Dynamic radioastronomic image reconstruction using Kalman filter under ionospheric perturbations

G. Robert-Dautun

Contents

Part I

Introduction

When observing distant objects, instruments resolution matters the most. When using optical telescopes, the determining factor for resolution is the size of the primary mirror. When observing really distant objects, the required size to have a decent resolution becomes unrealistically big. To address this problem, radio interferometry replaces a single optical instrument with multiple antennas collecting electromagnetic waves. When gathering these measures together, we can obtain a better angular resolution than any real single optical instrument could get.

Those distant objects are not fixed in time and a lot of calculations have been done to create evolution models for them. The goal here is not to use observations to confirm models, but to assume a model is known and get better results considering the knowledge of both observations and model.

A (non extended) Kalman filter is a linear quadratic estimator, using observations, state-transition model and knowledge about noise and errors to compute a more accurate result than observations alone. Radio astronomical measures being prone to noise and errors, usage of this filtering method seems *a priori* a good way to improve accuracy of reconstruction.

A sensitive part of the filter is the observations model, needing precise knowledge of antennas positions and directions. An objective of this work is to determine how inaccuracies in these measurements affect image reconstruction, and if it is realistically usable in radio astronomical imagery.

Part II
SATIE

SATIE stands for *laboratoire des Systèmes et Applications des Technologies de l'Information et de l'Energie* (Information and Energy Technology Systems and Applications Laboratory).
Created in 1975 under the name *Laboratoire d'Electricité de l'ENSET*, then renamed *LESiR* in 1984 at the time of its association with CNRS, then finally SATIE in 2002, the laboratory covers a large field of study including signal processing, electronics, and energy conversion.

The SATIE is linked to the following academic structures : ENS Paris-Saclay, ENS Rennes, Université Paris-Saclay, Cergy Paris Université, Université Gustave Eiffel, CNAM Paris and CNRS. It counts five geographical sites near the aforementioned structures.
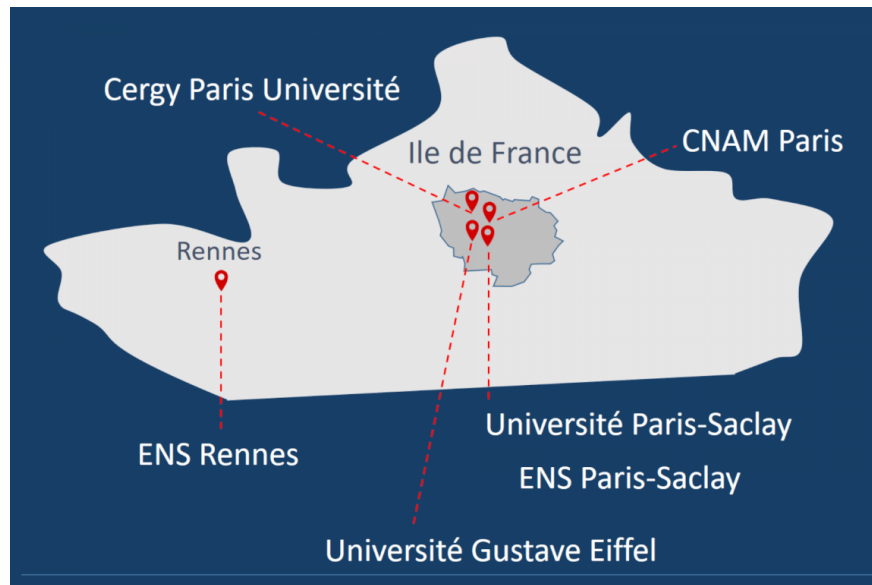

Figure 1: Geographic sites of SATIE (copyright E. Vourc'h)

Part III
Notations and problem

1   Notations

1.1   General notations

$E[\bullet]$ denotes the expected value of a random variable/vector/matrix.

Bold symbols represent vectors, normal ones represent scalars except when stated otherwise.

Operator $\mathbf{m_{\bullet}}$ represents the mean value/vector/matrix of the given $\bullet$ set.

When a 2-D matrix element is indexed with a single coordinate, it designates the reshaped matrix in a 1-D vector.

