Kalman filter for radio source power and direction of arrival estimation

Cyril Cano, Nawel Arab, Éric Chaumette, Pascal Larzabal, Mohammed Nabil El Korso, Isabelle Vin

Kalman filter for radio source power and direction of arrival estimation

Cyril Cano[1*], Nawel Arab[2], Éric Chaumette[1], Pascal Larzabal[2], Mohammed Nabil El Korso[3], Isabelle Vin[2]

[1*]ISAE-SUPAERO, Université de Toulouse, 10, Avenue Edouard Belin, Toulouse, 31000, France.
[2]SATIE, ENS Paris-Saclay, Université Paris-Saclay, 4, Avenue des sciences, Gif-sur-Yvette, 91190, France.
[3]L2S, CentraleSupelec, Université Paris-Saclay, 3, Rue Joliot-Curie, Gif-sur-Yvette, 91190, France.

*Corresponding author(s). E-mail(s): cyril.cano@isae-supero.fr;
Contributing authors: nawel.arab@ens-paris-saclay.fr;
eric.chaumette@isae-supaero.fr; pascal.larzabal@ens-paris-saclay.fr;
mohammed.nabil.el-korso@centralesupelec.fr;
isabelle.vin@universite-paris-saclay.fr;

Abstract

Images are an important source of information for spacecraft navigation. Based on an image and a known attitude, triangulation techniques (intersection or resection) are often used for positioning and navigation. In the resection problem, an observer estimate its unknown location by using angle measurements to points at known locations (i.e. landmarks), the localization performance depending on the accuracy of the angle measurements. As a contribution to resection for spacecraft navigation, we considers the dynamic image estimation problem based on radio interferometry, i.e. image of radio source power, where the measurements are sample covariance matrices (SCMs). Considering the case where several measurements are available as well as a known dynamic linear model of image evolution, a.k.a a linear state model, the minimum mean-squared error image estimator (MMSE) is given by the Kalman filter (KF) or one of its variants. However standard Kalman-like filters are not a priori suitable for the problem at hand since the measurements (i.e. SCMs) cannot be formulated analytically as a function of state parameters to be estimated (i.e. radio source power). In fact, this lack of

analytical formulation can be circumvented by a statistical linear fitting allowing the SCMs to be expressed in terms of the state. This linear fitting introduces an additive residual noise, equivalent to a measurement noise, whose covariance matrix depends on the current state, a non-standard case for a measurement model. The covariance matrix of the residual noise is derived whatever the distributions of the radio sources and of the additive noise at the samples level, unveiling the contribution of their multivariate kurtosis. The proposed method is evaluated on simulated data representative of a dynamic radio interferometry framework. The results show that the proposed method is capable of effectively tracking moving radio sources in complex scenes with theoretical guaranties when the signal multivariate kurtosis is known.

Keywords: Kalman filter, Sample covariance matrix, Interferometry, Resection problem

1 Introduction

Cameras, telescopes, and similar tools provide essential information for today's spacecraft. These devices serve various purposes, such as guiding, navigation and positioning. Relying on a known object attitude, digital images are used to identify known landmarks in the observed scene to their apparent pixel locations. In the context of guiding spacecraft, it is also known as "angles-only" optical navigation [1]. Triangulation (intersection or resection) from images is used in space navigation challenges like finding your way based on known landmarks, tracking objects using angles, and even guiding spacecraft with the help of stars for interstellar journeys [1]. We place ourselves in the context of resection, i.e. the observer estimates its unknown location by using angle measurements to points at known locations (i.e. landmarks). For instance, using resection, one can figure out where a spacecraft is in cislunar space by looking at known satellites or during interplanetary missions by tracking asteroids and planets [2, 3]. As a contribution to resection for spacecraft navigation, we tackle the dynamic stars image estimation problem based on radio interferometry, i.e. image of radio source power, where the measurements are sample covariance matrices (SCMs), a.k.a visibility matrices [4]. Indeed, such radio interferometric images estimate both power and direction of arrival (DOA) of a radio source, and allow to identify the DOA of radio sources of known power, provided that the power estimation is accurate enough. We consider the case where several measurements are available as well as a known dynamic linear model of image evolution, a.k.a a linear state model.

The design and use of state estimation techniques is fundamental in a plethora of applications, such as robotics, tracking, guidance and navigation systems [5–7]. For a linear dynamic system, the Kalman filter (KF) is the best linear minimum mean square error (MSE) estimator. The most widespread solution for nonlinear systems is to resort to system linearization, leading to the so-called linearized or extended KF (EKF) [7]. In both cases, as well as for more advanced techniques such as sigma-point filters [8], the main assumption is a perfect system knowledge [9]: (i) known process and measurement functions, including their parameters, (ii) known inputs and (iii)

noise statistics (i.e., first and second order moments for the KF and EKF). Thus, usage of KF may not be possible or of poor performance if one (or more) of the above requirements is not met [10].

Consequently, at first sight, Kalman-like filters do not seem to be usable for the problem at hand since, although each individual sample may adhere to a linear parametric model, the finite horizon SCMs cannot be formulated analytically as a function of state parameters to be estimated (radio source power). Fortunately, this lack of analytical formulation can be circumvented by a statistical linear fitting, at least under the assumptions of: a) a deterministic dynamic state model (no state noise), and b) instantaneous linear observations from multiple radio sources in the presence of additive noise (stochastic observation model), when the radio sources are mutually independent and independent from the noise. The proposed linear fitting allows the SCMs to be expressed in terms of the state parameters and introduces an additive residual noise, equivalent to a measurement noise, whose covariance matrix depends on the current state parameters, a non-standard case for a measurement model.
The covariance matrix of the residual noise is derived whatever the distributions of the radio sources and of the additive noise at the samples level, unveiling the contribution of their multivariate kurtosis [11] whose values depend on whether the radio sources and noise distributions are heavy or light tailed. To support the discussion, the proposed method is evaluated on simulated data representative of a dynamic radio interferometic imaging framework.

2 Measurement Model

Let us consider a network composed of M antennas receiving signals at consecutive short time integration (STI) intervals $[t_k, t_k + \epsilon]$, $k \geqslant 0$. During the $k-$th STI interval, observations are i.i.d. realizations of a stochastic variable $\mathbf{z}_k \in \mathbb{C}^{M\times 1}$ consisting of a linear mixture of signals coming from Q radio sources, $\mathbf{s}_k \in \mathbb{C}^{Q\times 1}$, in the presence of an additive noise $\mathbf{n}_k$:

$$\mathbf{z}_k = \mathbf{A}_k \mathbf{s}_k + \mathbf{n}_k, \tag{1}$$

with $\mathbf{A}_k \in \mathbb{C}^{M\times Q}$ the system response matrix, and $\mathbf{s}_k$ and $\mathbf{n}_k$ being independent and centered complex circular random vectors. Especially, we consider the case where we only have access to the sample covariance matrix (SCM)

$$\hat{\mathbf{C}}_{\mathbf{z}_k} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{z}_k(n)\mathbf{z}_k(n)^H \in \mathbb{C}^{M\times M}, \tag{2}$$

where $\mathbf{z}_k(n)$ is the $n-$th realization of $\mathbf{z}_k$, N is the number of samples and the subscript H denotes the hermitian operator, i.e. $\mathbf{z}_k^H \triangleq (\mathbf{z}_k^*)^T$. Asymptotically, i.e. when N tends to infinity, $\hat{\mathbf{C}}_{\mathbf{z}_k}$ converges in probability to

$$\begin{aligned} \mathbf{C}_{\mathbf{z}_k} &= \mathbb{E}\left[\mathbf{z}_k \mathbf{z}_k^H\right] \\ &= \mathbf{A}_k \mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] \mathbf{A}_k^H + \mathbf{C}_{\mathbf{n}_k}. \end{aligned} \tag{3}$$

In practice, $\mathbf{A}_k$ is known and $\mathbf{C}_{\mathbf{n}_k} \in \mathbb{C}^{M\times M}$ is known or can be measured at desired precision.

Considering mutually independent source signals and denoting $\mathbf{x}_k \in \mathbb{R}_+^{Q\times 1}$ the vector composed of individual source signal powers, a.k.a. intensities, the source signals covariance matrix is diagonal and writes

$$\mathbb{E}[\mathbf{s}_k\mathbf{s}_k^H] \triangleq \operatorname{diag}(\mathbf{x}_k) \in \mathbb{R}^{Q\times Q}, \tag{4}$$

using the vector-to-matrix diag operator. Assuming independence among all $\mathbf{n}_k$ and $\mathbf{s}_l$ for $k, l \geqslant 1$, the objective is to estimate the intensities $\mathbf{x}_k$, under the state model

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1}, \tag{5}$$

where the state-transition matrix $\mathbf{F}_{k-1} \in \mathbb{R}^{Q\times Q}$ is known a priori.

2.1 Linear discrete state-space model

One can construct a linear observation model in the asymptotic case, using the vectorized covariance matrix as observation, i.e.

$$\operatorname{vec}(\mathbf{C}_{\mathbf{z}_k}) = [\mathbf{A}_k^* * \mathbf{A}_k]\,\mathbf{x}_k + \operatorname{vec}(\mathbf{C}_{\mathbf{n}_k}), \tag{6}$$

with $*$ being the column-wise Kronecker product, also called the Khatri–Rao product [12]. However, such linear model does not apply in the non asymptotic case. In order to circumvent this limitation, we propose to apply a statistical linear fitting model based on (6) where measurements are the vectorized SCMs, which are concatenated with their complex conjugates for the sake of optimality (see Appendix A), i.e.

$$\mathbf{y}_k \triangleq \begin{pmatrix} \operatorname{vec}\left(\hat{\mathbf{C}}_{\mathbf{z}_k}\right) \\ \operatorname{vec}\left(\hat{\mathbf{C}}_{\mathbf{z}_k}^*\right) \end{pmatrix} \in \mathbb{C}^{2M^2\times 1}. \tag{7}$$

At iteration k, the observation model is defined as

$$\mathbf{H}_k \triangleq \begin{bmatrix} \mathbf{A}_k^* * \mathbf{A}_k \\ \mathbf{A}_k * \mathbf{A}_k^* \end{bmatrix} \in \mathbb{C}^{2M^2\times Q}, \tag{8}$$

and the observation residual is defined as

$$\mathbf{v}_k \triangleq \mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k. \tag{9}$$

The observation residual $\mathbf{v}_k$ corresponds to the approximation error of the vectorized SCM by the linear model $\mathbf{H}_k\mathbf{x}_k$.

In the following, a Kalman filter is derived for the estimation of $\mathbf{x}_k$ based on SCMs measurements. $\mathbf{v}_k$ is considered as the observation noise according to the standard KF formalism. This corresponds to a non-standard linear discrete state space (LDSS) model for which the quantities of interest $\mathbf{x}_k$ are involved in the observation noise.

Section 3 design the KF in accordance with the definition of $\mathbf{v}_k$, that will induce $\mathbf{x}_k$ to appear in the KF algorithm. When needed, in particular in the covariance matrix of $\mathbf{v}_k$, it is proposed to replace $\mathbf{x}_k$ by an estimate in order to conserve the underlying structure.

