

Parameter estimation using a sensor array in a Ricean fading channel*

K V S HARI¹ and BJÖRN OTTERSTEN²

¹Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India

²Department of Signals, Sensors and Systems, Royal Institute of Technology, S-10044 Stockholm, Sweden

e-mail: hari@ece.iisc.ernet.in; otterste@s3.kth.se

Abstract. The estimation of the Direction-Of-Arrival (DOA) and the variance of the angular spread, using an array of sensors in the case of a Ricean channel is considered, using the Maximum-Likelihood, Least-Squares and Weighted Least Squares criteria. The Cramér–Rao bound is also obtained for the problem of interest. Simplification of the cost functions to reduce the dimension of the problem has been carried out and the performance of the methods has been studied based on numerical experiments.

Keywords. Antenna arrays; fading channels; spatially distributed sources.

1. Introduction

The use of sensor arrays in the field of mobile communication to improve the performance of a cellular system has been an active area of research recently. An important issue is to efficiently use the available channel bandwidth to provide services to as many users as possible. Techniques using an array of sensors have been proposed which estimate the Direction-Of-Arrival (DOA) of the received signal from a mobile unit and the associated angular spread due to scattering, to form multiple beams on the same channel and increase user capacity (Yeh & Reudink 1982; Anderson *et al* 1991; Balaban & Salz 1992; Ohgane *et al* 1993; Zetterberg & Ottersten 1995). Estimation of DOA and angular spread of scattered field using a Uniform Linear Array (ULA) has been carried out for a Rayleigh fading channel using the Maximum-Likelihood and least squares criteria (Trump & Ottersten 1996). In this paper, a fading channel whose amplitude response has a Rice distribution is considered, and techniques to estimate the DOA and other relevant parameters are presented based on the Maximum-Likelihood and Least Squares

*A major part of the work was carried out when K V S Hari was visiting the Department of Signals, Sensors and Systems during Jul–Sep 1995, on leave from the Indian Institute of Science

criteria. A comparison of the performance of these methods is carried out using numerical simulations.

2. Datamodel

In a mobile communication scenario, the narrowband signal from a single mobile unit received at the base station using a sensor-array, is assumed to be a superposition of a *large* number of signals with different strengths and arriving from angular directions *close* to the direction of the source. This model has been verified by experiments to characterise the scattering effects of the channel (Adachi *et al* 1986). In this report, a ULA of L sensors with an inter-sensor distance, d_λ (in wavelengths) is considered. The response of the array to a unit-amplitude narrowband signal impinging from a direction θ (with respect to the broadside of the array), is known as the *array response vector* for the direction θ . It is denoted by $\mathbf{a}(\theta)$ and its k th element is defined as¹

$$[\mathbf{a}(\theta)]_k = \exp[j2\pi d_\lambda(k - 1) \sin(\theta)].$$

Assume that a single source is transmitting a narrowband signal $s(t)$, from a direction θ which is being received by the ULA due to scatterers in the vicinity of the source. The noiseless output of the k th sensor as a function of time t , can be written as

$$y_k(t) = \left(\sum_n g_n(t) e^{j\alpha_n(t)} e^{(j2\pi d_\lambda(k-1) \sin(\theta + \theta_n))} \right) s(t),$$

where $g_n(t)$, $\alpha_n(t)$ are the amplitude and phase factors due to the n th scatterer and θ_n is the deviation with respect to θ due to the n th scatterer. Stacking the outputs of the sensors into a vector, $\mathbf{y}(t)$, the *snapshot* of the array output can be written as

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{h}(t) s(t), \\ \mathbf{h}(t) &= \sum_n g_n(t) e^{j\alpha_n(t)} \mathbf{a}(\theta + \theta_n), \end{aligned}$$

where $\mathbf{h}(t)$ denotes the *channel response vector* at time instant t .