$\mathbb{I}$ is the identity matrix.

1.2   Hadamard product

Hadamard product, denoted $\odot$, is the element-wise product of same sized matrices :

$$\forall (A,B) \in \left(\mathbb{C}^{m\times n}\right)^2, \quad A\odot B \in \mathbb{C}^{m\times n} \quad \text{and} \quad \forall (i,j) \in [\![1;m]\!] \times [\![1;n]\!], (A\odot B)_{i,j} = A_{i,j} \times B_{i,j}$$

We can further introduce Hadamard inversion :

$$\forall A \in (\mathbb{C}\backslash\{0\})^{m\times n}, \quad A^{\circ-1} \in (\mathbb{C}\backslash\{0\})^{m\times n} \quad \text{and} \quad \forall (i,j) \in [\![1;m]\!] \times [\![1;n]\!], \left(A^{\circ-1}\right)_{i,j} = \left(A_{i,j}\right)^{-1}$$

And notation $\oslash$ being Hadamard division :

$$\forall (A,B) \in \mathbb{C}^{m\times n} \times (\mathbb{C}\backslash 0)^{m\times n}, \quad A\oslash B \in \mathbb{C}^{m\times n} \quad \text{and} \quad \forall (i,j) \in [\![1;m]\!] \times [\![1;n]\!], (A\oslash B)_{i,j} = \frac{A_{i,j}}{B_{i,j}}$$

1.3   Norms and distances

In order to compute distances between matrices, we note $||\bullet||_F$ Frobenius norm :

$$A \in \mathcal{M}_{m,n}(\mathbb{K}) \quad ||A||_F := \sqrt{\text{tr}\{AA^H\}} = \sqrt{\sum_{\substack{1\le i\le m \\ 1\le j\le n}} \left|A_{ij}\right|^2}$$

Inducing the distance d_1 :

$$\forall (A,B) \in \left(\mathcal{M}_{m,n}(\mathbb{K})\right)^2 \quad d_1(A,B) = \frac{||A-B||_F}{mn}$$

In order to measure influence of a specific element, the additive distance d_2 is the following :

$$\forall (A,B) \in \left(\mathcal{M}_{m,n}(\mathbb{K})\right)^2 \quad d_2(A,B) = \sum_{\substack{1\le i\le m \\ 1\le j\le n}} \left|A_{ij} - B_{i,j}\right|$$

2   Measure and parameters

In order to measure precisely the effects, the following parameters will be used :

- The antennas distribution is a 8.57m square 6×6 grid. The number of antennas is thus $J = 36$

- The observed directions grid is a 10×10 square. The number of directions is thus $D = 100$

- A normally distributed observation noise will be introduced with Signal to Noise Ratio (SNR) of 10.

- The first image used for Kalman will be computed using Minimum Variance Distortionless Response (MVDR) method [1], deconvoluted with Maximum Entropy Method (MEM) [6]

- Measured wavelength is $\lambda = 0.3$m

The following matrices and sets will be used to describe this setup :

- $\boldsymbol{z}$ represents the positions of each antenna in the 2-D grid

- $\Delta \boldsymbol{z}_{ij}$ is the baseline : $\Delta \boldsymbol{z}_{ij} = \boldsymbol{z}_i - \boldsymbol{z}_j$.

- l is the directions matrix, represented as normalised direction cosines.

3   Evolution model and Kalman filtering

We consider a set $\{\boldsymbol{x}_k\}$ of n vector as theoretical images taken at a regular time interval. Those images are measured by an antenna mesh $\{\boldsymbol{z}_k\}$ which UV plane will be written H, more generally written as observation model. Measurements vectors will be designated by n measures $\{\boldsymbol{y}_k\}$.
H can be computed by the following :

$$\mathsf{H}_{j,q} = \exp\left(-i\frac{2\pi}{\lambda}\Delta \boldsymbol{z}_j \cdot \mathsf{l}_q\right) \tag{1}$$

We will consider the following model to describe time evolution and observation equation :

$$\begin{cases} \boldsymbol{x}_k = \mathsf{A}\boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1} \\ \boldsymbol{y}_k = \mathsf{H}\boldsymbol{x}_k + \boldsymbol{v}_k \end{cases}$$

Where :

- A is the linear state-transition model between time $k-1$ and k, assuming evolution is time invariant

- $\{\boldsymbol{w}_k\}$ is the state noise sequence

- $\{\boldsymbol{v}_k\}$ is the measurements noise sequence

In order to build images dynamically with better precision, a Kalman filter has been implemented.

