

MITIGATION OF CONTINUOUS INTERFERENCE IN RADIO ASTRONOMY USING SPATIAL FILTERING

Sebastiaan van der Tol⁽¹⁾, Alle-Jan van der Veen⁽¹⁾ and Albert-Jan Boonstra^(1,2)

⁽¹⁾*Information Technology and Systems, Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands. S.vanderTol@its.tudelft.nl, allejan@cas.et.tudelft.nl*

⁽²⁾*ASTRON, PO Box 2, 7990 AA Dwingeloo, The Netherlands. boonstra@astron.nl*

The contamination of radio astronomical measurements by man-made Radio Frequency Interference (RFI) is becoming an increasingly serious problem and therefore the application of interference mitigation techniques is essential. Most current techniques address impulsive or intermittent interference and are based on time-frequency detection and blanking. Continuously present interferers cannot be cut out in the time-frequency plane and have to be removed using spatial filtering. One technique is based on the estimation of the spatial signature vector of the interferer from short-term spatial covariance matrices followed by a subspace projection to remove that dimension from the covariance matrix, and by further averaging. The projections will also modify the astronomical data, and hence a correction has to be applied to the long-term average to compensate for this. In this paper we analyse the performance of this spatial filtering algorithm.

1. INTRODUCTION

In interferometric radio astronomy the distribution of the intensity of radiation is measured by cross-correlating the signals from a number of radio telescopes. Unfortunately, observations are nowadays often corrupted by man-made interfering signals from sources in the same or adjacent bands, and this situation will get worse in the future. The signal from an interferer is spatially correlated and will therefore not average out completely. If the interferer is continuously present, it is not possible to filter out its contribution by detection and blanking of the contaminated samples [1].

Spatial filtering can null the energy received from the direction of the interferer. The projections will also modify the astronomical data, and hence a correction has to be applied to the long-term average to compensate for this. This algorithm was introduced in [2]. In this paper we summarize the algorithm and analyse its performance.

2. DATA MODEL AND SPATIAL FILTERING ALGORITHM

Assume we have a telescope array with p elements. For the interference free case the array output vector $\mathbf{x}_0(t)$ is modeled in complex baseband form as $\mathbf{x}_0(t) = \mathbf{v}(t) + \mathbf{n}(t)$ where $\mathbf{x}_0(t) = [x_{0,1}(t), \dots, x_{0,p}(t)]^T$ is the $p \times 1$ vector of output signals at time t , $\mathbf{v}(t)$ is the received sky signal, assumed a stationary Gaussian vector process with covariance matrix \mathbf{R}_v , and $\mathbf{n}(t)$ is the $p \times 1$ noise vector with independent identically distributed Gaussian entries and covariance matrix $\sigma^2 \mathbf{I}$. If an interferer is present the array output vector is modeled as $\mathbf{x}(t) = \mathbf{x}_0(t) + \mathbf{a}(t)s(t)$, where $s(t)$ is the interferer signal with spatial signature vector $\mathbf{a}(t)$ which is assumed stationary only over short time intervals. The astronomer is interested in \mathbf{R}_v . We assume that σ^2 is known from calibration and that $\mathbf{R}_v \ll \sigma^2 \mathbf{I}$.

Given observations $\mathbf{x}_n := \mathbf{x}(nT_s)$, where T_s is the sampling period, the objective is to estimate $\mathbf{R}_0 = \mathbf{R}_v + \sigma^2 \mathbf{I}$. We first construct short-term covariance estimates $\hat{\mathbf{R}}_k$,

$$\hat{\mathbf{R}}_k = \frac{1}{M} \sum_{n=kM}^{(k+1)M} \mathbf{x}_n \mathbf{x}_n^H,$$

where M is the number of samples per short-term average. MT_s is in the order of 1-100 millisecond. Suppose that the spatial signature \mathbf{a}_k of the interferer is known (it can be estimated from $\hat{\mathbf{R}}_k$ using an eigenvalue decomposition). We can then form a spatial filter $\mathbf{P}_k := \mathbf{I} - \mathbf{a}_k (\mathbf{a}_k^H \mathbf{a}_k)^{-1} \mathbf{a}_k^H$, which is such that $\mathbf{P}_k \mathbf{a}_k = 0$. When this spatial filter is applied to the data covariance matrix, $\hat{\mathbf{Q}}_k := \mathbf{P}_k \hat{\mathbf{R}}_k \mathbf{P}_k$, all the energy due to the interferer will be nulled. We subsequently average the modified covariance matrices to a long-term (say $T_{int} = NMT_s = 10$ seconds) estimate, $\hat{\mathbf{Q}} := \frac{1}{N} \sum_{k=1}^N \hat{\mathbf{Q}}_k$. This gives an estimate of