2.1 Fading channels

A channel is said to be a *fading channel* if the amplitudes and phases introduced by the scatterers and the directions of the scatterers are random and vary with time (Proakis 1989; Braun & Dersch 1991). Further, if these are independent and identically distributed and if the number of scatterers is very large, the Central-Limit Theorem can be invoked and the channel response vector $\mathbf{h}(t)$, can be thought of as a complex Gaussian random process. In order to obtain the statistical description of $\mathbf{y}(t)$, a statistical description of the scattered

¹Notation: $E[.]$ denotes the expectation operator. Superscripts T, * and H denote transpose, complex conjugate and complex-conjugate transpose respectively. Tr(.) and det(.) denote the trace and the determinant of a matrix respectively. \bullet denotes the Schur product operation on matrices and \mathbf{I} is the identity matrix. Boldfaced lower case and upper case letters denote vectors and matrices respectively.

signal amplitudes, phases and the directions needs to be done. Trump & Ottersten (1996) have assumed that

$$\begin{aligned} E[g_n(t)e^{j\alpha_n(t)}g_m(s)e^{-j\alpha_m(s)}] &= 0, \quad n \neq m, \quad t \neq s, \\ &= 1, \quad \text{otherwise}, \\ E[e^{j\alpha_n(t)}] &= 0. \end{aligned}$$

Based on these assumptions, $\text{Cov}[\mathbf{h}(t), \mathbf{h}(\tau)] = \mathbf{C}_h \delta(t - \tau)$. Many distributions of θ_n have been proposed by Anderson *et al* (1991), Parsons & Turknani (1991), Proakis (1991) and Trump & Ottersten (1996) and in this paper it is assumed that θ_n is small and $\theta_n \sim \mathcal{N}(m_\theta, \sigma_\theta)$.

2.1a Rayleigh fading channel: In an urban environment, usually, there is no direct path between the mobile unit and the base station sensor may and it is assumed that $\mathbf{m}_h(t) = E[\mathbf{h}(t)] = \mathbf{0}$ (Zetterberg & Ottersten 1995; Trump & Ottersten 1996) and such a channel is known as a Rayleigh fading channel (Proakis 1989). Assuming, $m_\theta = 0$, \mathbf{C}_h is given as (Trump & Ottersten 1996)

$$\mathbf{C}_h = \mathbf{R}_a \bullet \mathbf{B}(\theta, \sigma_\theta),$$

where the k th element of \mathbf{B} is given by

$$[\mathbf{B}]_{kl} = \exp\{-2[\pi(k-l)d_\lambda]^2\sigma_\theta \cos^2(\theta)\},$$

and

$$\mathbf{R}_a = \mathbf{a}(\theta) \mathbf{a}^H(\theta).$$

Model M1 – Let $s(t)$ be a *random* signal uncorrelated with $\mathbf{h}(t)$ and $s(t) = a \exp[j\phi_s(t)]$ with a being a deterministic quantity and $\phi_s(t)$ being distributed *uniformly* between 0 and 2π . Then $\mathbf{m}_y(t) = \mathbf{0}$, $\mathbf{C}_y(t) = \mathbf{C}_h |\alpha|^2$.

It can be easily shown using elementary probability theory, that $\mathbf{y}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_y(t))$ for model M1.

Remark 1. It is to be noted that if $s(t)$ is Gaussian, $\mathbf{y}(t)$ is no longer Gaussian.

Model M2 – If $s(t)$ is a *deterministic* signal, then $\mathbf{m}_y(t) = \mathbf{0}$, $\mathbf{C}_y(t) = \mathbf{C}_h |\mathbf{s}(t)|^2$ and $\mathbf{y}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_y(t))$.

2.2 Ricean fading channel

It is assumed that in a suburban or rural environment, the signal received has an additional component due to a direct path, also known as the *line-of sight* (LOS) path. Let γ denote the factor which controls the strength of the signal in the LOS direction, θ . Then the output of the array can be written as

$$\mathbf{y}(t) = (\gamma \mathbf{a}(\theta) + \mathbf{h}(t)) s(t).$$

Case 1: Let γ be a deterministic but unknown quantity.

