By ignoring the partial Hermitian symmetry in (8), a simple alternating least squares method can be used to compute $\mathbf{A}$ and $\mathbf{B}$. However, more sophisticated optimization based methods can be derived. More precisely, we can write (8) as an unconstrained Nonlinear Least Squares (NLS) minimization problem

$$\min_{\mathbf{A}_{re}, \mathbf{A}_{im}, \mathbf{C}} \|\mathbf{R}_{re} - (\mathbf{A}_{re} \odot \mathbf{A}_{re} + \mathbf{A}_{im} \odot \mathbf{A}_{im})(\mathbf{C} * \mathbf{C})^T\|_F^2 + \|\mathbf{R}_{im} - (\mathbf{A}_{im} \odot \mathbf{A}_{re} + \mathbf{A}_{im} \odot \mathbf{A}_{re})(\mathbf{C} * \mathbf{C})^T\|_F^2, \quad (9)$$

where $\mathbf{R}_{re}$ and $\mathbf{R}_{im}$ denote the real and imaginary parts of $\mathbf{R}$, $\mathbf{A}_{re}$ and $\mathbf{A}_{im}$ denote the real and imaginary parts of $\mathbf{A}$, and $\mathbf{B} = \mathbf{C} * \mathbf{C}$ for some real matrix $\mathbf{C} \in \mathbb{R}^{K \times Q}$ and where $*$ denotes the Hadamard (elementwise) product. The minimizer of (9) can be computed using standard optimization methods.

The Constant Modulus (CM) property of $\mathbf{A}$ can be taken into account by adding the term $\|\mathbf{A}_{re} * \mathbf{A}_{re} + \mathbf{A}_{im} * \mathbf{A}_{im} - \mathbf{1}_I \mathbf{1}_Q^T\|_F^2$ to the cost function (9), where $\mathbf{1}_n \in \mathbb{R}^n$ denotes an all-ones vector. Overall, we obtain the unconstrained NLS minimization problem

$$\min_{\mathbf{A}_{re}, \mathbf{A}_{im}, \mathbf{C}} \|\mathbf{R}_{re} - (\mathbf{A}_{re} \odot \mathbf{A}_{re} + \mathbf{A}_{im} \odot \mathbf{A}_{im})(\mathbf{C} * \mathbf{C})^T\|_F^2 + \|\mathbf{R}_{im} - (\mathbf{A}_{im} \odot \mathbf{A}_{re} + \mathbf{A}_{im} \odot \mathbf{A}_{re})(\mathbf{C} * \mathbf{C})^T\|_F^2 + \|\mathbf{A}_{re} * \mathbf{A}_{re} + \mathbf{A}_{im} * \mathbf{A}_{im} - \mathbf{1}_I \mathbf{1}_Q^T\|_F^2. \quad (10)$$

An added benefit of taking the CM property into account is that if $\mathbf{B} = \mathbf{C} * \mathbf{C}$ has nearly colinear columns or gets close to a rank-one matrix, the CM constrained CPD of $\mathbf{R}$ can still be unique. We compute $\mathbf{A}_{re}$, $\mathbf{A}_{im}$ and $\mathbf{C}$ by solving (10) using the limited memory BFGS method in Pytorch [9]. Typically the limited memory BFGS method is randomly initialized, or using the best of, say 10 random initializations. However, when $\mathbf{A}$ or $\mathbf{B}$ has full column rank, then algebraic initialization methods can be used. In short, if $\mathbf{A}$ or $\mathbf{B}$ has full column rank, then the CPD factor matrices in (7) can in the exact case be computed via an eigenvalue decomposition; see e.g., [10, 11, 12, 8, 13, 14] for details. This can be used to initialize $\mathbf{A}_{re}$ and $\mathbf{A}_{im}$ in the proposed optimization based method. Matrix $\mathbf{C}$ can be initialized by computing the elementwise square root $\mathbf{B}^{\frac{1}{2}}$, where $\mathbf{B}$ is the matrix obtained by solving (8) conditioned on the initial estimate of $\mathbf{A}$, which is a standard quadratic programming problem.