\mathbf{R}_0 , but it is biased due to the projection. To correct for this we first write the two-sided multiplication by \mathbf{P}_k as a single-sided multiplication, employing the matrix identity $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$, where $\text{vec}(\cdot)$ denotes the stacking of the columns of a matrix in a vector and \otimes the Kronecker product. This gives

$$\text{vec}(\hat{\mathbf{Q}}) := \frac{1}{N} \sum_{k=1}^N \mathbf{C}_k \text{vec}(\hat{\mathbf{R}}_k) \quad \text{where} \quad \mathbf{C}_k := (\mathbf{P}_k^T \otimes \mathbf{P}_k). \quad (1)$$

The bias on $\text{vec}(\hat{\mathbf{Q}})$ if the interference is completely removed is

$$\mathbb{E} [\text{vec}(\hat{\mathbf{Q}})] = \frac{1}{N} \sum_{k=1}^N \mathbf{C}_k \mathbb{E} [\text{vec}(\hat{\mathbf{R}}_{k,0})] = \mathbf{C} \text{vec}(\mathbf{R}_0) \quad \text{where} \quad \mathbf{C} := \frac{1}{N} \sum_{k=1}^N \mathbf{C}_k, \quad \hat{\mathbf{R}}_{0,k} := \frac{1}{M} \sum_{n=kM}^{(k+1)M} \mathbf{x}_{0,n} \mathbf{x}_{0,n}^H. \quad (2)$$

We can apply a correction \mathbf{C}^{-1} to $\hat{\mathbf{Q}}$ to obtain the final estimate $\hat{\mathbf{R}} := \text{unvec}(\mathbf{C}^{-1} \text{vec}(\hat{\mathbf{Q}}))$. This is the estimate of \mathbf{R}_0 produced by the algorithm. If the \mathbf{a}_k are known and completely projected out then $\hat{\mathbf{R}}$ is an unbiased estimate of \mathbf{R}_0 , i.e., $\mathbb{E}[\hat{\mathbf{R}}] = \mathbf{R}_0$. Two issues are the invertibility of \mathbf{C} and the noise enhancement due to \mathbf{C}^{-1} . Another issue is the effect of residual interference due to an *estimated* spatial signature.

An experimental result, described in more detail in [2], is shown in figure 1(a). The used data set is a $p = 8$ -channel recording at the Westerbork Radio Telescope of a 1.25 MHz-wide band at 434 MHz, containing the astronomical source 3C48 contaminated by narrow-band amateur radio broadcasts, which are both intermittent and continuous.

3. PERFORMANCE ANALYSIS

The result of the algorithm is $\hat{\mathbf{R}}$, an estimate of the true covariance matrix \mathbf{R}_0 . The quality of an estimator is determined by its covariance. In the following sections we will determine the covariance of $\hat{\mathbf{R}}$ in three cases: (I) interference free case, (II) the spatial signatures \mathbf{a}_k are known, and (III) the spatial signatures \mathbf{a}_k are estimated. We use the following notation. With $\hat{\mathbf{X}}$ we denote an estimate, with $\mathbf{X} = \mathbb{E}[\hat{\mathbf{X}}]$ the expected value of $\hat{\mathbf{X}}$ and with $\mathbf{X}' = \hat{\mathbf{X}} - \mathbf{X}$ the estimation error. The covariance of an estimate is defined as $\text{cov}\{\hat{\mathbf{X}}\} := \mathbb{E}[\text{vec}(\mathbf{X}') \text{vec}(\mathbf{X}')^H]$, and $\text{var}\{\hat{\mathbf{X}}\} := \mathbb{E}[\mathbf{X}' \odot \overline{\mathbf{X}'}] = \text{unvec}(\text{diag}(\text{cov}\{\hat{\mathbf{X}}\}))$, where \odot denotes entrywise multiplication of two matrices.