3 Design of the Kalman filter

3.1 Kalman filter existence

For such a LDSS model, without state noise, one needs to verify that

$$\mathbf{C}_{\mathbf{v}_k,\mathbf{y}_l} = \mathbb{E}\left[\left(\mathbf{v}_k - \mathbb{E}\left[\mathbf{v}_k\right]\right)\left(\mathbf{y}_l - \mathbb{E}\left[\mathbf{y}_l\right]\right)^H\right] \tag{10}$$

is null for $k \geqslant 2$ and $l < k$ in order to prove the existence of a Kalman filter [9]. The verification is straightforward since $\mathbf{y}_k$ and $\mathbf{y}_l$ are independent and $\mathbf{x}_k$ is deterministic, i.e.

$$\begin{aligned}
\mathbf{C}_{\mathbf{v}_k,\mathbf{y}_l} =& \mathbb{E}\left[\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k - \mathbb{E}\left[\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k\right]\right)\left(\mathbf{y}_l - \mathbb{E}\left[\mathbf{y}_l\right]\right)^H\right] \\
=& \mathbb{E}\left[\left(\mathbf{y}_k - \mathbb{E}\left[\mathbf{y}_k\right]\right)\left(\mathbf{y}_l - \mathbb{E}\left[\mathbf{y}_l\right]\right)^H\right] \\
=& \mathbf{C}_{\mathbf{y}_k,\mathbf{y}_l} \\
=& \mathbf{0}.
\end{aligned} \tag{11}$$

3.2 Identification of first and second order statistics

The next step consists in evaluating the first and second order statistics of the observation noise.

3.2.1 Mean of the observation noise

Since $\hat{\mathbf{C}}_{\mathbf{z}_k}$ is an unbiased estimate of $\mathbf{C}_{\mathbf{z}_k}$, one has that

$$\begin{aligned}
\mathbb{E}[\mathbf{y}_k] &= \left[\mathrm{vec}\left(\mathbf{C}_{\mathbf{z}_k}\right)^T, \mathrm{vec}\left(\mathbf{C}^*_{\mathbf{z}_k}\right)^T\right]^T \\
&= \mathbf{H}_k\mathbf{x}_k + \mathbf{v}^a_k,
\end{aligned} \tag{12}$$

where

$$\mathbf{v}^a_k = \left[\mathrm{vec}\left(\mathbf{C}_{\mathbf{n}_k}\right)^T, \mathrm{vec}\left(\mathbf{C}^*_{\mathbf{n}_k}\right)^T\right]^T \tag{13}$$

is the asymptotic measurement noise. From (9), one obtains that $\mathbb{E}[\mathbf{v}_k] = \mathbb{E}[\mathbf{y}_k] - \mathbf{H}_k\mathbf{x}_k$ and therefore

$$\mathbb{E}[\mathbf{v}_k] = \mathbf{v}^a_k. \tag{14}$$

3.2.2 Covariance of the observation noise

As $\mathbf{x}_k$ is deterministic, one obtains $\mathbf{C}_{\mathbf{v}_k} = \mathbf{C}_{\mathbf{y}_k}$, with

$$\begin{aligned}
\mathbf{C}_{\mathbf{y}_k} =& \mathbb{E}\left[\left(\mathbf{y}_k - \mathbb{E}\left[\mathbf{y}_k\right]\right)\left(\mathbf{y}_k - \mathbb{E}\left[\mathbf{y}_k\right]\right)^H\right] \\
=& \mathbb{E}\left[\begin{pmatrix} \mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)\mathbf{z}_k^H(n)\right) - \mathbb{E}\left[\mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)\mathbf{z}_k^H(n)\right)\right] \\ \mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)^*\mathbf{z}_k^T(n)\right) - \mathbb{E}\left[\mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k^*(n)\mathbf{z}_k^T(n)\right)\right] \end{pmatrix}\right. \\
& \left.\times \begin{pmatrix} \mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)\mathbf{z}_k^H(n)\right) - \mathbb{E}\left[\mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)\mathbf{z}_k^H(n)\right)\right] \\ \mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k(n)^*\mathbf{z}_k^T(n)\right) - \mathbb{E}\left[\mathrm{vec}\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{z}_k^*(n)\mathbf{z}_k^T(n)\right)\right] \end{pmatrix}^H\right] \\
=& \frac{1}{N}\mathbb{E}\left[\begin{pmatrix} \mathrm{vec}\left(\mathbf{z}_k\mathbf{z}_k^H\right) - \mathbb{E}\left[\mathrm{vec}\left(\mathbf{z}_k\mathbf{z}_k^H\right)\right] \\ \mathrm{vec}\left(\mathbf{z}_k^*\mathbf{z}_k^T\right) - \mathbb{E}\left[\mathrm{vec}\left(\mathbf{z}_k^*\mathbf{z}_k^T\right)\right] \end{pmatrix}\begin{pmatrix} \mathrm{vec}\left(\mathbf{z}_k\mathbf{z}_k^H\right) - \mathbb{E}\left[\mathrm{vec}\left(\mathbf{z}_k\mathbf{z}_k^H\right)\right] \\ \mathrm{vec}\left(\mathbf{z}_k^*\mathbf{z}_k^T\right) - \mathbb{E}\left[\mathrm{vec}\left(\mathbf{z}_k^*\mathbf{z}_k^T\right)\right] \end{pmatrix}^H\right] \\
=& \frac{1}{N}\mathbf{C}_{\begin{pmatrix}\mathbf{z}_k^*\otimes\mathbf{z}_k \\ \mathbf{z}_k\otimes\mathbf{z}_k^*\end{pmatrix}},
\end{aligned} \tag{15}$$

where $\otimes$ is the usual Kronecker product, such that $\mathbf{z}_k^* \otimes \mathbf{z}_k = \mathrm{vec}(\mathbf{z}_k\mathbf{z}_k^H)$, $\mathbf{z}_k \otimes \mathbf{z}_k^* = \mathrm{vec}(\mathbf{z}_k^*\mathbf{z}_k^T)$ and $\mathbf{z}_k(n)$ are N independent replicates of $\mathbf{z}_k$. Finally, one has that

$$\mathbf{C}_{\mathbf{v}_k} = \frac{1}{N}\begin{bmatrix} \mathbf{C}_{\mathbf{z}_k^*\otimes\mathbf{z}_k} & \mathbf{C}_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*} \\ \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*} & \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}_k} \end{bmatrix}. \tag{16}$$

Based on the independence of source signals and noises as well as their circularity, Appendix B develops expression of $\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}}$ with respect to Kronecker product properties. After dropping the index k in order to shorten expressions, one obtains

$$\begin{aligned}
\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}} = \mathbf{C}_{\mathbf{z}}^T \otimes \mathbf{C}_{\mathbf{z}} + \sum_{q=1}^{Q}\left(\mathbf{a}_q^*\otimes\mathbf{a}_q\right)\left(\mathbf{a}_q^*\otimes\mathbf{a}_q\right)^H \rho_{s_q} {x_q}^2 \\
+ \mathbf{C}_{\mathbf{n}^*\otimes\mathbf{n}} - \mathbf{C}_{\mathbf{n}}^T \otimes \mathbf{C}_{\mathbf{n}}.
\end{aligned} \tag{17}$$

for source signals $(s_q)_{q=1,\ldots,Q}$ and noises $(n_m)_{m=1,\ldots,M}$ following any distributions, where $x_q = \mathbb{E}\left[|s_q|^2\right]$ is the $q-$th coordinate of $\mathbf{x}$, $\mathbf{a}_q$ is the $q-$th column of $\mathbf{A}$ and

$$\rho_{s_q} = \frac{\mathbb{E}\left[\left(s_q^* s_q\right)^2\right]}{{x_q}^2} - 2 \tag{18}$$

is the normalized multivariate kurtosis of s_q [11].

Finally, note that for

$$\mathbf{P} = \sum_{m=1}^{M} \sum_{m'=1}^{M} \left(\mathbf{e}_m \otimes \mathbf{e}_{m'}\right) \left(\mathbf{e}_{m'} \otimes \mathbf{e}_m\right)^T, \tag{19}$$

with $(\mathbf{e}_1, \ldots, \mathbf{e}_M)$ the canonical basis of $\mathbb{R}^M$, one obtains that

$$\mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} \mathbf{P} \mathbf{e}_{l+M(l'-1)} = \mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} \mathbf{e}_{l'+M(l-1)} \tag{20}$$

from which the following equality holds:

$$\mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}, \mathbf{z} \otimes \mathbf{z}^*} = \mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} \mathbf{P}, \tag{21}$$

for any distribution.

Independent noise components

In particular, for a vector $\mathbf{n}$ composed of independent variables n_m, Appendix C shows that the kurtosis of the noise also appears in $\mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}}$, which is a diagonal matrix such that

$$\left(\mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}}\right)_{m+M(m'-1),\, m+M(m'-1)} = \begin{cases} (\rho_{n_m} + 1)\sigma_{n_m}^4, & m = m'; \\ \sigma_{n_m}^2 \sigma_{n_{m'}}^2, & m \neq m', \end{cases} \tag{22}$$

where

$$\rho_{n_m} = \frac{\mathbb{E}\left[\left(n_m^* n_m\right)^2\right]}{\sigma_{n_m}^4} - 2 \tag{23}$$

and σ_{n_m} are respectively the normalized multivariate kurtosis and the variance of n_m.

Gaussian noise distribution

For instance, for n_m such that $\mathfrak{Re}(n_m)$ and $\mathfrak{Im}(n_m)$ are independent and follows the same distribution $\mathcal{D}$,

$$\rho_{n_m} = \frac{1}{2} \left(\text{kurtosis}\left(\mathcal{D}\right) - 3\right). \tag{24}$$

Since the kurtosis of a univariate real-valued Gaussian random variable is 3, $\rho_{n_m} = 0$ for Gaussian noise, but in general $\rho_{n_m} \neq 0$. Hence, for centered and Gaussian complex circular noise, one has that

$$\mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}} = \mathbf{C}_{\mathbf{n}}^T \otimes \mathbf{C}_{\mathbf{n}}, \tag{25}$$

from which

$$\mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} = \mathbf{C}_{\mathbf{z}}^T \otimes \mathbf{C}_{\mathbf{z}} + \sum_{q=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right) \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)^H \rho_{s_q} {x_q}^2, \tag{26}$$

Gaussian noise and signal distributions

It simplifies for centered and Gaussian complex circular signal and noise, for which

$$\mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} = \mathbf{C}_{\mathbf{z}}^T \otimes \mathbf{C}_{\mathbf{z}}. \tag{27}$$

The latter can be proven by remarking that $\rho_{s_q} = 0$ in (26), or following Appendix C since $\mathbf{z}$ is a vector of centered and Gaussian complex circular random variables.

3.3 Kalman filter recursion

At iteration k, the current estimate writes

$$\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}_k - \widehat{\mathbf{y}}_{k|k-1} \right), \tag{28}$$

where state and measurement predictions are

$$\widehat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1} \widehat{\mathbf{x}}_{k-1|k-1}, \quad \widehat{\mathbf{y}}_{k|k-1} = \mathbf{H}_k \widehat{\mathbf{x}}_{k|k-1} + \mathbf{v}_k^a. \tag{29}$$

The observation noise covariance $\mathbf{C}_{\mathbf{v}_k}$ is unknown since it depends on the current state $\mathbf{x}_k$ that has to be estimated. An estimate $\widehat{\mathbf{C}}_{\mathbf{v}_k}$ is used, from which follows that the innovation covariance matrix $\mathbf{S}_{k|k-1} \triangleq \mathbf{C}_{\mathbf{y}_k - \widehat{\mathbf{y}}_{k|k-1}}$ and hence the predicted and a posteriori estimate error covariance matrices, i.e. $\mathbf{P}_{k|k-1} \triangleq \mathbf{C}_{\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1}}$ and $\mathbf{P}_{k|k} \triangleq \mathbf{C}_{\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k}}$ respectively, are estimated. The optimal Kalman gain $\mathbf{K}_k$ is computed with the recursion

$$\begin{cases} \widehat{\mathbf{P}}_{k|k-1} &= \mathbf{F}_{k-1} \widehat{\mathbf{P}}_{k-1|k-1} \mathbf{F}_{k-1}^H; \\ \widehat{\mathbf{S}}_{k|k-1} &= \mathbf{H}_k \widehat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^H + \widehat{\mathbf{C}}_{\mathbf{v}_k}(\pi_D \left(\widehat{\mathbf{x}}_{k|k-1} \right)); \\ \mathbf{K}_k &= \widehat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^H \left(\widehat{\mathbf{S}}_{k|k-1} \right)^{-1}; \\ \widehat{\mathbf{P}}_{k|k} &= \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \widehat{\mathbf{P}}_{k|k-1}. \end{cases} \tag{30}$$

The measurement noise covariance estimate is constructed with the state prediction $\widehat{\mathbf{x}}_{k|k-1}$. Since there is no non-negativity constraint on the estimation, one expect that $\widehat{\mathbf{x}}_{k|k}$ and $\widehat{\mathbf{x}}_{k|k-1}$ fluctuate around the mean value $\mathbf{x}_k$, such that the estimation of the KF can include negative intensities. Typically, this corresponds to actual low or null intensities and those negative estimates are kept in the recursion. On the other hand, one needs to conserve the positiveness of the estimated measurement noise covariance matrix, that is to use a projection of $\widehat{\mathbf{x}}_{k|k-1}$ on a certain domain $D \subseteq \mathbb{R}_+^{Q \times 1}$ (which remains to apply a threshold on $\widehat{\mathbf{x}}_{k|k-1}$ coordinates). Hence, the considered noise covariance estimate is $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\pi_D \left(\widehat{\mathbf{x}}_{k|k-1} \right))$, where π_D is the projector on D. Even though it is crucial that negative values of $\widehat{\mathbf{x}}_{k|k}$ and $\widehat{\mathbf{x}}_{k|k-1}$ are retained in the recursion (29) to fit with the analytic recursion (30) [5], the estimate that must be considered in physical application is indeed its thresholded version $\pi_D \left(\widehat{\mathbf{x}}_{k|k} \right)$ (as in Section 4).