Case 2: Let $\gamma \sim N(\gamma_0, \sigma_\gamma)$ independent of other parameters. Since a LOS component is assumed, $\gamma_0 \neq 0$. It is clear that the new channel response vector in both cases is Gaussian and has a non-zero mean of the form $\gamma \mathbf{a}(\theta)$. This channel is said to be a *Ricean fading channel* because the envelope has a Rice distribution (Proakis 1989).

Model M3 – Let $s(t)$ be a *random* signal uncorrelated with $\mathbf{h}(t)$ and as given in model M1. Then $\mathbf{m}_y(t) = \mathbf{0}$, $\mathbf{C}_y(t) = (\mathbf{C}_h + \mathbf{m}_h \mathbf{m}_h^H) |\alpha|^2$. As in the case of model M1, $\mathbf{y}(t)$ is Gaussian.

Model M4 – $s(t)$ is a *deterministic* signal. Then $\mathbf{m}_y(t) = \mathbf{m}_h s(t)$, $\mathbf{C}_y(t) = \mathbf{C}_h |s(t)|^2$ and $\mathbf{y}(t) \sim \mathcal{N}(\mathbf{m}_y(t), \mathbf{C}_y(t))$.

2.3 Noisy data

Consider the noisy array output as

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t),$$

where $\mathbf{n}(t)$ is the additive noise vector which is zero-mean Gaussian and satisfies $E[\mathbf{n}(t)s(t)] = \mathbf{0}$, $E[\mathbf{n}(t)\mathbf{n}^H(t)] = \sigma_n \mathbf{I}$ and $E[\mathbf{n}(t)\mathbf{n}^T(\tau)] = \mathbf{0}$. Then $\mathbf{C}_x(t) = \mathbf{C}_y(t) + \sigma_n \mathbf{I}$ for any of the above models. In this paper, the Ricean fading channel is assumed (models M3 and M4) with γ being deterministic but unknown. The covariance and correlation matrices for $\mathbf{x}(t)$ belonging to these models can be obtained as

$$\begin{aligned}\mathbf{R}_x &= \sigma_\gamma \sigma_s \mathbf{R}_a + \sigma_s \mathbf{R}_b + \sigma_n \mathbf{I}, \\ \mathbf{C}_x &= \sigma_s \mathbf{R}_b + \sigma_n \mathbf{I},\end{aligned}$$

where $\sigma_\gamma = |\gamma|^2$, $\sigma_s = E|s(t)|^2$ and \mathbf{m}_x is the mean of $\mathbf{x}(t)$.

3. Parameter estimation

It is clear that there are two cases of interest: (i) Random source (model M3), and (ii) deterministic source (model M4). The estimation of parameters for the Rayleigh fading channel assuming a random source (model M1) was carried out (Trump & Ottersten 1996). In this paper, the estimation of the parameters for models M3 and M4 using the Maximum Likelihood and Weighted Least Squares criteria is presented.

Problem statement. Given N snapshots of the array outputs, $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$, obtain the parameter vector $\boldsymbol{\eta} = [\theta, \sigma_\theta, \sigma_s, \sigma_n, \mathbf{d}]$ for model M3 and $\boldsymbol{\eta} = [\theta, \sigma_\theta, \sigma_n, \sigma_\gamma, s(t), t = 1, 2, \dots, N]^T$ for model M4.