This filter requires more inputs to estimate the state of the system :

- $\{\mathbf{Q}_k\}$, the covariance of the process noise sequence

- $\{\mathbf{R}_k\}$, the covariance of the observation noise sequence

Which can be rewritten as

$$\begin{cases} \mathbf{Q}_k = \mathrm{E}\left[\boldsymbol{w}_{k-1}\boldsymbol{w}_{k-1}^{\mathrm{H}}\right] - \mathrm{E}\left[\boldsymbol{w}_{k-1}\right]\mathrm{E}\left[\boldsymbol{w}_{k-1}\right]^{\mathrm{H}} \\ \mathbf{R}_k = \mathrm{E}\left[\boldsymbol{v}_k\boldsymbol{v}_k^{\mathrm{H}}\right] - \mathrm{E}\left[\boldsymbol{v}_k\right]\mathrm{E}\left[\boldsymbol{v}_k\right]^{\mathrm{H}} \end{cases}$$

Those inputs are not *a priori* known, but it has been shown that using the *autocovariance least-squares* (ALS) method provides a good estimate of those matrices. [CITATION NEEDED]
We will assume that these inputs are known perfectly except when specified otherwise.

Kalman filtering works in 2 steps :

Prediction : This step predicts the state of the system at next time-step through de transition-state model, and the predicted measurement :

$$\begin{cases} \widehat{\boldsymbol{x}}_{k|k-1} = A\widehat{\boldsymbol{x}}_{k-1|k-1} + \mathbf{m}_{\boldsymbol{w}_k} \\ \widehat{\boldsymbol{y}}_{k|k-1} = H\widehat{\boldsymbol{x}}_{k|k-1} + \mathbf{m}_{\boldsymbol{v}_k} \end{cases}$$

And the covariance matrix :

$$\mathbf{P}_{k|k-1} = \mathrm{E}\left[(\widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_k)(\widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_k)^{\mathrm{H}}\right] = \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^{\mathrm{H}} + \mathbf{Q}_k$$

This result can be computed using least-square estimation of the prediction error covariance[2]

Update : Using prior knowledge about noise and observation model, the main goal is to build an image from both prediction and measurement using maximum likelihood estimation in order to enhance the resulting image. This can be written as looking for K_k, known as *Kalman gain*, such as :

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \mathrm{K}_k\left(\boldsymbol{y}_k - \widehat{\boldsymbol{y}}_{k|k-1}\right) \tag{2}$$

minimises the quadratic error :

$$\mathrm{K}_k = \arg\min_{\mathrm{K}} \mathbf{P}_{k|k} = \arg\min_{\mathrm{K}} \mathrm{E}\left[(\widehat{\boldsymbol{x}}_{k|k} - \boldsymbol{x}_k)(\widehat{\boldsymbol{x}}_{k|k} - \boldsymbol{x}_k)^{\mathrm{H}}\right] \tag{3}$$

The resulting expression for K_k is [2] :

$$\mathrm{K}_k = \mathbf{P}_{k|k-1}\mathrm{H}^{\mathrm{H}}\left(\mathrm{H}\mathbf{P}_{k|k-1}\mathrm{H}^{\mathrm{H}} + \mathbf{R}_k\right)^{-1} \tag{4}$$

Thus the update phase can be written :

$$\begin{cases} \widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \mathrm{K}_k\left(\boldsymbol{y}_k - \widehat{\boldsymbol{y}}_{k|k-1}\right) \\ \mathbf{P}_{k|k} = (\mathbb{I} - \mathrm{K}_k\mathrm{H})\mathbf{P}_{k|k-1} \end{cases} \tag{5}$$

4 Data simulation and Image reconstruction

4.1 Simulation

In order to simulate data, an image is introduced as the "reference image", this image will only be used in order to compute simulated observations. The selected picture must respect the following conditions :