3.2. Step 2: Extraction of directional vectors via NLS

Once $\mathbf{A}$ has been computed we can extract the directional vectors $\mathbf{z}_q = [\cos(\theta_q) \sin(\phi_q), \sin(\theta_q) \sin(\phi_q)]^T$ from its columns. Let $\mathbf{a}_q$ denote the q -th column of $\mathbf{A}$ and let $\hat{\mathbf{a}}_q(\theta_q, \phi_q) = e^{-i2\pi f \mathbf{P}^T \mathbf{z}_q}$ denote the elementwise exponential function, in which $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_I]$ and we recall that $\mathbf{p}_i$ denotes the location of the i -th sensor. We will compute $\mathbf{z}_q$ via the NLS cost function

$$\min_{\theta_q \in \Omega_\theta, \phi_q \in \Omega_\phi, \alpha_q \in \mathbb{C}} \|\mathbf{a}_q - \alpha_q \hat{\mathbf{a}}_q(\theta_q, \phi_q)\|_2^2, \quad (11)$$

where $\|\cdot\|_2$ denotes the Euclidean norm, and Ω_θ and Ω_ϕ denote the ranges of θ_q and ϕ_q (e.g., $\Omega_\theta = [0, \pi]$ and $\Omega_\phi = [-\pi/2, \pi/2]$). Minimizing (11) is equivalent to maximizing the cost function

$$\max_{\theta_q \in \Omega_\theta, \phi_q \in \Omega_\phi} |\hat{\mathbf{a}}_q(\theta_q, \phi_q)^H \mathbf{a}_q|^2 \quad (12)$$

with $\alpha_q = \hat{\mathbf{a}}_q(\theta_q, \phi_q)^H \mathbf{a}_q / (\hat{\mathbf{a}}_q(\theta_q, \phi_q)^H \hat{\mathbf{a}}_q(\theta_q, \phi_q))$. An initial estimate of θ_q and ϕ_q can now be obtained by a two-dimensional grid search in the area $\Omega_\theta \times \Omega_\phi$. The estimate obtained using grid search can be refined utilizing an optimization-based method. More precisely, the cost function (12) can be expressed as

$$\max_{\theta_q \in \Omega_\theta, \phi_q \in \Omega_\phi} b(\theta_q, \phi_q)^2 + c(\theta_q, \phi_q)^2, \quad (13)$$

where $b(\theta_q, \phi_q) = \mathbf{a}_{q,re}^T \hat{\mathbf{a}}_{q,re}(\theta_q, \phi_q) + \mathbf{a}_{q,im}^T \hat{\mathbf{a}}_{q,im}(\theta_q, \phi_q)$ and $c(\theta_q, \phi_q) = \mathbf{a}_{q,re}^T \hat{\mathbf{a}}_{q,im}(\theta_q, \phi_q) - \mathbf{a}_{q,im}^T \hat{\mathbf{a}}_{q,re}(\theta_q, \phi_q)$, in which $\mathbf{a}_{q,re}$ and $\mathbf{a}_{q,im}$ denote the real and imaginary part of $\mathbf{a}_q$, respectively, and the elementwise cosine function $\hat{\mathbf{a}}_{q,re}(\theta_q, \phi_q) = \cos(2\pi f \mathbf{P}^T \mathbf{z}_q)$ and the elementwise sine function $\hat{\mathbf{a}}_{q,im}(\theta_q, \phi_q) = \sin(2\pi f \mathbf{P}^T \mathbf{z}_q)$ denote the real and imaginary part of $\hat{\mathbf{a}}_q(\theta_q, \phi_q)$, respectively. Due to the Cauchy-Schwartz inequality and the CM property of $\hat{\mathbf{a}}_q(\theta_q, \phi_q)$, maximizing (13) is equivalent to minimizing $\min_{\theta_q \in \Omega_\theta, \phi_q \in \Omega_\phi} I \cdot \|\mathbf{a}_q\|_2^2 - b(\theta_q, \phi_q)^2 - c(\theta_q, \phi_q)^2$. In practice, we relax the constraints on θ_q and ϕ_q and instead solve the unconstrained minimization problem

$$\min_{\theta_q, \phi_q \in \mathbb{R}} I \cdot \|\mathbf{a}_q\|_2^2 - b(\theta_q, \phi_q)^2 - c(\theta_q, \phi_q)^2. \quad (14)$$

We compute θ_q and ϕ_q by solving (14) using the gradient descent method in Pytorch [9] that in turn is initialized by the aforementioned two-dimensional grid-search procedure.