4. Cramér–Rao bound

Let $\hat{\boldsymbol{\eta}}$ denote an unbiased estimate of $\boldsymbol{\eta}$. The Cramér–Rao bound on the error covariance of the estimated parameter vector can be written as

$$E\{(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})^T\} \geq \mathbf{J}_{\boldsymbol{\eta}}^{-1},$$

where $\mathbf{J}_{\boldsymbol{\eta}}$ is the *Fisher Information Matrix*. For model M3, the kl th element of $\mathbf{J}_{\boldsymbol{\eta}}$ is (Trump & Ottersten 1996)

$$[\mathbf{J}_{\boldsymbol{\eta}}]_{kl} = N \operatorname{Tr} \left[\mathbf{R}_x^{-1} \frac{\partial \mathbf{R}_x}{\partial \boldsymbol{\eta}_k} \mathbf{R}_x^{-1} \frac{\partial \mathbf{R}_x}{\partial \boldsymbol{\eta}_l} \right], \quad (1)$$

where $\partial(\cdot)/\partial \boldsymbol{\eta}_k$ denotes differentiation with respect to the k th parameter of $\boldsymbol{\eta}$. For model M4, the kl th element of $\mathbf{J}_{\boldsymbol{\eta}}$ can be easily shown to be

$$[\mathbf{J}_{\boldsymbol{\eta}}]_{kl} = N \operatorname{Tr} \left[\mathbf{C}_x^{-1} \frac{\partial \mathbf{C}_x}{\partial \boldsymbol{\eta}_k} \mathbf{C}_x^{-1} \frac{\partial \mathbf{C}_x}{\partial \boldsymbol{\eta}_l} + 2 \mathbf{C}_x^{-1} \frac{\partial \mathbf{m}_x}{\partial \boldsymbol{\eta}_k} \frac{\partial \mathbf{m}_x^H}{\partial \boldsymbol{\eta}_l} \right]$$

where \mathbf{m}_x , \mathbf{C}_x and \mathbf{R}_x denote the mean, covariance matrix and the correlation matrix of $\mathbf{x}(t)$.

5. Maximum-likelihood estimation

Given the Gaussian nature of $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$, the negative log-likelihood function for model M3 is given as

$$l_{ML}(\theta, \sigma_\theta, \sigma_s, \sigma_n, \sigma_\gamma) = \log(\det(\mathbf{R}_x)) + \operatorname{Tr}[\mathbf{R}_x^{-1} \hat{\mathbf{R}}]$$

and the conditional negative log-likelihood function for model M4 is given as

$$\begin{aligned} l_{ML}(\theta, \sigma_\theta, \sigma_n, \sigma_\gamma, s(t), t = 1, \dots, N) \\ = \log(\det(\mathbf{C}_x)) + \operatorname{Tr}[\mathbf{C}_x^{-1} \mathbf{M}], \end{aligned}$$

where \mathbf{M} is defined as

$$\mathbf{M} = \hat{\mathbf{R}} + \mathbf{m}\mathbf{m}^H - 2\mathbf{m}\hat{\mathbf{m}}^H$$

with

$$\begin{aligned} \hat{\mathbf{R}} &= \frac{1}{N} \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^H(t), \\ \hat{\mathbf{m}} &= \frac{1}{N} \sum_{t=1}^N \mathbf{x}(t). \end{aligned}$$

\mathbf{R} is the data correlation matrix² and \mathbf{m} is the sample mean of $\mathbf{x}(t)$. The maximum-likelihood (ML) estimate of $\boldsymbol{\eta}$ is obtained by minimizing l_{ML} in the parameter space of $\boldsymbol{\eta}$. Results from estimation theory guarantee that the ML estimates are asymptotically efficient (Mendel 1989). Since there is no closed form solution to this function, search methods need to be used. Usually, obtaining the ML solution is computationally expensive as the dimension of the parameter space increases. Also, good initial estimates are necessary to avoid convergence to the local minima.

²It is assumed that \mathbf{R} tends to \mathbf{R}_x as N tends to infinity and $E(\hat{\mathbf{R}}) = \mathbf{R}_x$.

6. Least squares estimation

As the solution to the ML problem is computationally expensive, a look at other simpler criteria like the least squares criterion is worthwhile.