- mostly black background, to give realistic results for space imagery

- no convolution effect, as all directions are considered sources of different intensities, only real sources must appear

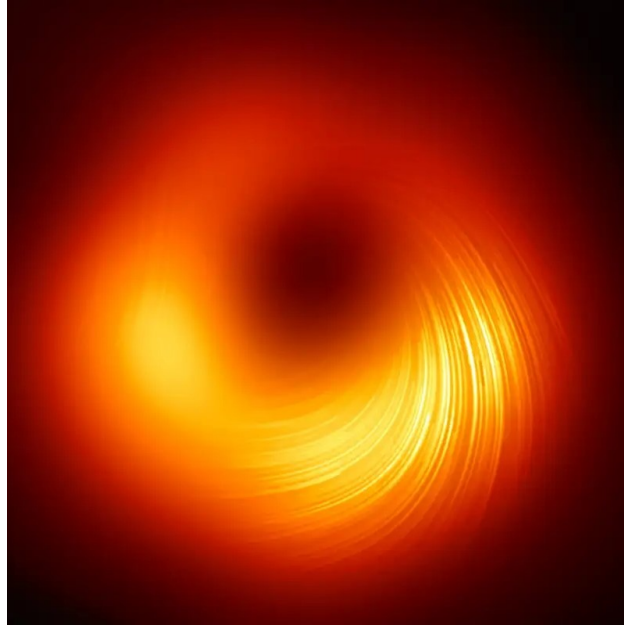The following picture is the one used :

Figure 2: Selected image for simulation

Different images gave similar results if they followed the conditions cited above.

Data simulation will be done as stated in [3]. First the visibilities matrix is computed using the reference image and the UV-plane of the antenna mesh. To simulate visibilities at different snapshots, we apply the wanted transformation without noise and use the same UV-plane matrix to compute visibilities.

To add noise, a normally distributed random matrix is generated and the covariance of that matrix is added to the visibilities one. This matrix could also be directly generated using Wishart distribution.

4.2   Reconstruction

To obtain the initial image x_0 two methods are used depending on the goal of the measure.

When reviewing performances under certain errors, all other inputs of the filter - including x_0 - must be errorless. The simplest way of doing that for x_0 is to use the reference image as input.

When reviewing Kalman filter performances under real conditions, the reference image is not considered as knwon, and we need to reconstruct the image from the visibilities matrix. While different methods exist [3], the ones used here are MVDR reconstructions followed by Maximum Entropy deconvolution.

Part IV

Reliability of Kalman Filter under perturbations

1   Perturbations sources

In radio astronomy, many factors can and will affect measures. Those can be :

Initialisation error   : Kalman filtering requires initializing the first image and the first correlation matrix. Those cannot be exact and will present errors.
Concerning the initial image, multiple precise and efficient methods exist [1], and from this initial image the initial correlation matrix can ben initialised[2] with precision as

$$\mathbf{P}_0 = \mathrm{E}\left[(\boldsymbol{x}_0)(\boldsymbol{x}_0)^{\mathrm{H}}\right] - \mathbf{m}_{\boldsymbol{x}_0}\mathbf{m}_{\boldsymbol{x}_0}^{\mathrm{H}} \tag{6}$$

Error covariance matrices   : as mentioned above, these matrices set can be estimated with good accuracy using ALS algorithms [CITATION NEEDED]. It is to be noted that errors on these matrices will not induce new errors in the estimated images, as they are only used to compute the Kalman gain. This implies that inaccurate error matrices can only spread already existing errors on estimated images.

Antennas informations   : Multiple equations in Kalman filtering relies on the observation model $\mathbf{H}$, which is computed as the UV plane FFT covered by the antenna mesh. The limited knowledge of this mesh can create errors in $\mathbf{H}$, particularly when antennas positions and directions cannot be perfectly measured.

Ionospheric perturbations   : Radio astronomical observations can be affected by the passage of the observed electromagnetic waves through the ionosphere [7]. This can affect the observation model $\mathbf{H}$ as it modifies the perceived direction of the electromagnetic wave measured, and calibration or correction must be made to compensate these perturbations.