3.1. Case I: The variance of $\hat{\mathbf{R}}$ for the interference free case

Let $\hat{\mathbf{R}}_0 = \frac{1}{N} \sum_{k=1}^N \hat{\mathbf{R}}_{0,k}$ be the long-term average of interference free samples $\hat{\mathbf{R}}_{0,k}$. For Gaussian sources, it is known that

$$\text{cov}\{\hat{\mathbf{R}}_0\} = \frac{1}{MN} \mathbf{R}_0^T \otimes \mathbf{R}_0 \approx \frac{\sigma^4}{MN} \mathbf{I}, \quad (3)$$

where the approximation follows from $\mathbf{R}_0 \approx \sigma^2 \mathbf{I}$ (weak sky signal). This is the best performance expected for $\hat{\mathbf{R}}$.

3.2. Case II: The variance of $\hat{\mathbf{R}}$ for interference with known spatial signatures

Suppose the spatial signatures \mathbf{a}_k of the interferers are known. In that case the algorithm is unbiased by design. The covariance of the estimate is

$$\text{cov}\{\hat{\mathbf{R}}\} := \mathbb{E}[\text{vec}(\hat{\mathbf{R}}') \text{vec}(\hat{\mathbf{R}}')^H] = \mathbf{C}^{-1} \text{cov}\{\hat{\mathbf{Q}}\} (\mathbf{C}^{-1})^H, \quad (4)$$

where, using (1)

$$\text{cov}\{\hat{\mathbf{Q}}\} = \mathbb{E} \left[\frac{1}{N^2} \sum_{k=1}^N \sum_{l=1}^N \mathbf{C}_k \text{vec}(\mathbf{R}'_k) \text{vec}(\mathbf{R}'_l)^H \mathbf{C}_l^H \right]. \quad (5)$$

The estimation errors \mathbf{R}'_k and \mathbf{R}'_l are uncorrelated for $k \neq l$. Since $\mathbf{C} \text{vec}(\mathbf{R}'_k) = \mathbf{C} \text{vec}(\mathbf{R}'_{k,0})$,

$$\text{cov}\{\hat{\mathbf{Q}}\} = \mathbb{E} \left[\frac{1}{N^2} \sum_{k=1}^N \mathbf{C}_k \text{vec}(\mathbf{R}'_{0,k}) \text{vec}(\mathbf{R}'_{0,k})^H \mathbf{C}_k^H \right] = \frac{1}{N^2} \sum_{k=1}^N \mathbf{C}_k \text{cov}\{\hat{\mathbf{R}}_{0,k}\} \mathbf{C}_k.$$

$\hat{\mathbf{R}}_{0,k}$ is the covariance matrix of a complex Gaussian signal vector, so $\text{cov}\{\hat{\mathbf{R}}_{0,k}\} = \frac{1}{M} \mathbf{R}_0^T \otimes \mathbf{R}_0$, and

$$\text{cov}\{\hat{\mathbf{Q}}\} = \frac{1}{MN^2} \sum_{k=1}^N \mathbf{C}_k (\mathbf{R}_0^T \otimes \mathbf{R}_0) \mathbf{C}_k^H = \frac{1}{MN^2} \sum_{k=1}^N (\mathbf{P}_k \mathbf{R}_0 \mathbf{P}_k)^T \otimes \mathbf{P}_k \mathbf{R}_0 \mathbf{P}_k \approx \frac{\sigma^4}{MN^2} \sum_{k=1}^N (\mathbf{P}_k^T \otimes \mathbf{P}_k) = \frac{\sigma^4}{MN} \mathbf{C},$$

where we used that $\mathbf{R}_0 \approx \sigma^2 \mathbf{I}$ and \mathbf{P}_k is a projection. It follows that

$$\text{cov}\{\hat{\mathbf{R}}\} \approx \frac{\sigma^4}{MN} \mathbf{C}^{-1} \mathbf{C} (\mathbf{C}^{-1})^H = \frac{\sigma^4}{MN} \mathbf{C}^{-1}. \quad (6)$$

The value of \mathbf{C}^{-1} depends on \mathbf{a}_k , the spatial signatures of the interferer. Compared to (3), this indicates that \mathbf{C}^{-1} determines the relative performance of the spatial filtering algorithm.