6.1 Weighted least squares

The general form of least squares is the weighted least squares (WLS) cost function which can be expressed as

$$\begin{aligned} l = l_{\text{WLS}}(\boldsymbol{\eta}) &= \|\mathbf{W}^{\text{H}/2}(\mathbf{R}_x - \hat{\mathbf{R}})\mathbf{W}^{1/2}\|_F^2 \\ &= \text{Tr}[(\mathbf{R}_x - \hat{\mathbf{R}})\mathbf{W}(\mathbf{R}_x - \hat{\mathbf{R}})\mathbf{W}], \end{aligned}$$

where \mathbf{W} is a positive definite *weighting* matrix. The choice of \mathbf{W} is usually made such that the error-covariance of the parameter vector, $\boldsymbol{\eta}$, is minimized. Denoting the estimate of $\boldsymbol{\eta}_0$ as $\hat{\boldsymbol{\eta}}$, the error in the parameter vector is given by

$$\tilde{\boldsymbol{\eta}} = (\boldsymbol{\eta}_0 - \hat{\boldsymbol{\eta}}) \approx -\mathbf{H}^{-1}(\boldsymbol{\eta}_0) \frac{\partial l}{\partial \boldsymbol{\eta}}.$$

where \mathbf{H} is the Hessian of the cost function given by

$$[\mathbf{H}]_{ij} = \frac{\partial^2 l}{\partial \boldsymbol{\eta}_i \partial \boldsymbol{\eta}_j}.$$

The i th element of the gradient vector \mathbf{l}' of the cost function can be obtained as

$$\begin{aligned} \mathbf{l}' &= \frac{\partial l}{\partial \boldsymbol{\eta}_i} = 2 \text{Tr}[(\mathbf{R}_x - \mathbf{R}) \mathbf{D}_i], \\ \mathbf{D}_i &= \mathbf{W} \frac{\partial \mathbf{R}_x}{\partial \boldsymbol{\eta}_i} \mathbf{W}. \end{aligned}$$

Following the development by Trump & Ottersten (1996),

$$\sqrt{N} \mathbf{l}' \sim \text{As } \mathcal{N}(\mathbf{0}, \mathbf{Q}),$$

where

$$\mathbf{Q} = \lim_{N \rightarrow \infty} N E[\mathbf{l}' (\mathbf{l}')^T].$$

Hence, the asymptotic distribution of the estimation error is given by

$$\sqrt{N} \tilde{\boldsymbol{\eta}} \sim \text{As } \mathcal{N}(\mathbf{0}, \mathbf{C}),$$

where

$$\mathbf{C} = \mathbf{H}^{-1} \mathbf{Q} \mathbf{H}^{-1}.$$

The ij th element of the limiting Hessian is given as

$$[\mathbf{H}]_{ij} = 2 \text{Tr} \left\{ \frac{\partial \mathbf{R}_x}{\partial \boldsymbol{\eta}_j} \mathbf{W} \frac{\partial \mathbf{R}_x}{\partial \boldsymbol{\eta}_i} \right\}.$$

The ij th element of \mathbf{Q} is given as

$$\begin{aligned} E\left[\frac{\partial l}{\partial \eta_i} \frac{\partial l}{\partial \eta_j}\right] &= 4 E[\text{Tr}\{(\mathbf{R}_x - \hat{\mathbf{R}}) \mathbf{D}_i\} \\ &\quad \times \text{Tr}\{(\mathbf{R}_x - \hat{\mathbf{R}}) \mathbf{D}_j\}] \\ &= 4 \text{Tr}\{\mathbf{I} - \mathbf{D}_i\} \text{Tr}\{\mathbf{R}_x \mathbf{D}_j\} \\ &\quad - 4 \text{Tr}\{\mathbf{R}_x \mathbf{D}_i\} E[\text{Tr}\{\hat{\mathbf{R}}_x \mathbf{D}_j\}] \\ &\quad - 4 E[\text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_i\}] \text{Tr}\{\mathbf{R}_x \mathbf{D}_j\} \\ &\quad + 4 E[\text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_i\} \text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_j\}] \\ &= 4 E[\text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_i\} \text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_j\}] \\ &\quad - 4 \text{Tr}\{\mathbf{R}_x \mathbf{D}_i\} \text{Tr}\{\mathbf{R}_x \mathbf{D}_j\} \end{aligned}$$