3.3. Step 3: Estimation of source intensities using NNLS

Once the angles $\{\theta_q\}$ and $\{\phi_q\}$ have been obtained, the source intensities $\{\mathbf{B}(f)\}$ can be obtained by solving the following standard NonNegative Least Squares (NNLS) problems [15]:

$$\min_{\tilde{\mathbf{b}}_k \geq 0} \left\| \begin{bmatrix} \text{vec}(\mathbf{R}_{re,k}) \\ \text{vec}(\mathbf{R}_{im,k}) \end{bmatrix} - \begin{bmatrix} \mathbf{A}_{re,k} \odot \mathbf{A}_{re,k} + \mathbf{A}_{im,k} \odot \mathbf{A}_{im,k} \\ \mathbf{A}_{re,k} \odot \mathbf{A}_{im,k} - \mathbf{A}_{im,k} \odot \mathbf{A}_{re,k} \end{bmatrix} \tilde{\mathbf{b}}_k \right\|_2$$

for every $k \in \{1, \dots, K\}$, where $\tilde{\mathbf{b}}_k$ denotes the k -th row of $\mathbf{B}$, $\mathbf{R}_{re,k}$ and $\mathbf{R}_{im,k}$ denote the real and imaginary parts of $\mathbf{R}(f_k)$, and $\mathbf{A}_{re,k}$ and $\mathbf{A}_{im,k}$ denote the real and imaginary parts of $\mathbf{A}(f_k) = [\mathbf{a}_q(f_k, \theta_1, \phi_1), \dots, \mathbf{a}_q(f_k, \theta_Q, \phi_Q)]$, in which $\{\theta_q\}$ and $\{\phi_q\}$ are obtained in the previous step.

4. MITIGATION OF NARROWBAND INTERFERENCE

Assume now that narrowband interference signals are present so that the observation model (1) admits the decomposition

$$\mathbf{x}(n, f_k) = \begin{cases} \mathbf{A}(n, f_k) \mathbf{s}(n, f_k), & k \in \Omega, \\ \mathbf{A}(n, f_k) \mathbf{s}(n, f_k) + \tilde{\mathbf{A}}(n, f_k) \tilde{\mathbf{s}}(n, f_k), & k \in \Omega^c, \end{cases} \quad (15)$$

where $\tilde{\mathbf{s}}(n, f_k) = [\tilde{s}_1(n, f_k), \dots, \tilde{s}_{L_k}(n, f_k)]^T$ denotes the n -th sample of the narrowband interference signals associated with frequency f_k and $\tilde{\mathbf{A}}(n, f_k) = [\tilde{\mathbf{a}}_1(n, f_k), \dots, \tilde{\mathbf{a}}_{L_k}(n, f_k)]$ denotes the associated antenna response matrix, Ω denotes the set of indices associated with frequencies not affected by narrowband interference terms and $\Omega^c = \{1, \dots, K\} \setminus \Omega$. Assume that all the signals are uncorrelated and have zero mean. Then the covariance matrix (6) now admits the decomposition

$$\mathbf{R}(f_k) = \begin{cases} \mathbf{A} D_k(\mathbf{B}) \mathbf{A}^H, & k \in \Omega, \\ \mathbf{A} D_k(\mathbf{B}) \mathbf{A}^H + \tilde{\mathbf{A}} D_k(\tilde{\mathbf{B}}) \tilde{\mathbf{A}}^H, & k \in \Omega^c, \end{cases} \quad (16)$$

where again we used the assumption that $\tilde{\mathbf{A}} := \tilde{\mathbf{A}}(f_1) = \dots = \tilde{\mathbf{A}}(f_K)$ and $\tilde{\mathbf{B}} \in \mathbb{R}^{K \times L}$ is a nonnegative matrix with entries $(\tilde{\mathbf{B}})_{kl} = \tilde{\sigma}_l^2(f_k)$, in which $\tilde{\sigma}_l^2(f_k)$ denotes the intensity of the l -th interference term at frequency f_k , and $D_k(\tilde{\mathbf{B}}) = \mathbf{0}$ if $k \in \Omega$.