2   Effects of perturbations

These perturbations create errors on the matrices used in reconstruction and prediction with Kalman filtering. The goal of this part is to evaluate how much these error affect prediction and reconstruction, to determine if Kalman filtering can be realistically used in radio-astronomic context.

2.1   Methodology and Parameters

First the basic parameters and matrices are computed : antenna grid, directions cosines, baseline vectors.
Next, perfect base matrices are computed : errorless observation matrix $\mathbf{H}$, reference observations using the errorless $\mathbf{H}$, error covariance matrices. The latter are computed perfectly using known simulated noise, and errors from ALS estimation are neglected. A set of perfect images are also computed to measure reconstruction absolute error. Initial image $\boldsymbol{x}_0$ and covariance $\mathbf{P}_0$ are computed as stated in equation (6).

The next step is the generation of an inaccurate matrix normally distributed around the true one - position or direction here. From this matrix is generated an inaccurate $\mathbf{H}^*$ observations matrix if needed, and Kalman filter is run two times, the first using perfect parameters and the second one using inaccurate ones. In order to smooth out random noise effects this step is repeated 100 times.

We can then get the mean error values, measured using a distance defined above, and standard deviations needed.

2.2   Position

The first error studied is the position error. A normally distributed error $\boldsymbol{\varepsilon}$ is introduced on the antennas positions matrix, giving :

$$z_k^* = z_k + \boldsymbol{\varepsilon}_k$$

The new inaccurate baseline is :

$$\Delta z_{ij}^* = z_i + \boldsymbol{\varepsilon}_i - z_j - \boldsymbol{\varepsilon}_j = \Delta z_{ij} + \boldsymbol{\varepsilon}_{ij}$$

The new inaccurate observations matrix's expression is the following :

$$\mathbf{H}_{j,q}^* = \exp\left(-i\frac{2\pi}{\lambda}\Delta z_j^* \cdot \mathbf{I}_q\right) = \exp\left(-i\frac{2\pi}{\lambda}\left(\Delta z_j + \boldsymbol{\varepsilon}_j\right) \cdot \mathbf{I}_q\right) = \exp\left(-i\frac{2\pi}{\lambda}\Delta z_j \cdot \mathbf{I}_q\right)\exp\left(-i\frac{2\pi}{\lambda}\boldsymbol{\varepsilon}_j \cdot \mathbf{I}_q\right) \tag{7}$$

The distortion matrix $\mathscr{C}$ can thus be defined such as :

$$\mathbf{H}^* = \mathbf{H} \odot \mathscr{C} \tag{8}$$

By the same calculation than before [2], position error being decorrelated with other errors or noise, the equations become :

$$\widehat{x}_{k|k-1}^* = \mathbf{A}\widehat{x}_{k-1|k-1} + \mathbf{m}_{w_k} \tag{9a}$$

$$\widehat{y}_{k|k-1}^* = \mathbf{H}^*\widehat{x}_{k|k-1}^* + \mathbf{m}_{v_k} \tag{9b}$$

$$\mathbf{P}_{k|k-1}^* = \mathbf{A}\mathbf{P}_{k-1|k-1}^*\mathbf{A}^{\mathrm{H}} + \mathbf{Q}_k \tag{9c}$$

$$\mathbf{K}_k^* = \mathbf{P}_{k|k-1}^* \left(\mathbf{H}^*\right)^{\mathrm{H}}\left(\mathbf{H}^*\mathbf{P}_{k|k-1}^*\left(\mathbf{H}^*\right)^{\mathrm{H}}\right)^{-1} \tag{9d}$$

$$\mathbf{P}_{k|k}^* = \mathbf{P}_{k|k-1}^* + \mathbf{K}_k^*\mathbf{H}^*\mathbf{P}_{k|k-1}^*\left(\mathbf{H}^*\right)^{\mathrm{H}}\left(\mathbf{K}_k^*\right)^{\mathrm{H}} + \mathbf{K}_k^*\mathbf{R}_k\left(\mathbf{K}_k^*\right)^{\mathrm{H}} \tag{9e}$$

$$\widehat{x}_{k|k}^* = \widehat{x}_{k|k-1}^* + \mathbf{K}_k^*(y_k - \widehat{y}_{k|k-1}^*) \tag{9f}$$