3.4 Performance analysis

Since the measurement noise covariance matrix $\mathbf{C}_{\mathbf{v}_k}$ involves the actual state to be estimated and is thus unknown, the proposed KF is misspecified. Performances obtained

by applying the designed KF on the actual signal model are suboptimal with respect to the MSE, i.e. compared to what can ideally achieve the KF with a perfect knowledge of the system (i.e. when $\mathbf{C}_{\mathbf{v}_k}$ is known). Following the discussion in [13], the performance obtained by applying the designed KF on the actual signal model is

$$\mathbf{P}^{\mathrm{p}}_{k|k} = \mathbf{C}_{\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k}}, \tag{31}$$

and can be estimated by Monte Carlo simulations[1]. The error covariance matrix $\widehat{\mathbf{P}}_{k|k}$ computed by KF recursion (30) is an estimation of $\mathbf{P}^{\mathrm{p}}_{k|k}$. Typically, poor-quality estimator of $\mathbf{C}_{\mathbf{v}_k}$ leads to poor-quality performance estimator $\widehat{\mathbf{P}}_{k|k}$ which decreases the true performances $\mathbf{P}^{\mathrm{p}}_{k|k}$.

An ideal KF, corresponding to the best linear unbiased estimator (BLUE) in the MSE sense, is constructed by considering the actual measurement noise covariance $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\mathbf{x}_k) \triangleq \mathbf{C}_{\mathbf{v}_k}$ (not feasible in practice). The error covariance matrix of the ideal KF, denoted $\mathbf{P}^{\mathrm{a}}_{k|k}$, is computed by KF recursion and verifies that

$$\mathbf{P}^{\mathrm{p}}_{k|k} \geqslant \mathbf{P}^{\mathrm{a}}_{k|k}. \tag{32}$$

There are different sources of misspecification that can combine and decrease KF performances. Indeed, the very first hypotheses mentioned in Section 2 must be verified in order for the considered LDSS to apply: independence properties and validity of the observation and state transition models (i.e. $\mathbf{A}_k$ and $\mathbf{F}_k$). Systematic tests can be applied in order to verify the validity of these hypothesis. Besides this, the purpose of this work lies in the analytical expression of $\mathbf{C}_{\mathbf{v}_k}$ and its usage in the KF: using $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\pi_D\left(\widehat{\mathbf{x}}_{k|k-1}\right))$ constructed from (16) improves performances compared with default model of $\mathbf{C}_{\mathbf{v}_k}$ such as $\sigma^2\mathbf{I}_{2M^2}$ (it also improves accuracy of $\widehat{\mathbf{P}}_{k|k}$). This becomes especially true for smaller samples size since estimation errors of $\mathbf{C}_{\mathbf{z}^*_k \otimes \mathbf{z}_k}$ becomes more significant in $\mathbf{C}_{\mathbf{v}_k}$ expression ($\mathbf{C}_{\mathbf{v}_k} \approx \mathbf{0}$ in the asymptotic regime). Once expression (16) is used, a second source of error concerns $\mathbf{C}_{\mathbf{n}_k}$ misspecification (e.g. validity of the independence assumption of $\mathbf{n}_k$ coordinates and $(\mathbf{n}_k)_m$ kurtosis) and the signal kurtosis misspecification. The latter are also more significant for smaller sample sizes.

While $\widehat{\mathbf{x}}_{k|k}$ is not thresholded during the KF recursion, one may consider $\pi_D\left(\widehat{\mathbf{x}}_{k|k}\right)$ as a final estimator. Yet performance of $\pi_D\left(\widehat{\mathbf{x}}_{k|k}\right)$ is not given analytically, but it can be estimated by Monte Carlo simulations. These can be significantly better than for $\widehat{\mathbf{x}}_{k|k}$, in particular for small or null values.

3.5 Kalman filter initialization

KF formalism supposes that the first and second order moments of $\mathbf{x}_0$ and $\mathbf{v}_k$ for $k \geqslant 0$ are known, which in the case of a deterministic state vector $\mathbf{x}_k$ amounts to know $\mathbf{x}_0$. However, $\mathbf{x}_0$ is unknown and must be estimated (which is the aim of the

[1] Note that the recursive estimation procedure of $\mathbf{P}^{\mathrm{p}}_{k|k}$ presented in [13] is not tractable in practice since $\mathbf{C}_{\mathbf{v}_k}$ is unknown.

present paper); therefore, the filter must be initialized with an estimate $\widehat{\mathbf{x}}_{0|0}$ and its corresponding mean square error $\mathbf{P}_{0|0} = \mathbf{C}_{\widehat{\mathbf{x}}_{0|0}}$.

A plethora of estimators can be used [14], provided that the associated covariance is known. Among them, a linear unbiased estimator $\widehat{\mathbf{x}}_{0|0}$ is of particular interest since the unbiased property is conserved by the KF. It can be obtained by a Distortionless Response Filter (DRF) $\mathbf{K}_0$ verifying $\mathbf{K}_0\mathbf{H}_0 = \mathbf{I}_Q$, which would give

$$\widehat{\mathbf{x}}_{0|0} = \mathbf{K}_0 \left(\mathbf{y}_0 - \mathbb{E}\left[\mathbf{y}_0\right]\right) \text{ and } \mathbf{P}_{0|0} = \mathbf{K}_0\mathbf{C}_{\mathbf{v}_0}\mathbf{K}_0^H. \tag{33}$$

Among the DRFs, the Minimum Variance Distortionless Filter (MVDRF) given by

$$\mathbf{K}_0 = \left(\mathbf{H}_0^H\mathbf{C}_{\mathbf{v}_0}^{-1}\mathbf{H}_0\right)^{-1}\mathbf{H}_0^H\mathbf{C}_{\mathbf{v}_0}^{-1}, \tag{34}$$

minimizes the covariance (hence the MSE), is used in this work.

However, in some practical cases (including the one under consideration), $\mathbf{C}_{\mathbf{v}_0}$ is not invertible, and a specific form must be used. Based on a compact eigen value decomposition (EVD) of the measurement noise covariance and a compact singular value decomposition (SVD) of the measurement model, i.e.

$$\mathbf{C}_{\mathbf{v}_0} \overset{\text{compact EVD}}{=} \mathbf{W}\mathbf{D}\mathbf{W}^H \quad \text{and} \quad \mathbf{H}_0 \overset{\text{compact SVD}}{=} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H; \tag{35}$$

for $\mathbb{U}$ such that $[\mathbf{U}\ \mathbb{U}]$ forms an orthonormal basis of $\mathbb{R}^{2M^2}$ and

$$\mathbf{W}^H\mathbb{U} \overset{\text{compact SVD}}{=} \underline{\mathbf{U}}\underline{\mathbf{\Sigma}}\underline{\mathbf{V}}^H, \tag{36}$$

Appendix D shows that the MVDRF writes

$$\mathbf{K}_0 = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H - \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{W}\mathbf{D}\underline{\mathbf{U}}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{\Sigma}}^{-1}\underline{\mathbf{V}}^H\mathbb{U}^H. \tag{37}$$

It is clear from (33) that, any DRF implementation supposes to know the observation noise covariance matrix $\mathbf{C}_{\mathbf{v}_0}$, which, due to the unconventional construction of the KF, depends on the current state $\mathbf{x}_0$. As for $k \geqslant 1$, the KF must be computed by substituting $\mathbf{C}_{\mathbf{v}_0}$ by an estimate $\widehat{\mathbf{C}}_{\mathbf{v}_0}$.

In the following results, at initialization, the observation noise covariance matrix $\widehat{\mathbf{C}}_{\mathbf{v}_0}$ injected in the MVDRF initialization is built with a beamforming estimator

$$\widehat{\mathbf{x}}_{0|0}^{\text{BF}} = \text{diag}\left(\mathbf{A}_k^H\left(\widehat{\mathbf{C}}_{\mathbf{z}_0} - \mathbf{C}_{\mathbf{n}_0}\right)\mathbf{A}_k\right) \odot \text{diag}\left(\mathbf{A}_k^*\mathbf{A}_k\right)^{-1} \tag{38}$$

where $\odot$ is the Hadamar product. By construction

$$\mathbb{E}\left[\left(\widehat{\mathbf{x}}_{0|0}^{\mathrm{BF}}\right)_q\right] = (\mathbf{x}_0)_q + \sum_{q'=1,q'\neq q}^{Q} (\mathbf{x}_0)_{q'} \frac{\left|\mathbf{a}_q^H \mathbf{a}_{q'}\right|^2}{\left|\mathbf{a}_q^H \mathbf{a}_q\right|^2} \geqslant (\mathbf{x}_0)_q \tag{39}$$

for $q = 1, \ldots, Q$, and $\left(\widehat{\mathbf{x}}_{0|0}^{\mathrm{BF}}\right)_q$ can be considered as an unbiased estimator of an upper bound of $(\mathbf{x}_0)_q$ and in practice $\left(\widehat{\mathbf{x}}_{0|0}^{\mathrm{BF}}\right)_q \gg (\mathbf{x}_0)_q$. Then from Appendix E one obtains

$$\widehat{\mathbf{C}}_{\mathbf{v}_0}\left(\widehat{\mathbf{x}}_{0|0}^{\mathrm{BF}}\right) \geqslant \mathbf{C}_{\mathbf{v}_0}, \tag{40}$$

which implies that the performance of the KF are upper bounded by $\widehat{\mathbf{P}}_{0|0}$ [13] (i.e. the KF estimate $\widehat{\mathbf{P}}_{0|0}$ is pessimistic). In particular,

$$\widehat{\mathbf{P}}_{0|0} \geqslant \mathbf{P}_{0|0}^{\mathrm{p}} \geqslant \mathbf{P}_{0|0}^{\mathrm{a}}, \tag{41}$$

which is not necessarily verified for $k > 0$.

4 Application to dynamic imaging for spacecraft navigation

As a contribution to resection for spacecraft navigation, we tackle the dynamic stars image estimation problem based on radio interferometry, i.e. image of radio source power. Indeed, as higlighted below, such radio interfeometric images estimate both power (i.e. x_q) and DOA of a radio source (i.e. $\mathbf{u}_q$ in (42)), and allow to identify the DOA of radio sources of known power, provided that the power estimation is accurate enough.