In this section we will discuss two ways GCCA can be used to mitigate the impact of the interference signals when $K \geq 2$ frequencies are considered. For simplicity, we consider the case where $\text{rank}([\mathbf{A} D_k(\mathbf{B}), \tilde{\mathbf{A}} D_k(\tilde{\mathbf{B}})]) = Q + L_k$ for all $k \in \{1, \dots, K\}$ (necessitating $Q + L_k \leq I$). This implies that, under certain conditions [16], we can obtain $\text{range}(\mathbf{A})$ via $\bigcap_{k=1}^K \text{range}(\mathbf{R}(f_k))$, where $\text{range}(\mathbf{A})$ denotes the span of $\mathbf{A}$. Specifically, we will present a method for computing $\text{range}(\mathbf{A})$ that can deal with interference terms. Let $\mathbf{P}_k^\perp = \mathbf{I} - \mathbf{U}_k \mathbf{U}_k^H$, where $\mathbf{U}_k \in \mathbb{C}^{I \times (Q+Z_k)}$ is a columnwise orthonormal matrix ($\mathbf{U}_k^H \mathbf{U}_k = \mathbf{I}$) with $Z_k \geq L_k$. It can be shown that, under certain conditions [16], if $\text{range}(\mathbf{U}_k) = \text{range}(\mathbf{R}(f_k))$, then

$$\text{range}(\mathbf{A}) = \bigcap_{k=1}^K \text{range}(\mathbf{R}(f_k)) = \text{range}\left(\sum_{k=1}^K \ker(\mathbf{P}_k^\perp)\right), \quad (17)$$

where $\ker(\mathbf{P}_k^\perp)$ denotes the kernel of $\mathbf{P}_k^\perp$; see [16, 17] for details. Perhaps more interesting, even if $\text{range}(\mathbf{A}) \subseteq \text{range}(\mathbf{U}_k)$ and $Z_k > L_k$, then relation (17) can still hold; see [16] for details. This can be useful when L_k is not known in advance. To summarize, when $K \geq 2$, we can (under certain conditions) obtain a basis for

$\text{range}(\mathbf{A})$ by (i) first computing the SVDs $\mathbf{R}(f_k) = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H$, (ii) next computing $\mathbf{P}_k^\perp = \mathbf{I} - \mathbf{U}_k(:, 1 : Q + Z_k) \mathbf{U}_k(:, 1 : Q + Z_k)^H$ with $Z_k \geq L_k$, $k \in \{1, \dots, K\}$, (iii) computing the SVD $\sum_{k=1}^K \mathbf{P}_k^\perp = \mathbf{U}_{\text{sum}} \mathbf{\Sigma}_{\text{sum}} \mathbf{V}_{\text{sum}}^H$, and (iv) finally constructing $\mathbf{U} = \mathbf{V}_{\text{sum}}(:, I - Q : I)$ which ideally has $\text{range}(\mathbf{U}) = \text{range}(\mathbf{A})$. Here Matlab array indexing notation was used. Alternatively, by computing the SVD $[\mathbf{U}_k(:, 1 : Q + Z_1), \dots, \mathbf{U}_k(:, 1 : Q + Z_K)] = \mathbf{U}_{\text{stacked}} \mathbf{\Sigma}_{\text{stacked}} \mathbf{V}_{\text{stacked}}^H$, a basis for $\text{range}(\mathbf{A})$ can be obtained by constructing $\mathbf{U} = \mathbf{U}_{\text{stacked}}(:, 1 : Q)$ which ideally has the property $\text{range}(\mathbf{U}) = \text{range}(\mathbf{A})$. Finally, we note that Q can be determined as the dimension of $\bigcap_{k=1}^K \text{range}(\mathbf{R}(f_k))$.