as $E[\hat{\mathbf{R}}] = \mathbf{R}$. The second term in the above equation can be written as

$$\begin{aligned} &= 4 E[\text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_i\} \text{Tr}\{\hat{\mathbf{R}} \mathbf{D}_j\}] \\ &= \frac{4}{N^2} \sum_{t \tau l m o p} E[\mathbf{x}_l^*(t) \mathbf{x}_m(t) \mathbf{x}_o^*(\tau) \mathbf{x}_p(\tau)] [\mathbf{D}_i]_{op} [\mathbf{D}_j]_{lm}. \end{aligned} \quad (2)$$

6.1a *Random source (model M3):* For the random source case, since $\mathbf{m}_x(t) = \mathbf{0}$, the above product of four Gaussian random variables with zero mean can be expressed as

$$\begin{aligned} &= \frac{4}{N^2} \sum_{t \tau l m o p} (E[\mathbf{x}_l^*(t) \mathbf{x}_m(t)] E[\mathbf{x}_o^*(\tau) \mathbf{x}_p(\tau)] \\ &\quad + E[\mathbf{x}_l^*(t) \mathbf{x}_p(\tau)] E[\mathbf{x}_o^*(\tau) \mathbf{x}_m(t)]) [\mathbf{D}_i]_{op} [\mathbf{D}_j]_{lm} \\ &= 4 \text{Tr}\{\mathbf{R}_x \mathbf{D}_i\} \text{Tr}\{\mathbf{R}_x \mathbf{D}_j\} + \frac{4}{N^2} \text{Tr}\{\mathbf{R}_x \mathbf{D}_i \mathbf{R}_x \mathbf{D}_j\}. \end{aligned}$$

Thus

$$E\left[\frac{\partial l}{\partial \eta_i} \frac{\partial l}{\partial \eta_j}\right] = \frac{4}{N} \text{Tr}\left\{\mathbf{R}_x \mathbf{W} \frac{\partial \mathbf{R}_x}{\partial \eta_i} \mathbf{W} \mathbf{R}_x \mathbf{W} \frac{\partial \mathbf{R}_x}{\partial \eta_j} \mathbf{W}\right\}.$$

It was shown (Goransson 1995) that $\mathbf{W} = \mathbf{R}_x^{-1}$ would yield

$$[\mathbf{C}^{-1}]_{ij} = \text{Tr}\left\{\frac{\partial \mathbf{R}_x}{\partial \eta_j} \mathbf{R}_x^{-1} \frac{\partial \mathbf{R}_x}{\partial \eta_i} \mathbf{R}_x^{-1}\right\},$$

which achieves the CR bound given by (1). In practice, $\mathbf{W} = \hat{\mathbf{R}}^{-1}$ is used which is a consistent estimator of \mathbf{R}_x^{-1} .

Therefore, the WLS criterion can be simplified to

$$l = \text{Tr}\{(\mathbf{R}_x \hat{\mathbf{R}}^{-1} - \mathbf{I})^2\}.$$

To obtain the estimates, a multidimensional search in the parameter space is required. Since the cost function is quadratic in $\sigma_s, \sigma_y, \sigma_n$, these parameters can be separated as follows.

Differentiating w.r.t. $\sigma_s, \sigma_\gamma, \sigma_n$, setting to zero, and simplifying, the following equations hold,

$$\begin{aligned} a_{11}\sigma_\gamma\sigma_s + a_{12}\sigma_s + a_{13}\sigma_n &= b_1, \\ a_{21}\sigma_\gamma\sigma_s + a_{22}\sigma_s + a_{11}\sigma_n &= b_2, \\ a_{22}\sigma_\gamma\sigma_s + a_{32}\sigma_s + a_{12}\sigma_n &= b_3, \end{aligned}$$