For instance, in order to achieve a satisfactory DOA resolution, modern radio astronomy observatories are interferometers which consist of a network of M antennas, possibly scattered around the globe. Considering their respective positions $\mathbf{r}_m \in \mathbb{R}^2$, wavelength λ and gain function $g_m(.)$ for $m = 1, \ldots, M$, the transfer function associated to a band-limited signal modulated by a carrier frequency $f_c = c/\lambda$ is given by

$$\mathbf{a}(\mathbf{u}) = \left(g_1(\mathbf{u})\, e^{2j\pi \frac{\mathbf{r}_1^T \mathbf{u}}{\lambda}}, \ldots, g_M(\mathbf{u})\, e^{2j\pi \frac{\mathbf{r}_M^T \mathbf{u}}{\lambda}}\right)^T, \tag{42}$$

where $\mathbf{u} \in \mathbb{R}^2$ is the DOA, i.e. the unit vector oriented towards the source of the signal, and c is the speed of light in a vacuum. Given an image of Q pixels, there are respectively Q direction of arrivals $\mathbf{u}_q$ defining the array response matrix $\mathbf{A} = (\mathbf{a}(\mathbf{u}_1), \ldots, \mathbf{a}(\mathbf{u}_Q)) \in \mathbb{C}^{M \times Q}$ [4]. During the $k-$th STI interval, signals from the different radio sources are complex circular, mutually independent. Hereinafter, they are concatenated in a vector $\mathbf{s}_k \in \mathbb{C}^{Q \times 1}$. The signal received by the network of antennas $\mathbf{z}_k \in \mathbb{C}^{M \times 1}$ is modeled as (1), where $\mathbf{n}_k \in \mathbb{C}^{M \times 1}$ is the measurement noise of the

interferometer, which is complex circular, Gaussian distributed and independent of the source signals $\mathbf{s}_k$. The observed visibility matrix, i.e. the SCM $\hat{\mathbf{C}}_{\mathbf{z}_k} \in \mathbb{C}^{M \times M}$, is computed as in (2) from N independent and identically distributed realizations of $\mathbf{z}_k$.

The proposed formalization of the KF provides an estimate of the time-varying power of the radio sources based on $\hat{\mathbf{C}}_{\mathbf{z}_l}$ measurements for $0 \leqslant l \leqslant k$ in the case of a deterministic state model (5) and which can be computed iteratively. The source power estimation can be used to identify and track landmarks of known power, their DOAs being used for resection.

4.1 Results

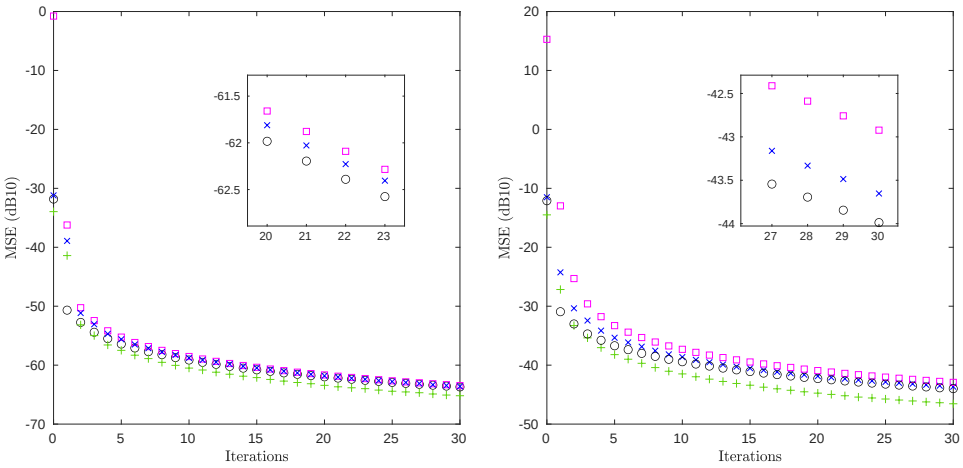
The considered network has a "Y" shape, similarly to the Karl G. Jansky Very Large Array (VLA) observatory [15] and is composed of 27 independent antennas divided in three branches. The method is evaluated on a synthetic image composed of $Q = 22 \times 22$ pixels, with a state-transition model $\mathbf{F}_k \in \mathbb{R}^{Q \times Q}$ being a rotation matrix which corresponds to a rotation of the image by a fixed angle of 90 deg between each STI interval. $\mathbf{F}_k$ is supposed to be known from an inertial motion sensor. Without loss of generality, all the gains g_m are considered equal to 1. Different sample sizes are considered for the SCM construction, that is $N = 10^5$ and $N = 10^3$ samples of bivariate signals $\mathbf{s}_k$ and Gaussian noises $\mathbf{n}_k$. A constant normalized noise covariance matrix is considered as $\mathbf{C}_{\mathbf{n}_k} \triangleq \mathbf{I}_M$. Signals have independent real and imaginary parts following a Laplace distribution, i.e. $\rho_{s_q} = 3/2$. The wavelength is taken such that $\lambda \triangleq 1$, which remains to express coordinates of antennas as function of the wavelength.

Performances of the KF are illustrated in Figure 1 in terms of MSE, defined as

$$\mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right) = \mathbb{E}\left(\left\|\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k}\right\|_2^2\right). \tag{43}$$

Considered images and KF estimates are displayed in Figure 3. The assumed KF performance obtained from KF recursion $\widehat{\mathrm{MSE}}\left(\widehat{\mathbf{x}}_{k|k}\right) = \mathrm{trace}\left(\mathbb{E}\left[\widehat{\mathbf{P}}_{k|k}\right]\right)$, the true KF performance $\mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right) = \mathrm{trace}\left(\mathbf{P}^p_{k|k}\right)$ (computed by Monte Carlo simulations) and the lower bound $\mathrm{MSE}^a\left(\widehat{\mathbf{x}}_{k|k}\right) = \mathrm{trace}\left(\mathbf{P}^a_{k|k}\right)$ (computed by KF with $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\mathbf{x}_k) \triangleq \mathbf{C}_{\mathbf{v}_k}$) are given with respect to the number of iterations. It is observed that for such reasonable configuration, $\widehat{\mathrm{MSE}}\left(\widehat{\mathbf{x}}_{k|k}\right) \geqslant \mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right) \geqslant \mathrm{MSE}^a\left(\widehat{\mathbf{x}}_{k|k}\right)$ for $k \geqslant 0$, although it was proved only for $k = 0$. One shows that those three (i.e. $\widehat{\mathrm{MSE}}\left(\widehat{\mathbf{x}}_{k|k}\right)$, $\mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right)$ and $\mathrm{MSE}^a\left(\widehat{\mathbf{x}}_{k|k}\right)$) appear to converge towards each other along time. $\mathrm{MSE}^p\left(\pi_D\left(\widehat{\mathbf{x}}_{k|k}\right)\right)$ is presented for $D = \mathbb{R}_+^{Q \times 1}$, illustrating that the truncated estimator can achieve better performances than the ideal KF.

Figures 1 and 3 present the results for $N = 10^3$ and $N = 10^5$. Notably, the KF can reach the same performance with a lower sample size but with an increasing iteration number. For $N = 10^5$, the true KF MSE is below -50 db10 for $k \geqslant 3$, while it is for $k \geqslant 120$ with $N = 10^3$. Indeed, performances increase with the number of samples, i.e. as the SCM $\widehat{\mathbf{C}}_{\mathbf{z}_k}$ converges to $\mathbf{C}_{\mathbf{z}_k}$. The effect of the sample size N is clearly visible in



(a) (b)

Fig. 1: MSE lower bound $\mathrm{MSE}^a\left(\widehat{\mathbf{x}}_{k|k}\right)$ (black circles), MSE predicted by KF $\widehat{\mathrm{MSE}}\left(\widehat{\mathbf{x}}_{k|k}\right)$ (pink boxes), MSE actually achieved $\mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right)$ (blue crosses) of the proposed FK and MSE achieved by the thresholded estimates $\mathrm{MSE}^p\left(\pi_{\mathbb{R}_+^{Q\times 1}}\left(\widehat{\mathbf{x}}_{k|k}\right)\right)$ (green plus signs), for **(a)** $N = 10^5$ and **(b)** $N = 10^3$ **s**amples. Initialization was performed with a MVDRF.

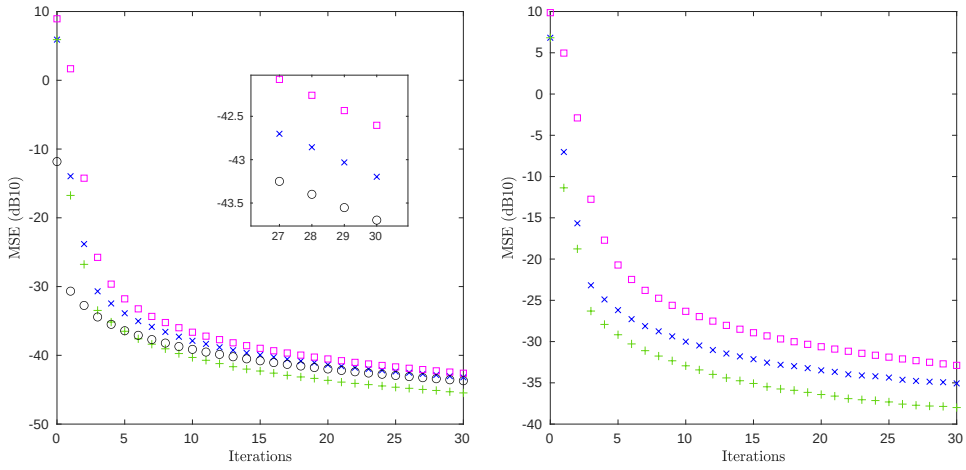
the expression of the observation noise covariance matrix (16): increasing N reduces the variance of the observation noise $\mathbf{v}_k$, and therefore reduces the estimation error.

The MVDRF is limited to a low number of pixels (it can go up to $M(M-2) = 675$ pixels). Other estimators could be used in a general context, e.g. $\widehat{\mathbf{x}}_{0|0} \triangleq \widehat{\mathbf{x}}_{0|0}^{\mathrm{BF}}$. Performances of the latter are illustrated in Figures 2 and 4. The associated error covariance matrix expression is not known, and it is considered that $\widehat{\mathbf{P}}_{0|0} \triangleq 2\mathrm{diag}\left(\widehat{\mathbf{x}}_{0|0} \odot \widehat{\mathbf{x}}_{0|0}\right)$. Two sets of direction of arrival are considered: the same as previously with $Q = 22\times 22$ and one with a better resolution of $Q = 30 \times 30$. While there is no performance lower bound for $Q = 30 \times 30$ since $Q > M(M-2)$, it is shown that the KF still improves over recursion.

A key model property, in order for the KF to converge, is the validity of $\mathbf{P}_{k|k}$. For instance, the proposed filter may not work for high resolution images as the observation model $\mathbf{H}_k$ will be to bad conditioned, corrupting the estimation of $\mathbf{P}_{k|k}$ at each step (the number of considered direction of arrivals depends on the number of antennas).