4.1. Approach 1: Detection of corrupt channels

In the special case where only a few of the frequency channels are corrupted by narrowband interference signals, the knowledge of $\text{range}(\mathbf{A})$ can be used to determine the index set Ω . In short, let $\mathbf{P}^\perp = \mathbf{I} - \mathbf{U} \mathbf{U}^H$ denote the orthogonal projector onto $\text{range}(\mathbf{A})^\perp$, where $^\perp$ denotes the orthogonal complement and $\mathbf{U} \in \mathbb{C}^{I \times Q}$ is the matrix obtained via the earlier described SVD procedure. Then $\mathbf{P}^\perp \mathbf{A} = \mathbf{0}$ and $\mathbf{P}^\perp \tilde{\mathbf{A}} \neq \mathbf{0}$. Consequently,

$$\mathbf{P}^\perp \mathbf{R}(f_k) = \begin{cases} \mathbf{0}, & k \in \Omega, \\ \mathbf{P}^\perp \tilde{\mathbf{A}} \tilde{\mathbf{R}}(f_k) \tilde{\mathbf{A}}^H, & k \in \Omega^c, \end{cases} \quad (18)$$

From (18) we can determine Ω . By only considering the matrices $\mathbf{R}(f_k)$, $k \in \Omega$, we can now proceed as in Section 3. We also mention that in [18] a joint tensor factorization and corrupt slab suppression method has been proposed that can also be used to find Ω .

4.2. Approach 2: Constrained signal subspace fitting

Let us now consider the more general case where $\Omega = \emptyset$ is permitted. By exploiting that $\mathbf{A}$ is CM constrained, we can recover it from $\text{range}(\mathbf{U})$ using variants of ‘‘ACMA’’ [19], where $\mathbf{U} \in \mathbb{C}^{I \times Q}$ is the matrix obtained via the earlier described SVD procedure. We mention that the combination of CCA where $K = 2$ and ACMA has been considered in [20] and for more general cases $K \geq 1$ in [13]. However, we propose a new optimization method for computing $\mathbf{A}$ via $\text{range}(\mathbf{U})$. Since $\text{range}(\mathbf{U}) = \text{range}(\mathbf{A})$ there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{Q \times Q}$ such that

$$\mathbf{A} = \mathbf{U} \mathbf{F}. \quad (19)$$

Since $\mathbf{A}$ is CM constrained, it must satisfy the relations $a_{i_1 q}^* a_{i_2 q} a_{i_2 q} = 0$, $1 \leq i_1 < i_2 \leq I$. This property combined with (19) yields the relation $\mathbf{C}(\mathbf{F}^* \odot \mathbf{F}) = \mathbf{0}$, where the rows of $\mathbf{C}^{\frac{I(I+1)}{2} \times Q^2}$ are of the form $(\mathbf{e}_{i_1}^T \mathbf{U}^*) \otimes (\mathbf{e}_{i_1}^T \mathbf{U}) - (\mathbf{e}_{i_2}^T \mathbf{U}^*) \otimes (\mathbf{e}_{i_2}^T \mathbf{U})$ in which $\mathbf{e}_i \in \mathbb{R}^I$ denotes a unit vector with unit entry at position i and zero elsewhere. It can be shown that under certain conditions $\mathbf{C}$ has the property $\text{rank}(\mathbf{C}) = Q(Q-1)$; see [13] for details. Consequently, when $\text{rank}(\mathbf{C}) = Q(Q-1)$, then we can from the SVD of $\mathbf{C}$ compute a (columnwise orthonormal) matrix $\mathbf{N} \in \mathbb{C}^{Q^2 \times Q}$ with property $\text{range}(\mathbf{N}) = \ker(\mathbf{C})$ and factorization

$$\mathbf{N} = (\mathbf{F}^* \odot \mathbf{F}) \mathbf{G}^T, \quad (20)$$

where $\mathbf{G} \in \mathbb{C}^{Q \times Q}$ is nonsingular; see again [13] for details. We now recognize that the factorization (20) corresponds to a CPD, implying that $\mathbf{F}$ can be obtained via CPD computation. Once $\mathbf{F}$ has been obtained, then, due to relation (19) and the column scaling ambiguity of the CPD, we obtain $\mathbf{A} \mathbf{D} = \mathbf{U} \mathbf{F}$, where $\mathbf{D}$ is a diagonal nonsingular matrix. We can obtain $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_Q]$ by column scaling $\mathbf{a}_q \leftarrow \sqrt{((\mathbf{a}_q * \mathbf{a}_q)^H \mathbf{1}_I) / ((\mathbf{a}_q * \mathbf{a}_q)^H (\mathbf{a}_q * \mathbf{a}_q))} \mathbf{a}_q$,

$q \in \{1, \dots, Q\}$. Finally, we also compute $\mathbf{F} = \mathbf{U}^H \mathbf{A}$. An advantage of this approach is that it works even in the presence of strong interfering signals since they are not involved in the expression (19).