3.3. Case III: The variance of $\hat{\mathbf{R}}$ for interference with deterministic spatial signatures

If the spatial signatures are unknown, they need to be estimated, and hence the projection matrices are estimates too. \mathbf{P}_k , \mathbf{C}_k and \mathbf{C} are substituted by their estimates $\hat{\mathbf{P}}_k$, $\hat{\mathbf{C}}_k$ and $\hat{\mathbf{C}}$. In that case equation (2) does not hold because $\hat{\mathbf{C}}$ and $\hat{\mathbf{R}}$ are not independent. The algorithm is not unbiased anymore, but it can be shown that the bias of $\hat{\mathbf{R}}$ is $O(M^{-1})$. This bias can be neglected because the standard deviation is $O(M^{-1/2})$.

Recall that $\text{cov}\{\hat{\mathbf{R}}\} = \mathbb{E}[\text{vec}(\hat{\mathbf{R}}') \text{vec}(\hat{\mathbf{R}}')^H]$. In first order approximation, $\text{vec}(\hat{\mathbf{R}}') = (\mathbf{C}^{-1})' \text{vec}(\mathbf{Q}) + \mathbf{C}^{-1} \text{vec}(\mathbf{Q}')$, where

$$\begin{aligned} (\mathbf{C}^{-1})' &= -\mathbf{C}^{-1} \mathbf{C}' \mathbf{C}^{-1}, \quad \mathbf{C}' = \frac{1}{N} \sum_{k=1}^N \mathbf{C}'_k, \quad \mathbf{C}'_k = (\mathbf{P}'_k \otimes \mathbf{P}_k) + (\mathbf{P}_k \otimes \mathbf{P}'_k) \\ \text{vec}(\mathbf{Q}') &= \frac{1}{N} \sum_{k=1}^N \text{vec}(\mathbf{Q}'_k), \quad \text{vec}(\mathbf{Q}'_k) = \mathbf{C}'_k \text{vec}(\mathbf{R}) + \mathbf{C} \text{vec}(\mathbf{R}'). \end{aligned}$$

Working this out and using $\mathbf{C}^{-1} \text{vec}(\mathbf{Q}) = \text{vec}(\mathbf{R}_0)$ ultimately leads to

$$\text{cov}\{\hat{\mathbf{R}}\} = \mathbf{C}^{-1} \text{cov}\{\hat{\mathbf{Q}}\} (\mathbf{C}^{-1})^H, \quad (7)$$

where $\text{cov}\{\hat{\mathbf{Q}}\}$ is as given in (5). Equation (7) is equal to (4) so in first order approximation, replacing the true projections \mathbf{P}_k by the estimated projections $\hat{\mathbf{P}}_k$ does not change the covariance. Also in this case it follows that $\text{cov}\{\hat{\mathbf{R}}\} \approx \frac{\sigma^4}{MN} \mathbf{C}^{-1}$.

4. THE EXPECTED VALUE OF \mathbf{C}^{-1}

\mathbf{C}^{-1} determines the penalty due to spatial filtering. The main diagonal of \mathbf{C}^{-1} contains the factors by which the variance is multiplied compared to the interference free case. To describe the penalty in a single number we introduce the ‘‘quality factor’’ $\kappa := \max(\text{diag}(\mathbf{C}^{-1}))$, which is the worst case amplification of the variance. The value of κ is a function of \mathbf{a}_k . We will determine the asymptotic value of κ for two cases: (A) \mathbf{a}_k are normally distributed and (B) \mathbf{a}_k are the spatial signatures of a stationary interferer.