where $a_{11} = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_a\hat{\mathbf{R}}^{-1}]$, $a_{12} = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_b\hat{\mathbf{R}}^{-1}]$, $a_{13} = \text{Tr}[\hat{\mathbf{R}}^{-1}\hat{\mathbf{R}}^{-1}]$, $a_{21} = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_a\hat{\mathbf{R}}^{-1}\mathbf{R}_a]$, $a_{22} = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_b\hat{\mathbf{R}}^{-1}\mathbf{R}_a]$, $a_{32} = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_b\hat{\mathbf{R}}^{-1}\mathbf{R}_b]$, $b_1 = \text{Tr}[\hat{\mathbf{R}}^{-1}]$, $b_2 = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_a]$, $b_3 = \text{Tr}[\hat{\mathbf{R}}^{-1}\mathbf{R}_b]$. Using the above equations, one can obtain expressions for the three parameters as

$$\begin{aligned} \hat{\sigma}_n &= \frac{a_{22}b_1 - a_{11}b_3 - \sigma_s[a_{12}a_{22} - a_{32}a_{11}]}{[a_{13}a_{22} - a_{12}a_{11}]}, \\ \hat{\sigma}_\gamma &= \frac{a_{11}b_1 - a_{13}b_2 - \sigma_s[a_{11}a_{12} - a_{13}a_{22}]}{\sigma_s[a_{11}^2 - a_{13}a_{21}]}, \\ \hat{\sigma}_s &= \frac{c_1 + c_2 + c_3}{c_4}, \end{aligned}$$

where

$$\begin{aligned} c_1 &= b_1 \left(1 - \frac{a_{11}^2}{[a_{11}^2 - a_{13}a_{21}]} - \frac{a_{13}a_{22}}{[a_{13}a_{22} - a_{12}a_{11}]} \right), \\ c_2 &= b_2 \left(\frac{a_{13}a_{11}}{[a_{11}^2 - a_{13}a_{21}]} \right), c_3 = b_3 \left(\frac{a_{13}a_{11}}{[a_{13}a_{22} - a_{12}a_{11}]} \right), \\ c_4 &= \left(a_{12} - \frac{a_{11}[a_{11}a_{12} - a_{13}a_{22}]}{-a_{13}a_{21}} - \frac{a_{13}[a_{12}a_{22} - a_{32}a_{11}]}{[a_{13}a_{22} - a_{12}a_{11}]} \right). \end{aligned}$$

The weighted least-squares cost function can now be recast as

$$l_{WLS}(\theta, \sigma_\theta) = \text{Tr}((\hat{\sigma}_\gamma\hat{\sigma}_s\mathbf{R}_a + \hat{\sigma}_s\mathbf{R}_b + \hat{\sigma}_n\mathbf{I})\hat{\mathbf{R}}^{-1} - \mathbf{I})^2],$$

and the search is now over a two-dimensional space of $[\theta, \sigma_\theta]$ which is computationally less expensive than before.

6.1b Deterministic source (Model M4): For the model M4, the random variables in (2) have non-zero means and thus

$$\begin{aligned} E[\text{Tr}\{\hat{\mathbf{R}}\mathbf{D}_i\} \text{Tr}\{\hat{\mathbf{R}}\mathbf{D}_j\}] &= \frac{1}{N^2} \sum_{t\tau l m o p} \{ E[\mathbf{x}_l^*(t)\mathbf{x}_m(t)] E[\mathbf{x}_o^*(\tau)\mathbf{x}_p(\tau)] \\ &\quad + E[\mathbf{x}_l^*(t)\mathbf{x}_o^*(\tau)] E[\mathbf{x}_m(t)\mathbf{x}_p(\tau)] \\ &\quad + E[\mathbf{x}_l^*(t)\mathbf{x}_p(\tau)] E[\mathbf{x}_m(t)\mathbf{x}_o^*(\tau)] \\ &\quad - 2 E[\mathbf{x}_l^*(t)] E[\mathbf{x}_m(t)] E[\mathbf{x}_o^*(\tau)] E[\mathbf{x}_p(\tau)] \} \\ &\quad \times (\mathbf{D}_i)_{op} (\mathbf{D}_j)_{lm} \end{aligned}$$