5 Conclusion and perspectives

Autonomous spacecraft navigation relies on the availability of trustworthy measurements from which reference elements are tracked. For instance, radio interferometric images can be used to identify known radio sources from their power, provided that



(a) (b)

Fig. 2: KF performances with a beamforming initialization for different image size: $Q = 22 \times 22$ **(a)** and $Q = 30 \times 30$ **(b)**. The MSE lower bound $\mathrm{MSE}^a\left(\widehat{\mathbf{x}}_{k|k}\right)$ (black circles), MSE predicted by KF $\widehat{\mathrm{MSE}}\left(\widehat{\mathbf{x}}_{k|k}\right)$ (pink boxes), MSE actually achieved $\mathrm{MSE}^p\left(\widehat{\mathbf{x}}_{k|k}\right)$ (blue crosses) of the proposed FK and MSE achieved by the thresholded estimates $\mathrm{MSE}^p\left(\pi_{\mathbb{R}_+^{Q\times 1}}\left(\widehat{\mathbf{x}}_{k|k}\right)\right)$ (green plus signs) are given for $N = 10^3$.

the estimation is accurate enough. In this context, this work proposes a formalization of the Kalman filter for the dynamic estimation of radio source power based on empirical covariance measurements, considering any signal and noise distributions. The proposed filter is misspecified but conserves the structure of the observation noise covariance matrix during the recursion. As such, the state and the filter accuracy are estimated conjointly. It is shown that the observation noise covariance matrix expression involves the multivariate kurtosis of the source signals and noise. An application on simulated data representative of a dynamic radio interferometic imaging framework was presented, highlighting the applicability of the proposed filter. Given that the observation model is well conditioned, one can compute a lower performance bound towards which the predicted and true filter performances converge along iterations. Future work will focus on state-transition model with additive state noise.

Supplementary information. Not applicable

Acknowledgments. Not applicable

Declarations

Fundings

This work was partially supported by the DGA/AID project 2021.65.0070.

Conflict of interest/Competing interests

The authors declare that they have no competing interests.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Code availability

The code used to generate the presented results are available from the corresponding author upon request.

Authors' contributions

C.C. and N.A. designed the algorithm; E.C. is at the origin of the conceptualization; C.C. and E.C. wrote the manuscript and developed the methodology; C.C. completed simulation experiments; N.A., P.L., N.E.K. and I.V. provided suggestions for code and article details.

Appendix A Estimation of real-valued state with complex measurements

Let consider a linear observation model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \quad \mathbf{y}, \mathbf{v} \in \mathbb{C}^N, \quad \mathbf{x} \in \mathbb{R}^P. \tag{A1}$$

A complex, affine estimator of $\mathbf{x}$ writes $\widehat{\mathbf{x}} = \mathbf{K}\mathbf{y} + \mathbf{a}$ with $\widehat{\mathbf{x}} \in \mathbb{C}^P$, $\mathbf{K} \in \mathbb{C}^{P \times N}$, $\mathbf{a} \in \mathbb{C}^P$ and

$$\mathbf{K}\mathbf{y} = (\mathbf{K}_r + j\mathbf{K}_j)(\mathbf{y}_r + j\mathbf{y}_j) = (\mathbf{K}_r\mathbf{y}_r - \mathbf{K}_j\mathbf{y}_j) + j(\mathbf{K}_r\mathbf{y}_j + \mathbf{K}_j\mathbf{y}_r)$$

and thus

$$(\mathbf{K}\mathbf{y} + \mathbf{a}) - \mathbf{x} = (\mathbf{K}_r\mathbf{y}_r - \mathbf{K}_j\mathbf{y}_j + \mathbf{a}_r - \mathbf{x}) + j(\mathbf{K}_r\mathbf{y}_j + \mathbf{K}_j\mathbf{y}_r + \mathbf{a}_j). \tag{A2}$$

The mean square error matrix writes

$$\begin{aligned}
\mathbf{P}(\mathbf{K}) &= E\left[((\mathbf{K}\mathbf{y}+\mathbf{a})-\mathbf{x})((\mathbf{K}\mathbf{y}+\mathbf{a})-\mathbf{x})^H\right] \\
&= E\left[\begin{array}{l}((\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})+j(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)) \\ \times((\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})+j(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j))^H\end{array}\right] \\
&= E\left[\begin{array}{l}((\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})+j(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)) \\ \times\left((\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T-j(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)^T\right)\end{array}\right] \\
&= \begin{array}{l}E\left[(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T\right] \\ +E\left[(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)^T\right] \\ +j\left(\begin{array}{l}E\left[(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T\right] \\ -E\left[(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)^T\right]\end{array}\right)\end{array}
\end{aligned} \tag{A3}$$

Then $\forall \mathbf{w} \in \mathbb{R}^P$

$$\begin{aligned}
\mathbf{w}^T\mathbf{P}(\mathbf{K})\,\mathbf{w} = &\ \mathbf{w}^T E\left[(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T\right]\mathbf{w} \\
&+\mathbf{w}^T E\left[(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)^T\right]\mathbf{w} \\
&+j\Big(\mathbf{w}^T E\left[(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T\right]\mathbf{w} \\
&-\mathbf{w}^T E\left[(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)^T\right]\mathbf{w}\Big) \\
= &E\left[\left(\mathbf{w}^T(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})\right)^2\right]+E\left[\left(\mathbf{w}^T(\mathbf{K}_r\mathbf{y}_j+\mathbf{K}_j\mathbf{y}_r+\mathbf{a}_j)\right)^2\right].
\end{aligned} \tag{A4}$$

On the other hand, one may consider the following observation model:

$$\begin{pmatrix}\mathbf{y}_r\\ \mathbf{y}_j\end{pmatrix} = \begin{bmatrix}\mathbf{H}_r\\ \mathbf{H}_j\end{bmatrix}\mathbf{x}+\begin{pmatrix}\mathbf{v}_r\\ \mathbf{v}_j\end{pmatrix} \tag{A5}$$

for which a real and affine estimator of $\mathbf{x}$ writes $\widehat{\mathbf{x}} = [\mathbf{K}_r\ \mathbf{K}_j]\begin{pmatrix}\mathbf{y}_r\\ \mathbf{y}_j\end{pmatrix}+\mathbf{a}_r = \mathbf{K}_r\mathbf{y}_r+\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r$ $\widehat{\mathbf{x}} \in \mathbb{R}^P$, $\mathbf{K}_r, \mathbf{K}_j \in \mathbb{R}^{P\times N}$, $\mathbf{a}_r \in \mathbb{R}^P$. Its mean square error matrix is

$$\mathbf{P}([\mathbf{K}_r\ \mathbf{K}_j]) = E\left[(\mathbf{K}_r\mathbf{y}_r+\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})(\mathbf{K}_r\mathbf{y}_r+\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})^T\right].$$

Then, $\forall \mathbf{w} \in \mathbb{R}^P$,

$$\mathbf{w}^T\mathbf{P}([\mathbf{K}_r\ -\mathbf{K}_j])\,\mathbf{w} = E\left[\left(\mathbf{w}^T(\mathbf{K}_r\mathbf{y}_r-\mathbf{K}_j\mathbf{y}_j+\mathbf{a}_r-\mathbf{x})\right)^2\right] \le \mathbf{w}^T\mathbf{P}(\mathbf{K}_r+j\mathbf{K}_j)\,\mathbf{w} \tag{A6}$$

which leads to

$$\forall \mathbf{w} \in \mathbb{R}^P,\ \mathbf{w}^T\mathbf{P}\left(\left[\mathbf{K}_r\ \mathbf{K}_j\right]^b\right)\mathbf{w} \leq \mathbf{w}^T\mathbf{P}\left(\left[\mathbf{K}_r^b\ -\mathbf{K}_j^b\right]\right)\mathbf{w} \leq \mathbf{w}^T\mathbf{P}\left(\mathbf{K}^b = \mathbf{K}_r^b + j\mathbf{K}_j^b\right)\mathbf{w}. \tag{A7}$$

Finally, it was proved that the performance of a linear estimator based on the concatenation of real and imaginary parts of a complex valued observation is upper bounded by the performance of an estimator based only on complex measurements. The latter also apply for the concatenation of complex-valued observations and its complex conjugate, as presented in this work. The demonstration derives froms the relation

$$\begin{bmatrix}\mathbf{y}_r\\ \mathbf{y}_j\end{bmatrix} = \frac{1}{2}\begin{bmatrix}\mathbf{I_N} & \mathbf{I_N}\\ -j\mathbf{I_N} & j\mathbf{I_N}\end{bmatrix}\begin{bmatrix}\mathbf{y}\\ \mathbf{y}^*\end{bmatrix}. \tag{A8}$$

Appendix B Covariance of measurement's vectorized SCM

This section provides expression of $\mathbf{C}_{\mathbf{z}_k^*\otimes\mathbf{z}_k}$ in order to implement (15). To proceed, one starts to write

$$\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}} = \mathbf{R}_{\mathbf{z}^*\otimes\mathbf{z}} - \mathbf{m}_{\mathbf{z}^*\otimes\mathbf{z}}\mathbf{m}_{\mathbf{z}^*\otimes\mathbf{z}}^H, \tag{B9}$$

where the subscript k is dropped in order to simplify expressions, with

$$\mathbf{R}_{\mathbf{z}^*\otimes\mathbf{z}} = \mathbb{E}\left[\left(\mathbf{z}^*\otimes\mathbf{z}\right)\left(\mathbf{z}^*\otimes\mathbf{z}\right)^H\right], \tag{B10}$$

and

$$\mathbf{m}_{\mathbf{z}^*\otimes\mathbf{z}} = \mathbb{E}\left[\mathbf{z}^*\otimes\mathbf{z}\right]. \tag{B11}$$

The latter is developed from the general form of $\mathbf{z}$

$$\mathbf{z} = \sum_{q=1}^{Q}\mathbf{a}_q s_q + \mathbf{n} \tag{B12}$$

as

$$\mathbf{m}_{\mathbf{z}^*\otimes\mathbf{z}} = \mathrm{vec}\left(\mathbf{C}_\mathbf{z}\right), \tag{B13}$$

where

$$\begin{aligned}\mathrm{vec}\left(\mathbf{C}_\mathbf{z}\right) &= \mathrm{vec}\left(\sum_{q=1}^{Q} x_q\mathbf{a}_q\mathbf{a}_q^H + \mathrm{vec}\left(\mathbf{C}_\mathbf{n}\right)\right)\\ &= \sum_{q=1}^{Q} x_q\mathbf{a}_q^*\otimes\mathbf{a}_q + \mathrm{vec}\left(\mathbf{C}_\mathbf{n}\right).\end{aligned} \tag{B14}$$