4.1. Case A: The variance of $\hat{\mathbf{R}}$ for normally distributed spatial signatures

If we choose a temporally i.i.d. statistical model for \mathbf{a}_k we can determine $\mathbb{E}[\mathbf{C}_k]$. When $N \rightarrow \infty$ \mathbf{C} will converge to $\mathbb{E}[\mathbf{C}_k]$, and \mathbf{C}^{-1} to $\mathbb{E}[\mathbf{C}_k]^{-1}$. Let $\mathbf{a}_k \sim \mathcal{CN}(0, \mathbf{I})$ and i.i.d. for different k , and let $\mathbf{u}_k = \mathbf{a}_k / \|\mathbf{a}_k\|$, then \mathbf{u}_k is uniformly distributed over the unit-sphere in \mathbb{C}^p and $\mathbf{P} = \mathbf{I} - \mathbf{u}_k \mathbf{u}_k^H$. It follows that

$$\mathbb{E}[\mathbf{C}] = \mathbb{E}[\mathbf{P}^T \otimes \mathbf{P}] = \mathbb{E}[(\mathbf{I} - \mathbf{u}_k \mathbf{u}_k^H)^T \otimes (\mathbf{I} - \mathbf{u}_k \mathbf{u}_k^H)] = \mathbb{E}[\mathbf{I} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{u}_k \mathbf{u}_k^H - (\mathbf{u}_k \mathbf{u}_k^H)^T \otimes \mathbf{I} + (\mathbf{u}_k \mathbf{u}_k^H)^T \otimes \mathbf{u}_k \mathbf{u}_k^H],$$

where

$$\mathbb{E}[\mathbf{u}_k \mathbf{u}_k^H] = \frac{1}{p} \mathbf{I}, \quad \mathbb{E}[(\mathbf{u}_k \mathbf{u}_k^H)^T \otimes \mathbf{u}_k \mathbf{u}_k^H] = \frac{1}{p(p+1)} (\mathbf{I} + \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^T).$$

Hence,

$$\mathbb{E}[\mathbf{C}_k] = \mathbf{I} - \frac{2}{p} \mathbf{I} + \frac{1}{p(p+1)} (\mathbf{I} + \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^T) \quad \Leftrightarrow \quad \mathbb{E}[\mathbf{C}_k]^{-1} = \frac{p(p+1)}{p^2-p-1} (\mathbf{I} - \frac{1}{p^2-1} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^T).$$

So, for large N

$$\text{var}\{\hat{\mathbf{R}}\} = \frac{\sigma^4}{MN} \frac{p(p+1)}{p^2-p-1} (\mathbf{1}\mathbf{1}^T - \frac{1}{p^2-1} \mathbf{I}), \quad \kappa = \frac{p(p+1)}{p^2-p-1},$$

(where $\mathbf{1}$ is an all-one vector). E.g., if $p = 8$, then $\kappa = 72/55 \approx 1.3$, so the variance of the entries of $\hat{\mathbf{R}}$ increases with 30% in the worst case.

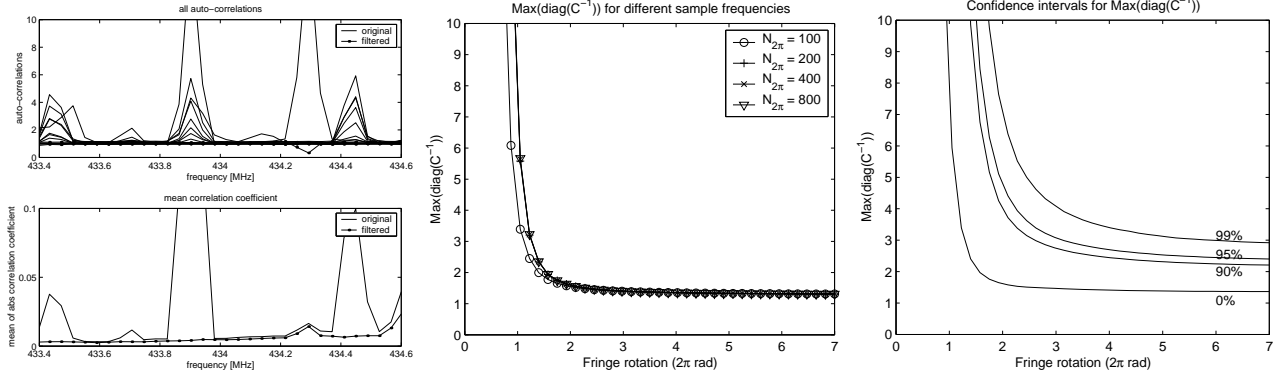


Figure 1. (a) Experimental data: correlation spectra before and after the spatial projection algorithm, (b) Quality factor κ for different $N_{2\pi}$, (c) Confidence intervals for κ for random \mathbf{a}_0 .