Compared the above CPD based approach of computing $\mathbf{A}$ via (19) and (20), a more direct way of computing $\mathbf{A}$ will now be put forward. More precisely, we propose here to compute $\mathbf{A}$ in (19) by

$$\min_{\mathbf{A}, \mathbf{H}} \|\mathbf{U} - \mathbf{A}\mathbf{H}\|_F + \|\mathbf{A}^* * \mathbf{A} - \mathbf{1}_I \mathbf{1}_Q^T\|_F, \quad (21)$$

where ideally $\mathbf{H} = \mathbf{F}^{-1} \in \mathbb{C}^{Q \times Q}$. Relation (21) can be written as

$$\begin{aligned} \min_{\mathbf{A}_{re}, \mathbf{A}_{im}, \mathbf{H}_{re}, \mathbf{H}_{im}} & \|\mathbf{U}_{re} - \mathbf{A}_{re} \mathbf{H}_{re} + \mathbf{A}_{im} \mathbf{H}_{im}\|_F + \\ & \|\mathbf{U}_{im} - \mathbf{A}_{re} \mathbf{H}_{im} - \mathbf{A}_{im} \mathbf{H}_{re}\|_F + \\ & \|\mathbf{A}_{re} * \mathbf{A}_{re} + \mathbf{A}_{im} * \mathbf{A}_{im} - \mathbf{1}_I \mathbf{1}_Q^T\|_F, \end{aligned} \quad (22)$$

where $\mathbf{U}_{re}$ and $\mathbf{U}_{im}$ denote the real and imaginary parts of $\mathbf{U}$, $\mathbf{A}_{re}$ and $\mathbf{A}_{im}$ denote the real and imaginary parts of $\mathbf{A}$, and $\mathbf{H}_{re}$ and $\mathbf{H}_{im}$ denote the real and imaginary parts of $\mathbf{H}$. Again, the limited memory BFGS method in Pytorch [9] can be used to solve the minimization problem (22). Note that $\mathbf{A}$ and $\mathbf{H} = \mathbf{F}^{-1}$ can be initialized by the previously discussed CPD based approach.

4.3. Recovery of source intensities when $L_k = 1$

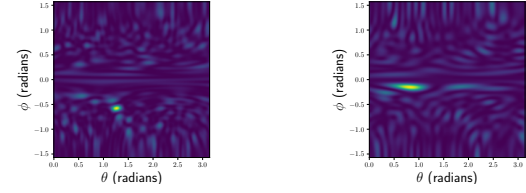
So far we have not discussed how to recover $D_k(\mathbf{B})$ when $k \in \Omega^c$. For simplicity, let us consider the case where $L_k = 1$. In this case the problem of estimating $\tilde{\mathbf{A}}$ can be reduced to a single isolated-source estimation problem. Indeed, if $\mathbf{A}$ has already been computed by for instance solving (21), then $\mathbf{P}^\perp \mathbf{R}(f_k) = \mathbf{P}^\perp \tilde{\mathbf{a}} \tilde{\sigma}^2(f_k) \tilde{\mathbf{a}}^H$ with $\mathbf{P}^\perp = \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$. Then, as explained in Section 3.2, the directional vector $\mathbf{z}$ associated with the interference term can now be determined by solving the NLS problem $\max_{\theta, \phi} |\mathbf{P}^\perp \mathbf{R}(f_k) \tilde{\mathbf{a}}(f_k, \theta, \phi)|^2$, where $\tilde{\mathbf{a}}(f_k, \theta, \phi)$ denotes the antenna response vector associated with the interference signal and parameterized by the angles (θ, ϕ) . Once $\tilde{\mathbf{a}}(f_k, \theta, \phi)$ has been obtained, the source intensities for the frequency channel affected by the interfering narrowband signal can be determined by solving the NNLS problem $\min_{\tilde{\mathbf{b}}_k \geq 0} \|\text{vec}(\mathbf{R}(f_k)) - [\mathbf{A}^* \odot \mathbf{A}, \tilde{\mathbf{a}}^* \otimes \tilde{\mathbf{a}}] \tilde{\mathbf{b}}_k\|_2$, where $\tilde{\mathbf{b}}_k = [\sigma_1^2(f_k), \dots, \sigma_Q^2(f_k), \tilde{\sigma}_1^2(f_k)]^T$.