The former is developed from

$$\mathbf{z}^* \otimes \mathbf{z} = \sum_{q=1}^{Q}\sum_{q'=1}^{Q} \mathbf{a}_q^* \otimes \mathbf{a}_{q'} \left(s_q^* s_{q'}\right) + \mathbf{n}^* \otimes \mathbf{n} + \sum_{q'=1}^{Q} \left(s_{q'}\mathbf{n}^*\right) \otimes \mathbf{a}_{q'} + \sum_{q=1}^{Q} \mathbf{a}_q^* \otimes \left(s_q^* \mathbf{n}\right), \quad \text{(B15)}$$

which, given that $\mathbf{s}$, $\mathbf{n}$ are independent and composed of respectively Q and M circular random variables, leads to

$$\begin{aligned}
\mathbf{R}_{\mathbf{z}^* \otimes \mathbf{z}} = & \sum_{l=1}^{Q}\sum_{l'=1}^{Q}\sum_{q=1}^{Q}\sum_{q'=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)\left(\mathbf{a}_l^* \otimes \mathbf{a}_{l'}\right)^H E\left[s_q^* s_l s_{q'} s_{l'}^*\right] + \left(\sum_{q=1}^{Q} \mathbf{a}_q^* \otimes \mathbf{a}_q x_q^2\right) E\left[\mathbf{n}^* \otimes \mathbf{n}\right]^H \\
& + \left(\sum_{q=1}^{Q} x_q^2 E\left[\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)^H\right]\right) + \sum_{q=1}^{Q} x_q^2 E\left[\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)^H\right] \\
& + \sum_{q=1}^{Q} x_q^2 E\left[\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)^H\right] + \sum_{q=1}^{Q} x_q^2 E\left[\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)^H\right] \\
& + E\left[\mathbf{n}^* \otimes \mathbf{n}\right]\left(\sum_{q=1}^{Q} \mathbf{a}_q^* \otimes \mathbf{a}_q x_q^2\right)^H + E\left[\left(\mathbf{n}^* \otimes \mathbf{n}\right)\left(\mathbf{n}^* \otimes \mathbf{n}\right)^H\right]
\end{aligned} \quad \text{(B16)}$$

where $x_q \triangleq \mathbb{E}\left[|s_q|^2\right]$. Since

$$E\left[s_q^* s_l s_{q'} s_{l'}^*\right] = \begin{cases} x_q x_{q'} & \text{, if } q = l \neq q' = l'; \\ x_q x_l & \text{, if } q = q' \neq l = l'; \\ \mathbb{E}\left[|s_q|^4\right] & \text{, if } m = l \neq l' = m'; \\ 0 & \text{, elsewhere,} \end{cases} \quad \text{(B17)}$$

and

$$\begin{aligned}
E\left[\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)^H\right] &= E\left[\mathbf{n}\mathbf{n}^H\right]^* \otimes \mathbf{a}_q \mathbf{a}_q^H \\
&= \mathbf{C}_{\mathbf{n}}^* \otimes \mathbf{a}_q \mathbf{a}_q^H \\
E\left[\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)^H\right] &= E\left[\left(\mathbf{n} \otimes \mathbf{a}_q^*\right)\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)^T\right]^* \\
&= \mathbf{0} \\
E\left[\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)\left(\mathbf{n}^* \otimes \mathbf{a}_q\right)^H\right] &= E\left[\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)\left(\mathbf{n} \otimes \mathbf{a}_q^*\right)^T\right] \\
&= \mathbf{0}
\end{aligned}$$

$$
\begin{aligned}
E\left[\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)\left(\mathbf{a}_q^* \otimes \mathbf{n}\right)^H\right] &= \mathbf{a}_q^* \mathbf{a}_q^T \otimes E\left[\mathbf{n}\mathbf{n}^H\right] \\
&= \left(\mathbf{a}_q \mathbf{a}_q^H\right)^* \otimes \mathbf{C}_\mathbf{n}
\end{aligned}
$$

one has that

$$
\begin{aligned}
\mathbf{R}_{\mathbf{z}^* \otimes \mathbf{z}} = & \sum_{q=1}^{Q} \sum_{q'=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)^H x_q x_{q'} \\
& + \sum_{q=1}^{Q} \sum_{q'=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_{q'}^* \otimes \mathbf{a}_{q'}\right)^H x_{q'} x_q \\
& + \sum_{q=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)^H \left(\mathbb{E}\left[\left(s_q^* s_q\right)^2\right] - 2{x_q}^2\right) \\
& + \left(\sum_{q=1}^{Q} x_q \mathbf{a}_q^* \otimes \mathbf{a}_q\right) \mathrm{vec}\left(\mathbf{C}_\mathbf{n}\right)^H + \left(\left(\sum_{q=1}^{Q} x_q \mathbf{a}_q^* \otimes \mathbf{a}_q\right) \mathrm{vec}\left(\mathbf{C}_\mathbf{n}\right)^H\right)^H \\
& + \mathbf{C}_\mathbf{n}^* \otimes \left(\sum_{q=1}^{Q} x_q \mathbf{a}_q \mathbf{a}_q^H\right) + \left(\sum_{q=1}^{Q} x_q \mathbf{a}_q \mathbf{a}_q^H\right)^* \otimes \mathbf{C}_\mathbf{n} \\
& + \mathbb{E}\left[\left(\mathbf{n}^* \otimes \mathbf{n}\right)\left(\mathbf{n}^* \otimes \mathbf{n}\right)^H\right].
\end{aligned}
\tag{B18}
$$

It results from (B9) that

$$
\begin{aligned}
\mathbf{C}_{\mathbf{z}^* \otimes \mathbf{z}} = & \sum_{q=1}^{Q} \sum_{q'=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)^H x_q x_{q'} \\
& + \sum_{q=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)^H \rho_{s_q} {x_q}^2 \\
& + \mathbf{C}_\mathbf{n}^* \otimes \left(\sum_{q=1}^{Q} x_q^2 \mathbf{a}_q \mathbf{a}_q^H\right) + \left(\sum_{q=1}^{Q} x_q^2 \mathbf{a}_q \mathbf{a}_q^H\right)^* \otimes \mathbf{C}_\mathbf{n} + \mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}}
\end{aligned}
\tag{B19}
$$

where ρ_{s_q} is defined as (18). On the other hand, remarking that

$$\begin{aligned}\mathbf{C}_{\mathbf{z}}^T \otimes \mathbf{C}_{\mathbf{z}} =& \sum_{q=1}^{Q}\sum_{q'=1}^{Q'} x_q x_{q'} \left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_{q'}\right)^H \\ &+ \left(\sum_{q=1}^{Q} x_q \mathbf{a}_q^* \mathbf{a}_q^T\right) \otimes \mathbf{C}_{\mathbf{n}} \\ &+ \mathbf{C}_{\mathbf{n}}^T \otimes \left(\sum_{q=1}^{Q} x_q \mathbf{a}_q \mathbf{a}_q^H\right) + \mathbf{C}_{\mathbf{n}}^T \otimes \mathbf{C}_{\mathbf{n}}\end{aligned} \tag{B20}$$

leads to the final equation:

$$\begin{aligned}\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}} =& \mathbf{C}_{\mathbf{z}}^T \otimes \mathbf{C}_{\mathbf{z}} + \sum_{q=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)^H {\rho_{s_q}}{x_q}^2 \\ &+ \left(\mathbf{C}_{\mathbf{n}^*\otimes\mathbf{n}} - \mathbf{C}_{\mathbf{n}}^T \otimes \mathbf{C}_{\mathbf{n}}\right).\end{aligned} \tag{B21}$$

Appendix C Covariance of a vectorized SCM

Let $\mathbf{n}$ be a vector of centered and independent complex circular random variables n_m for $m = 1, \ldots, M$, then for $\left(\mathbf{e}_1, \ldots, \mathbf{e}_M^2\right)$ being the canonical basis of $\mathbb{R}^{M^2}$, one has that

$$\begin{aligned}\mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{n}^*\otimes\mathbf{n}} \mathbf{e}_{l+M(l'-1)} &= \mathbf{C}_{\mathbf{e}_{m+M(m'-1)}^T(\mathbf{n}^*\otimes\mathbf{n}),\mathbf{e}_{l+M(l'-1)}^T(\mathbf{n}^*\otimes\mathbf{n})} \\ &= \mathbf{C}_{(\mathbf{e}_{m'}\otimes\mathbf{e}_m)^T(\mathbf{n}^*\otimes\mathbf{n}),(\mathbf{e}_{l'}\otimes\mathbf{e}_l)^T(\mathbf{n}^*\otimes\mathbf{n})} \\ &= \mathbf{C}_{\left(\mathbf{e}_{m'}^T\mathbf{n}^*\right)\otimes\left(\mathbf{e}_m^T\mathbf{n}\right),\left(\mathbf{e}_{l'}^T\mathbf{n}^*\right)\otimes\left(\mathbf{e}_l^T\mathbf{n}\right)}\end{aligned} \tag{C22}$$

and thus

$$\mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{n}^*\otimes\mathbf{n}} \mathbf{e}_{l+M(l'-1)} = \mathbb{E}\left[n_{m'}^* n_m n_l^* n_{l'}\right] - \mathbb{E}\left[n_m n_{m'}^*\right]\mathbb{E}\left[n_l^* n_{l'}\right], \tag{C23}$$

where m, m', l, $l' \in \{1, \ldots, M\}$.

On one hand, for $\mathbf{n}$ being a vector of centered and Gaussian complex circular random variables, the fourth-order cumulant

$$\kappa_{i,j,k,l} = E\left[n_i n_j^* n_k n_l^*\right] - E\left[n_i n_j^*\right] E\left[n_k n_l^*\right] - E\left[n_i n_l^*\right] E\left[n_k n_j^*\right] \tag{C24}$$

is null, i.e. $\kappa_{i,j,k,l} = 0$, which implies that

$$E\left[n_i n_j^* n_k n_l^*\right] = E\left[n_i n_j^*\right] E\left[n_k n_l^*\right] + E\left[n_i n_l^*\right] E\left[n_k n_j^*\right] \tag{C25}$$

and then

$$\mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{n}^*\otimes\mathbf{n}} \mathbf{e}_{l+M(l'-1)} = \mathbb{E}\left[n_m n_l^*\right]\mathbb{E}\left[n_{l'} n_{m'}^*\right]$$

$$= (\mathbf{C}_{\mathbf{n}})_{m,l} (\mathbf{C}_{\mathbf{n}})_{l',m'}$$

from (C23). Finally, one has directly that

$$\mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}} = \mathbf{C}_{\mathbf{n}}^T \otimes \mathbf{C}_{\mathbf{n}}. \tag{C26}$$

On the other hand, for $\mathbf{n}$ being a vector of independent and centered complex circular random variables following any distribution, one has from (C23) that

$$\mathbf{e}_{m+M(m'-1)}^T \mathbf{C}_{\mathbf{n}^* \otimes \mathbf{n}} \mathbf{e}_{l+M(l'-1)} = \begin{cases} \mathbb{E}\left[|n_m|^4\right] - \mathbb{E}\left[|n_m|^2\right]^2 & , \text{if } m = m' = l = l'; \\ \mathbb{E}\left[|n_m|^2\right] \mathbb{E}\left[|n_{l'}|^2\right] & , \text{if } m = l \neq l' = m'; \\ 0 & , \text{elsewhere,} \end{cases} \tag{C27}$$

where $\mathbb{E}\left[|n_m|^2\right] = \sigma_{n_m}^2$, $\mathbb{E}\left[|n_{l'}|^2\right] = \sigma_{n_{l'}}^2$ and

$$\mathbb{E}\left[|n_m|^4\right] - \mathbb{E}\left[|n_m|^2\right]^2 = (\rho_{n_m} + 1)\, \sigma_{n_m}^4 \tag{C28}$$

where

$$\rho_{n_m} = \frac{\mathbb{E}\left[|n_m|^4\right]}{\mathbb{E}\left[|n_m|^2\right]^2} - 2. \tag{C29}$$

In particular, $\rho_{n_m} = 0$ for Gaussian random variables, which implies (C26).

In summary, it was proved that (C26) applies for a vector $\mathbf{n}$ of centered and Gaussian complex circular random variables, and that (C27) applies for centered and independent complex circular random variables following any distribution.