4.2. Case B: The variance of $\hat{\mathbf{R}}$ for stationary interferers

For matrix \mathbf{C} to be invertible the spatial signatures \mathbf{a}_k need to be sufficiently variable. For stationary interferers (no own movement, no multi-path) the only source of variability is the geometric delay compensation (a delay placed between each telescope and the correlator to correct for the different path lengths of the astronomical signal). The geometric delays depend on the position of the observed field in the sky, and are time-varying due to the earth rotation. In narrow subbands, the delays become time varying phase-shifts, named fringe corrections. For a linear array of telescopes and an interferer fixed on earth, the effect of the fringe correction on its spatial signature $\mathbf{a}(t)$ can be modeled as (see [3] for the latter expressions)

$$\mathbf{a}(t) = \begin{bmatrix} a_1 \\ a_2 e^{j\varphi t} \\ \vdots \\ a_p e^{j(p-1)\varphi t} \end{bmatrix}, \quad \mathbf{a}_0 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}, \quad \varphi = \frac{2\pi f_F}{p-1}, \quad f_F = \frac{2\pi}{24 \cdot 3600} D_\lambda \cos \delta \cos h,$$

where f_F is the fringe frequency, D_λ is the longest baseline length in wavelengths, δ is the declination of the source and h is the hour angle of the source, which is time varying and has a period of 24 hours. For a stationary interferer \mathbf{C}^{-1} depends on (i) the fringe rotation per short term sample $MT_s f_F$, where T_s is the sampling time and M the number of samples per short-term average, (ii) the number of short-term averages per long-term average N , (iii) the number of antennas p , (iv) the spatial signature without fringe correction \mathbf{a}_0 .

The first two parameters can be converted to the total fringe rotation during integration, $\varphi_{tot} = MT_s 2\pi f_F N = T_{int} 2\pi f_F$, and the number of samples per fringe cycle, $N_{2\pi} = N / (T_{int} f_F)$. The lowest possible $N_{2\pi}$ is reached when f_F reaches its maximum value. If we choose $MT_s = 10\text{ms}$, $D_\lambda = 3000\text{m}/30\text{cm}$ then the minimum value for $N_{2\pi}$ is 135. The results of simulations in figure 1(b) show that within the range of possible values for $N_{2\pi}$ the effect on the quality factor κ is neglectable. A transition from poor to reasonably good performance occurs already after 1 to 2 fringe cycles. Further simulations are carried out with the parameters $N_{2\pi} = 200$ and $p = 8$. The curves in figure 1(c) show how the performance increases with increasing fringe rotation, and a minimum φ_{tot} for acceptable performance can be determined. This condition can be formulated as a division of the sky in an “observable” and an “unobservable” area. The unobservable area is a band from the East over the celestial pole to the West. The width of this band is given by $\alpha = 2 \arcsin[(\varphi_{min} \cdot 24 \cdot 3600) / (D_\lambda \cdot T_{int} \cdot 2\pi)]$. E.g., if $\varphi_{min} = 3$, $T_{int} = 30\text{s}$, $\lambda = 30\text{cm}$ and $D = 3000\text{m}$ then the width of this band is 16° .

References

- [1] A. Leshem, A.-J. van der Veen, and A.-J. Boonstra, “Multichannel interference mitigation techniques in radio astronomy,” *The Astrophysical Journal Supplement Series*, vol. 131, pp. 355–373, Nov. 2000.
- [2] J. Raza, A.J. Boonstra, and A.J. van der Veen, “Spatial filtering of RF interference in radio astronomy,” *IEEE Signal Processing Letters*, vol. 9, no. 3, Mar. 2002.
- [3] R.A. Perley, F. Schwab, and A.H. Bridle, Eds., *Synthesis imaging in radio astronomy*, Astronomical society of the pacific, 1989.