5. NUMERICAL EXPERIMENTS

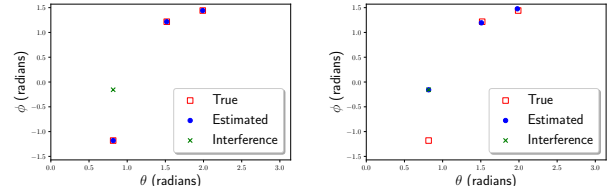
Consider an array composed of $I = 20$ sensors randomly placed in the $[0, 20\lambda] \times [0, 20\lambda]$ plane, where $\lambda = \frac{c}{2f_c}$ is the “half wavelength” and f_c is the center frequency. We set $Q = 3$, $K = 21$, $f \in \{0.999990 \cdot f_c, 0.999991 \cdot f_c, \dots, 1.000009 \cdot f_c, 1.000010 \cdot f_c\}$, $N = 200.000$ and randomly draw the real and imaginary parts of $s_q(n, f)$ and $\tilde{s}_q(n, f)$ from a normal distribution with zero mean and unit variance. Each signal $s_q(n, f)$ and $\tilde{s}_q(n, f)$ is scaled by a factor drawn from uniform distribution with support $[0, 1]$.

5.1. Case 1: Closely spaced (and weak) sources

In this experiment we will demonstrate the potential of the discussed tensor method for detecting closely spaced points in the sky. The azimuth and elevation angles of the point-sources are listed in the first column of Table 1. The image obtained using a Delay-And-Sum (DAS) beamformer [1, 21, 22], denoted by $I(\theta, \phi) = \mathbf{a}(\theta, \phi)^H \mathbf{R}(f_c) \mathbf{a}(\theta, \phi)$ and with discretization steps $\Delta\theta = 0.00628$ and $\Delta\phi = 0.00628$, in which $\mathbf{a}(\theta, \phi)$ is an antenna response vector characterized by the pair (θ, ϕ) , can be seen in Figure 1a. By inspection of the image it is difficult to distinguish sources one and two apart and detecting weak third source.



(a) Case 1. (b) Case 2.
Fig. 1: DAS beamformer images.



(a) Proposed method. (b) OMP.
Fig. 2: Performance of proposed method and OMP method.

A classical technique to resolve closely spaced sources is the Orthogonal Matching Pursuit (OMP) method [23, 24, 25]. In short, it tries to solve the sparse regression problem $\min_{\gamma} \|\text{Re}(\text{vec}(\mathbf{R}(f_c)) - \text{Re}(\mathbf{D})\gamma)\|_2 + \|\text{Im}(\text{vec}(\mathbf{R}(f_c)) - \text{Im}(\mathbf{D})\gamma)\|_2$ subject to $\|\gamma\|_0 \leq Q$, where $\mathbf{D}$ is a user-defined dictionary matrix and $\|\gamma\|_0 = \sum_i [\gamma_i \neq 0]$. We consider an OMP using a dictionary of size (400×10.000) with columns of the form $\mathbf{a}(\theta, \phi)^* \otimes \mathbf{a}(\theta, \phi)$ in which $\mathbf{a}(\theta, \phi)$ is an antenna response vector characterized by the pair (θ, ϕ) . The discretization steps used in the construction of $\mathbf{D}$ are $\Delta\theta = 0.031415$ and $\Delta\phi = 0.031415$. In Table 1 we compare the performance of the proposed and the OMP methods. We observe that the proposed method does a better job.

Angles	True	Proposed	OMP
(θ_1, ϕ_1)	(1.25879, -0.57380)	(1.256637, -0.57176)	(1.25663, -0.62831)
(θ_2, ϕ_2)	(1.26000, -0.58500)	(1.26292, -0.58433)	(1.25663, -0.56548)
(θ_3, ϕ_3)	(0.26360, 1.33558)	(0.26389, 1.33831)	(1.22522, -0.53407)

Table 1: True and estimated angles obtained when the proposed method and OMP are used.