Appendix D Computation of the MVDRF

Given a linear observation model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \quad \mathbf{y}, \mathbf{v} \in \mathbb{C}^{N \times 1},\ \mathbf{x} \in \mathbb{C}^{P \times 1},\ \mathbf{H} \in \mathbb{C}^{N \times P}, \tag{D30a}$$

such that

$$\operatorname{rank}(\mathbf{H}) = P \leq N, \quad \operatorname{rank}(\mathbf{C}_{\mathbf{v}}) = R \leq N, \tag{D30b}$$

the general form of a DRF is (denoting $\mathbf{L}^H = \mathbf{K}$)

$$\begin{aligned} \widehat{\mathbf{x}} &= \mathbf{m}_{\mathbf{x}} + \mathbf{L}^H \left(\mathbf{y} - \mathbf{m}_{\mathbf{y}}\right) \\ &= \left(\mathbf{I}_P - \mathbf{L}^H \mathbf{H}\right) \mathbf{m}_{\mathbf{x}} + \mathbf{L}^H \left(\mathbf{y} - \mathbf{m}_{\mathbf{v}}\right) \end{aligned} \tag{D31}$$

with

$$\begin{cases} \widehat{\mathbf{x}} = \mathbf{x} + \mathbf{L}^H \left(\mathbf{v} - \mathbf{m}_{\mathbf{v}} \right); \\ \mathbf{I}_P = \mathbf{L}^H \mathbf{H}; \\ E\left[\left(\widehat{\mathbf{x}} - \mathbf{x} \right) \left(\widehat{\mathbf{x}} - \mathbf{x} \right)^H \right] = \mathbf{L}^H \mathbf{C}_{\mathbf{v}} \mathbf{L}. \end{cases}$$

The MVDRF (provided that it exists) is then the DRF such that

$$\mathbf{L}^b = \arg\min_{\mathbf{L}} \left\{ \mathbf{L}^H \mathbf{C}_{\mathbf{v}} \mathbf{L} \text{ s.t. } \mathbf{L}^H \mathbf{H} = \mathbf{I}_P \right\}. \tag{D32}$$

Considering that nor $\mathbf{H}$ and $\mathbf{C}_{\mathbf{v}}$ are full rank, one denotes

$$\mathbf{H} \overset{\text{compact SVD}}{=} \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H, \tag{D33a}$$

$$\mathbf{C}_{\mathbf{v}} \overset{\text{compact EVD}}{=} \mathbf{W} \mathbf{D} \mathbf{W}^H \tag{D33b}$$

and $\mathbb{U}$, $\mathbb{W}$ such that $[\mathbf{U}\,\mathbb{U}]$, $[\mathbf{W}\,\mathbb{W}]$ are unitary matrices. The considered filter is of the form

$$\mathbf{L} = \mathbf{U}\mathbf{T} + \mathbb{U}\mathbb{T} \tag{D33c}$$

with

$$\mathbf{T} \in \mathbb{C}^{P \times P}, \ \mathbb{T} \in \mathbb{C}^{(N-P) \times P}. \tag{D33d}$$

Thus

$$\mathbf{L}^H \mathbf{H} = \mathbf{T}^H \mathbf{\Sigma} \mathbf{V}^H \tag{D34}$$

from which the distortionless condition is equivalent to

$$\mathbf{T} = \mathbf{\Sigma}^{-1} \mathbf{V}^H. \tag{D35}$$

One has that

$$\begin{aligned} \mathbf{L}^H \mathbf{C}_{\mathbf{v}} \mathbf{L} &= \left(\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbb{U}\mathbb{T} \right)^H \mathbf{C}_{\mathbf{v}} \left(\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbb{U}\mathbb{T} \right) \\ &= \left(\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbb{U}\mathbb{T} \right)^H \mathbf{W}\mathbf{D}\mathbf{W}^H \left(\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbb{U}\mathbb{T} \right) \\ &= \left(\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbf{W}^H\mathbb{U}\mathbb{T} \right)^H \mathbf{D} \left(\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \mathbf{W}^H\mathbb{U}\mathbb{T} \right) \end{aligned} \tag{D36}$$

with

$$\mathbf{W}^H\mathbb{U} \in \mathbb{C}^{R \times (N-P)}, \quad rank\left(\mathbf{W}^H\mathbb{U} \right) = Q \leq \min\left(R, N-P \right). \tag{D37}$$

Then, one applies a last decomposition

$$\mathbf{W}^H\mathbb{U} \overset{\text{compact SVD}}{=} \underline{\mathbf{U}\mathbf{\Sigma}\mathbf{V}}^H. \tag{D38a}$$

Given $\mathbf{g} \in \mathbb{C}^P$, one obtains

$$\mathbf{g}^H \left(\mathbf{L}^H \mathbf{C}_{\mathbf{v}} \mathbf{L} \right) \mathbf{g} = \left\| \left(\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \underline{\mathbf{U}} \left(\underline{\mathbf{\Sigma}\mathbf{V}}^H\mathbb{T} \right) \right) \mathbf{g} \right\|_{\mathbf{D}}^2 \tag{D38b}$$

where $\|\mathbf{z}\|_{\mathbf{D}}^2 = \mathbf{z}^H\mathbf{D}\mathbf{z}$. Considering the oblique projector

$$\mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}} = \underline{\mathbf{U}}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D} \in \mathbb{C}^{R\times R}, \tag{D38c}$$

then

$$\begin{aligned}\mathbf{g}^H\left(\mathbf{L}^H\mathbf{C}_{\mathbf{v}}\mathbf{L}\right)\mathbf{g} &= \left\|\mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\left(\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \underline{\mathbf{U}}\left(\underline{\mathbf{\Sigma}\mathbf{V}}^H\mathbb{T}\right)\right)\mathbf{g}\right\|_{\mathbf{D}}^2 \\ &\quad + \left\|\left(\mathbf{I}_R - \mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\right)\left(\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \underline{\mathbf{U}}\left(\underline{\mathbf{\Sigma}\mathbf{V}}^H\mathbb{T}\right)\right)\mathbf{g}\right\|_{\mathbf{D}}^2 \\ &= \left\|\underline{\mathbf{U}}\left(\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \underline{\mathbf{\Sigma}\mathbf{V}}^H\mathbb{T}\right)\mathbf{g}\right\|_{\mathbf{D}}^2 \\ &\quad + \left\|\left(\mathbf{I}_R - \mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\right)\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H\mathbf{g}\right\|_{\mathbf{D}}^2 \end{aligned} \tag{D38d}$$

from which one can conclude

$$\mathbf{g}^H\left(\mathbf{L}^H\mathbf{C}_{\mathbf{v}}\mathbf{L}\right)\mathbf{g} \geq \left\|\left(\mathbf{I}_R - \mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\right)\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H\mathbf{g}\right\|_{\mathbf{D}}^2,$$

i.e.

$$\mathbf{L}^H\mathbf{C}_{\mathbf{v}}\mathbf{L} \geq \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{W}\left(\mathbf{I}_R - \mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\right)^H\mathbf{D}\left(\mathbf{I}_R - \mathbf{\Pi}_{\underline{\mathbf{U}}}^{\mathbf{D}}\right)\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H. \tag{D38e}$$

The equality holds for

$$\begin{aligned}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H + \underline{\mathbf{\Sigma}\mathbf{V}}^H\mathbb{T} &= \mathbf{0} \\ &\Updownarrow \\ -\underline{\mathbf{\Sigma}}^{-1}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H &= \underline{\mathbf{V}}^H\mathbb{T}. \end{aligned} \tag{D38f}$$

Let $\underline{\mathbb{V}} \in \mathbb{C}^{(N-P)\times((N-P)-Q)}$ be such that $[\underline{\mathbf{V}}\ \underline{\mathbb{V}}]$ is a unitary matrix, then since $\mathbb{T} \in \mathbb{C}^{(N-P)\times P}$,

$$\mathbb{T} = \left(\underline{\mathbf{V}\mathbf{V}}^H + \underline{\mathbb{V}\mathbb{V}}^H\right)\mathbb{T} = [\underline{\mathbf{V}}\ \underline{\mathbb{V}}]\begin{bmatrix}\underline{\mathbf{V}}^H\mathbb{T} \\ \underline{\mathbb{V}}^H\mathbb{T}\end{bmatrix},$$

which leads to

$$\mathbb{T} = -\underline{\mathbf{V}}\left(\underline{\mathbf{\Sigma}}^{-1}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H\right) + \underline{\mathbb{V}\mathbb{A}}, \tag{D38g}$$

where

$$\underline{\mathbb{A}} \triangleq \underline{\mathbb{V}}^H\mathbb{T} \in \mathbb{C}^{((N-P)-Q)\times P} \tag{D38h}$$

and thus

$$\mathbf{L} = \mathbf{U}\left(\mathbf{\Sigma}^{-1}\mathbf{V}^H\right) - \mathbb{U}\underline{\mathbf{V}}\left(\underline{\mathbf{\Sigma}}^{-1}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H\right) + \mathbb{U}\underline{\mathbb{V}\mathbb{A}}. \tag{D38i}$$

One deduces that there is an infinity of solutions which minimize $\mathbf{L}^H\mathbf{C}_{\mathbf{v}}\mathbf{L}$ (D38i) : since there is no constraint on $\underline{\mathbb{A}}$, the solution has $((N-P)-Q)\times P$ degrees of freedom.

One may consider the solution with minimal Frobenius norm

$$\begin{aligned} tr\left(\mathbf{L}^H\mathbf{L}\right) &= tr\left(\left(\mathbf{UB}-\mathbb{U}\underline{\mathbf{V}}\mathbf{A}+\mathbb{U}\underline{\mathbb{V}}\mathbb{A}\right)^H\left(\mathbf{UB}-\mathbb{U}\underline{\mathbf{V}}\mathbf{A}+\mathbb{U}\underline{\mathbb{V}}\mathbb{A}\right)\right) \\ &= tr\left(\mathbf{B}^H\mathbf{B}\right)+tr\left(\mathbf{A}^H\mathbf{A}\right)+tr\left(\mathbb{A}^H\mathbb{A}\right) \end{aligned}$$

i.e.

$$\mathbf{L}=\mathbf{U}\left(\mathbf{\Sigma}^{-1}\mathbf{V}^H\right)-\mathbb{U}\underline{\mathbf{V}}\underline{\mathbf{\Sigma}}^{-1}\left(\underline{\mathbf{U}}^H\mathbf{D}\underline{\mathbf{U}}\right)^{-1}\underline{\mathbf{U}}^H\mathbf{D}\mathbf{W}^H\mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^H, \tag{D39a}$$

by setting $\underline{\mathbb{A}}=\mathbf{0}$ (i.e. no unnecessarily additional assumption is made).

Appendix E Upper bound of the measurement noise cocariance

This section proves that for a given estimation $\widehat{\mathbf{x}}_{k|k}$ verifying $\left(\widehat{\mathbf{x}}_{k|k}\right)_q \geqslant \left(\mathbf{x}_k\right)_q \,\forall q \in \{1,\ldots,Q\}$, one have that $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\widehat{\mathbf{x}}_{k|k}) \geqslant \widehat{\mathbf{C}}_{\mathbf{v}_k}(\mathbf{x}_k)$, where $\widehat{\mathbf{C}}_{\mathbf{v}_k}(\mathbf{x}_k) \triangleq \mathbf{C}_{\mathbf{v}_k}$.