To evaluate reconstruction quality, we can introduce the following errors :

$$\boldsymbol{\varepsilon}_k^{\mathbf{K}} = \mathbf{K}_k^* - \mathbf{K}_k \tag{10}$$

$$\boldsymbol{\varepsilon}_k^{x} = \widehat{x}_{k|k}^* - \widehat{x}_{k|k} \tag{11}$$

Respectively gain error and estimation error.
Using equation (9a), (11) can be rewritten :

$$\boldsymbol{\varepsilon}_k^{x} = \boldsymbol{\varepsilon}_k^{\mathbf{K}}y_k + \left(\mathbb{I} - \mathbf{K}_k^*\mathbf{H}^*\right)\widehat{x}_{k|k-1}^* - \left(\mathbb{I} - \mathbf{K}_k\mathbf{H}\right)\widehat{x}_{k|k-1} \tag{12}$$

Computing these equations gives the following results :

2.3   Direction

Direction error can be caused by a misaligned antenna, or a small displacement, but will mostly come from ionospheric perturbation[7]. To study the effects of these perturbations, a normally distributed error $\boldsymbol{\varepsilon}$ is introduced on the direction matrix :

$$\mathbf{I}_q^* = \mathbf{I}_q + \boldsymbol{\varepsilon}_q \tag{13}$$

The observations matrix becomes :

$$\mathbf{H}_{j,q}^* = \exp\left(-i\frac{2\pi}{\lambda}\Delta z_j \cdot \mathbf{I}_q^*\right) = \exp\left(-i\frac{2\pi}{\lambda}\Delta z_j \cdot \mathbf{I}_q\right)\exp\left(-i\frac{2\pi}{\lambda}\Delta z_j \cdot \boldsymbol{\varepsilon}_q\right)$$

A directional distortion matrix can be defined such as :

$$\mathbf{H}^* = \mathbf{H} \odot \mathscr{D} \tag{14}$$

Kalman equations from (9a) to (9f) stay the same.

To further measure the effect of this error, it will be introduced on a single direction first, then on all of them.

2.3.1   One direction

2.3.2   All directions

Part V
Experimental correction

1   implementation

2   Measures

Part VI

Further insights

Directional error on one antenna : on larger arrays, ionospheric perturbations might not affect all antennas the same, thus studying the reliability when one antenna has errors might be an idea to explore.

Use of smoothers : when using a Kalman filter, only prior knowledge of data is needed. When all measures at different snapshots are done, a *smoother*[5] can be used to give better Kalman results, their reliability under perturbations needs to be discussed.

Deep learning : In image processing, deep learning and artificial intelligence are getting more effective and reliable. In particular, model based deep learning [4] can be an effective way of rebuilding images when an evolution model is known.

References

[1] A. M. Sardarabadi A.-J. vand der Veen S. J. Wijnholds. "Signal Processing for Radio Astronomy". In: *Handbook of Signal Processing Systems.* 2018. DOI: https://doi.org/10.1007/978-3-319-91734-4_9.

[2] Eric Chaumette. *Introduction to Kalman Filtering and Extend Kalman Filtering.* Nov. 2018.

[3] G.Robert-Dautun M. Marchesseau. *Reconstruction dynamique d'images radioastronomiques de trous noirs.* Report. ENS Paris-Saclay, 2022. URL: example.com.

[4] Stephen P. Boyd Nir Shlezinger Yonina C. Eldar. *Model-Based Deep Learning: On the Intersection of Deep Learning and Optimization.* preprint. arXiv:2205.02640v1. arXiv, May 2022.

[5] Simo Särkkä. *Optimal Smoothing.* Lecture. Mar. 2011.

[6] K. F. Evans T. J. Cornwell. "A simple maximum entropy devoncolution algorithm". In: *Astro. Astrophys.* (1985).

[7] Spoelstra TAT. "The ionosphere and radio interferometry". In: *Ann. Geophys.* 40.4 (Nov. 1997). DOI: https://doi.org/10.4401/ag-3885.