5.2. Case 2: Narrowband interference term

Consider a similar set-up as before, but now a single frequency channel f_c is perturbed by a single interference signal ($L_{f_c} = 1$). The image obtained using a DAS beamformer can be seen in Figure 2b. The true interference signal locations and true and estimated source locations obtained using the proposed and OMP methods can be seen in Figure 2. We observe that a single-channel DAS beamformer or OMP method cannot directly identify the narrowband interference term while the proposed method based on the approach discussed in Section 4.2 was able to identify the location of the sources.

6. CONCLUSION

We have proposed a multi-frequency based covariance matrix model for radio astronomical imaging that is suitable for cases where the spectral profiles of the sources are varying. Based on the proposed model, we developed a tensor decomposition framework that can reduce a Q -point source model of the sky into a set of simpler single-source point models. We also explained that the proposed model enables GCCA to handle narrowband interference signals.

7. REFERENCES

- [1] D. H. Johnson and D. E. Dudgeon, *Array signal processing: Concepts and techniques*, Prentice Hall, 1993.
- [2] A.-J. van der Veen, A. Leshem, and A.-J. Boonstra, "Signal processing for radio astronomical arrays," in *Processing Workshop Proceedings, 2004 Sensor Array and Multichannel Signal*, 2004, pp. 1–10.
- [3] R. Levanda and A. Leshem, "Synthetic aperture radio telescopes," *IEEE Signal Processing Magazine*, vol. 27, no. 1, pp. 14–29, 2010.
- [4] A.-J. van der Veen, S. J. Wijnholds, and A. M. Sardarabadi, "Signal processing for radio astronomy," in *Handbook of signal processing systems*, S. S. Bhattacharyya, E. F. Deprettere, R. Leupers, and J. Takala, Eds., pp. 311–360. Springer international publishing, 2018.
- [5] S. Zhang, Y. Gu, C.-H. Won, and Y. D. Zhang, "Dimension-reduced radio astronomical imaging based on sparse reconstruction," in *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2018, pp. 470–474.
- [6] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart–Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
- [7] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.
- [8] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Processing*, vol. 65, no. 13, pp. 3551–3582, July 2017.
- [9] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, 2019.
- [10] E. Sanchez and B. R. Kowalski, "Tensorial resolution: A direct trilinear decomposition," *J. Chemometrics*, vol. 4, no. 1, pp. 29–45, 1990.
- [11] S. E. Leurgans, R. T. Ross, and R. B. Abel, "A decomposition of three-way arrays," *SIAM J. Matrix Anal. Appl.*, vol. 14, pp. 1064–1083, 1993.
- [12] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.
- [13] M. Sørensen and N. D. Sidiropoulos, "Multi-set low-rank factorizations with shared and unshared components," *IEEE Trans. Signal Processing*, vol. 68, pp. 5122–5137, 2020.
- [14] M. Sørensen, L. De Lathauwer, and N. D. Sidiropoulos, "Bilinear factorizations subject to monomial equality constraints via tensor decompositions," *Linear Algebra and its Applications*, vol. 621, pp. 296–333, 2021.
- [15] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Prentice Hall, 1974.
- [16] M. Sørensen, C. I. Kanatsoulis, and N. D. Sidiropoulos, "Generalized canonical correlation analysis: A subspace intersection approach," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2452–2467, 2021.
- [17] A. Ben-Israel, "Projectors on intersection of subspaces," in *Contemporary Mathematics*, vol. 636, pp. 41–50, 2015.
- [18] X. Fu, K. Huang, W.-K. Ma, N. D. Sidiropoulos, and R. Bro, "Joint tensor factorization and outlying slab suppression with applications," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6315–6327, 2015.
- [19] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, 1996.
- [20] M. Salah Ibrahim and N. D. Sidiropoulos, "Reliable detection of unknown cell-edge users via canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 19, no. 6, pp. 4170–4182, 2020.
- [21] H. L. Van Trees, *Detection, estimation, and modulation theory, optimum array processing (Part IV)*, Wiley-Interscience, 2002.
- [22] P. Stoica and R. L. Moses, *Spectral analysis of signals*, Pearson Prentice Hall, 2005.
- [23] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, July 1993.
- [24] M. Elad, *Sparse and redundant representations*, Springer-Verlag New York, 2010.
- [25] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, Birkhäuser Basel, 2013.