Let $\mathbf{x}$, $\mathbf{x}' \in \mathbb{R}_+^Q$ be such that $x'_q \geqslant x_q \,\forall q \in \{1,\ldots,Q\}$, where the subscript k is dropped in order to simplify formulas. Since

$$\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x})=\sum_{q=1}^{Q} x_q\mathbf{a}_q\mathbf{a}_q^H+\mathbf{C}_{\mathbf{n}}, \tag{E40}$$

one has that

$$\begin{aligned} \widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')^T\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')-\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x})\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}) = &\sum_{q=1}^{Q}\sum_{l=1}^{Q}\left(x'_qx'_l-x_qx_l\right)\left(\mathbf{a}_q\mathbf{a}_q^H\right)^T\otimes\left(\mathbf{a}_l\mathbf{a}_l^H\right) \\ &+\sum_{q=1}^{Q}\left(x'_q-x_q\right)\left(\mathbf{a}_q\mathbf{a}_q^H\right)^T\otimes\mathbf{C}_{\mathbf{n}} \\ &+\sum_{q=1}^{Q}\left(x'_q-x_q\right)\mathbf{C}_{\mathbf{n}}^T\otimes\left(\mathbf{a}_q\mathbf{a}_q^H\right), \end{aligned} \tag{E41}$$

with $\left(\mathbf{a}_q\mathbf{a}_q^H\right)^T\otimes\left(\mathbf{a}_l\mathbf{a}_l^H\right)$, $\left(\mathbf{a}_q\mathbf{a}_q^H\right)^T\otimes\mathbf{C}_{\mathbf{n}}$ and $\mathbf{C}_{\mathbf{n}}^T\otimes\left(\mathbf{a}_q\mathbf{a}_q^H\right)$ being three hermitian and positive matrices. Hence, $\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')^T\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')-\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x})\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x})$ is a linear combination of positive hermitian matrices with positive coefficients, and thus

$$\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')^T\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}')\geqslant\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x})\otimes\widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}). \tag{E42}$$

On the other hand, from (17), one has that

$$\begin{aligned}\widehat{\mathbf{C}}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}') - \widehat{\mathbf{C}}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) = &\widehat{\mathbf{C}}_{\mathbf{z}}^T(\mathbf{x}') \otimes \widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}') - \widehat{\mathbf{C}}_{\mathbf{z}}^T(\mathbf{x}) \otimes \widehat{\mathbf{C}}_{\mathbf{z}}(\mathbf{x}) \\ &+ \sum_{q=1}^{Q} \left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)\left(\mathbf{a}_q^* \otimes \mathbf{a}_q\right)^H \rho_{s_q}\left({x'_q}^2 - {x_q}^2\right)\end{aligned} \tag{E43}$$

which (from the same argument) yields to

$$\widehat{\mathbf{C}}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}') \geqslant \widehat{\mathbf{C}}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}). \tag{E44}$$

Finally, remarking that from (16), $\mathbf{C}_{\mathbf{v}}(\mathbf{x}') - \mathbf{C}_{\mathbf{v}}(\mathbf{x})$ is the sum of

$$\frac{1}{N}\begin{bmatrix} \mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}') - \mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}}(\mathbf{x}') - \mathbf{C}^*_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) \end{bmatrix} \tag{E45}$$

and

$$\frac{1}{N}\begin{bmatrix} \mathbf{0} & \mathbf{C}_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}') - \mathbf{C}_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}) \\ \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}') - \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}) & \mathbf{0} \end{bmatrix} \tag{E46}$$

which are positive hermitian matrices since

$$\begin{aligned}\mathbf{C}^H_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}) &= \mathbf{P}^T\mathbf{C}^H_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) \\ &= \mathbf{P}^T\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) \\ &= \mathbf{C}_{\mathbf{z}\otimes\mathbf{z}^*,\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x}) \\ &= \mathbf{C}_{\mathbf{z}\otimes\mathbf{z}^*}(\mathbf{x})\mathbf{P} \\ &= \left(\mathbf{C}_{\mathbf{z}^*\otimes\mathbf{z}}(\mathbf{x})\mathbf{P}\right)^* \\ &= \mathbf{C}^*_{\mathbf{z}_k^*\otimes\mathbf{z}_k,\mathbf{z}_k\otimes\mathbf{z}_k^*}(\mathbf{x}),\end{aligned} \tag{E47}$$

one can conclude that

$$\mathbf{C}_{\mathbf{v}}(\mathbf{x}') \geqslant \mathbf{C}_{\mathbf{v}}(\mathbf{x}). \tag{E48}$$

References

[1] Henry, S., Christian, J.A.: Absolute triangulation algorithms for space exploration. Journal of Guidance, Control, and Dynamics **46**(1), 21–46 (2023) https://doi.org/10.2514/1.g006989

[2] Poirot, J.L., McWilliams, G.V.: Navigation by back triangulation. IEEE Transactions on Aerospace and Electronic **12**(2), 270–274 (1976) https://doi.org/10.1109/TAES.1976.308304

[3] Driedger, M., Ferguson, P.: Feasibility study of an orbital navigation filter using resident space object observations. Journal of Guidance, Control, and Dynamics **44**(3), 622–628 (2021) https://doi.org/10.2514/1.G005210

[4] Veen, A.-J., Wijnholds, S.J., Sardarabadi, A.M.: In: Bhattacharyya, S.S., Deprettere, E.F., Leupers, R., Takala, J. (eds.) Signal Processing for Radio Astronomy, pp. 311–360. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-91734-4_9 . https://doi.org/10.1007/978-3-319-91734-4_9

[5] Crassidis, J.L., Junkins, J.L.: Optimal Estimation of Dynamic Systems (2nd Ed.). Chapman and Hall/CRC, (2011). https://doi.org/10.1201/b11154

[6] Diniz, P.S.R.: Adaptive Filtering: Algorithms and Practical Implementation (4 Ed.). Springer, (2013)

[7] Groves, P.D.: Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, Second Edition. Artech House, (2013)

[8] Arasaratnam, I., Haykin, S.: Cubature kalman filters. IEEE Transactions on Automatic Control **54**(6), 1254–1269 (2009) https://doi.org/10.1109/TAC.2009.2019800

[9] Chaumette, E., Vilà-Valls, J., Vincent, F.: On the general conditions of existence for linear mmse filters: Wiener and kalman. Signal Processing **184**, 108052 (2021) https://doi.org/10.1016/j.sigpro.2021.108052

[10] Vilà-Valls, J., Chaumette, E., Vincent, F., Closas, P.: Robust linearly constrained kalman filter for general mismatched linear state-space models. IEEE Transactions on Automatic Control **67**(12), 6794–6801 (2022) https://doi.org/10.1109/TAC.2021.3132890

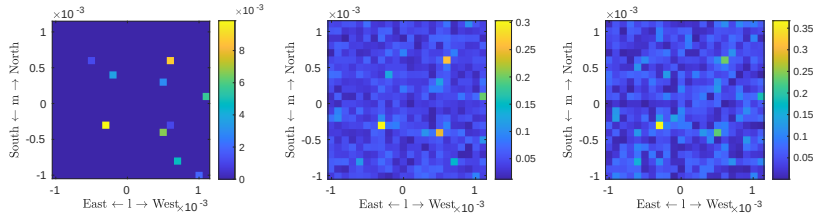
[11] MARDIA, K.V.: Measures of multivariate skewness and kurtosis with applications. Biometrika **57**(3), 519–530 (1970) https://doi.org/10.1093/biomet/57.3.519 https://academic.oup.com/biomet/article-pdf/57/3/519/702615/57-3-519.pdf

[12] Vanderveen, M.C., Ng, B.C., Papadias, C.B., Paulraj, A.: Joint angle and delay estimation (jade) for signals in multipath environments. In: Conference Record of The Thirtieth Asilomar Conference on Signals, Systems and Computers, pp. 1250–12542 (1996). https://doi.org/10.1109/ACSSC.1996.599145

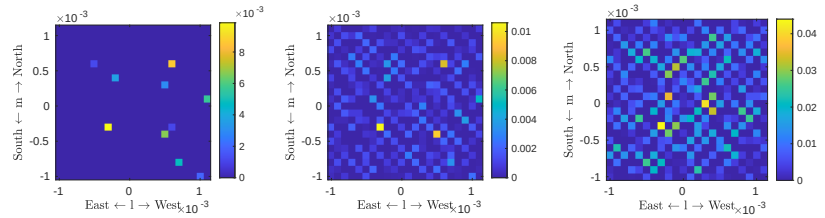
[13] Anderson, B.D.O., Moore, J.B.: Optimal Filtering. Information and system sciences series. Prentice-Hall, (1979)

[14] Van Trees, H.L.: Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory. John Wiley & Sons, (2002). https://doi.org/10.1002/0471221104.ch2

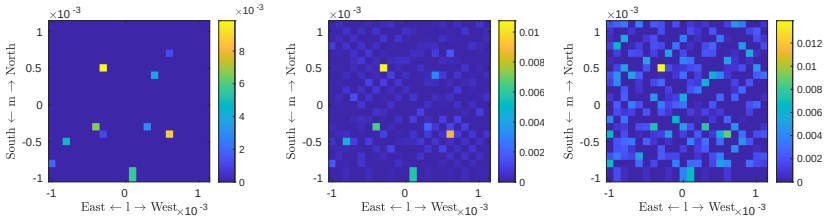
[15] Napier, P.J., Thompson, A.R., Ekers, R.D.: The very large array: Design and performance of a modern synthesis radio telescope. Proceedings of the IEEE **71**(11), 1295–1320 (1983) https://doi.org/10.1109/PROC.1983.12765



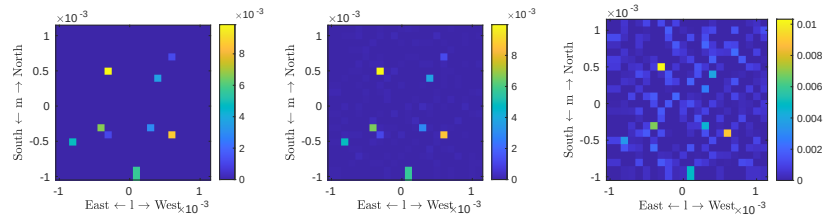
(a) $k = 0$ (beamforming)



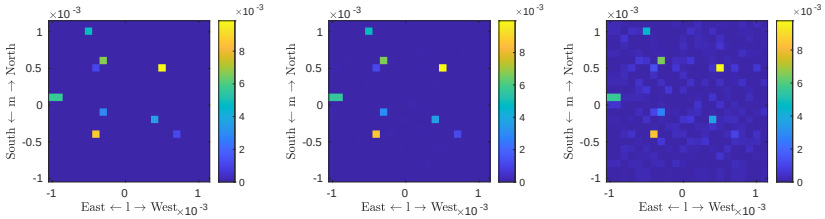
(b) $k = 0$ (MVDRF)



(c) $k = 1$

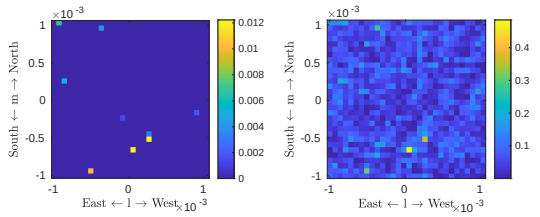


(d) $k = 5$

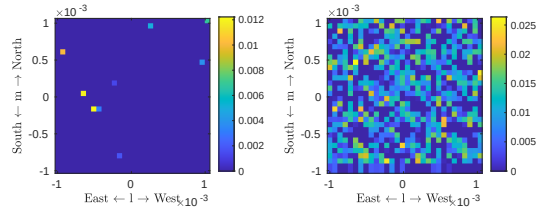


(e) $k = 30$

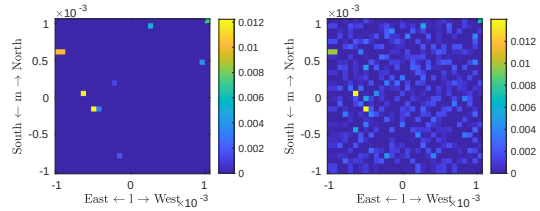
Fig. 3: Actual images (left) and KF estimates $\pi_{\mathbb{R}_+^Q}\left(\widehat{\mathbf{x}}_{k|k}\right)$ (right) at $k = 0, 1, 5$ and 30 (**(b)**, **(c)**, **(d)** and **(e)** respectively) for $N = 10^5$ (middle) and $N = 10^3$ (right) samples. The beamforming estimation **(a)** is used to initialize $\widehat{\mathbf{C}}_{\mathbf{v}_0}$ for the MVDRF estimation **(b)**.



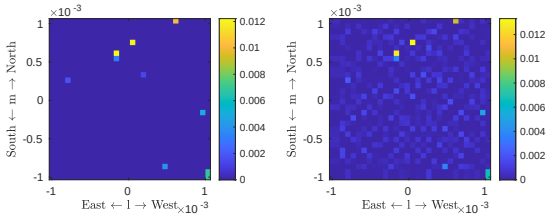
(a) $k = 0$ (beamforming)



(b) $k = 1$



(c) $k = 5$



(d) $k = 30$

Fig. 4: Actual images (left) and KF estimates $\pi_{\mathbb{R}_+^Q}\left(\widehat{\mathbf{x}}_{k|k}\right)$ (right) at $k = 0, 1, 5$ and 30 (**(a)**, **(b)**, **(c)** and **(d)** respectively) for $N = 10^3$ samples.