After some tedious calculations, the ij th element of \mathbf{Q} for this case is given as

$$[\mathbf{Q}]_{ij} = \frac{4}{N} \text{Tr}\{\mathbf{R}_x \mathbf{D}_i\} \text{Tr}\{\mathbf{R}_x \mathbf{D}_j\} + \frac{4}{N} \text{Tr}\{\mathbf{R}'_x \mathbf{D}_i\} \text{Tr}\{(\mathbf{R}'_x)^T \mathbf{D}_j\} \\ - 8 \text{Tr}[\mathbf{m}_x \mathbf{m}_x^H \mathbf{D}_i] \text{Tr}[\mathbf{m}_x \mathbf{m}_x^H \mathbf{D}_j],$$

where

$$\mathbf{R}'_x = E[\mathbf{x}(t) \mathbf{x}^T(t)] = \mathbf{m}_x(t) \mathbf{m}_x^T(t).$$

Remark 2.

- It is difficult to obtain \mathbf{W} from the above equation for \mathbf{Q} , which minimizes \mathbf{C} .
- The same result holds good for model M2 also, with the appropriate correlation matrix.

6.1c Least squares: A more popular criterion, which is simpler, is the least squares criterion, defined for $\mathbf{W} = \mathbf{I}$ as

$$l_{LS} = \text{Tr}[(\hat{\mathbf{R}} - \mathbf{R}_x)(\hat{\mathbf{R}} - \mathbf{R}_x)^H]$$

Random source (Model M3) – As in the case of the WLS criterion, since the cost function is quadratic in $\sigma_s, \sigma_\gamma, \sigma_n$, these parameters can be separated. Using some simple identities,³ one can obtain

$$\hat{\sigma}_n = \frac{\beta \text{Tr}[\hat{\mathbf{R}}] - L \text{Tr}[\hat{\mathbf{R}} \mathbf{R}_b] - \sigma_s L (\beta - \text{Tr}[\mathbf{R}_b \mathbf{R}_b^H])}{L(\beta - L)}, \\ \hat{\sigma}_\gamma = \frac{\text{Tr}[\hat{\mathbf{R}} \mathbf{R}_a] - \text{Tr}[\hat{\mathbf{R}}] - \sigma_s (\beta - L)}{\sigma_s L (L - 1)}, \\ \hat{\sigma}_s = \frac{\text{numerator}}{\text{denominator}},$$

where

$$\begin{aligned} \text{numerator} &= \text{Tr}[\hat{\mathbf{R}}]((L^2 - 2L + \beta)/((\beta - L)(L - 1))) \\ &\quad - \text{Tr}[\hat{\mathbf{R}} \mathbf{R}_a](1/(L - 1)) + \text{Tr}[\hat{\mathbf{R}} \mathbf{R}_b](L/(\beta - L)), \\ \text{denominator} &= L - (L(\beta - \text{Tr}[\mathbf{R}_b \mathbf{R}_b]))/(\beta - L) - (\beta - L)/(L - 1). \end{aligned}$$

The least-squares cost function can now be recast as

$$l_{LS}(\theta, \sigma_\theta) = \text{Tr}[(\hat{\mathbf{R}} - \hat{\sigma}_\gamma \hat{\sigma}_s \mathbf{R}_a - \hat{\sigma}_s \mathbf{R}_b - \hat{\sigma}_n \mathbf{I})^2]$$

and the search is now over a two-dimensional space of $[\theta, \sigma_\theta]$ which is computationally less expensive than before.

³ $\text{Tr}[\mathbf{R}_a] = \text{Tr}[\mathbf{R}_b] = L$, $\text{Tr}[\mathbf{R}_a \mathbf{R}_a^H] = L^2$, $\text{Tr}[\mathbf{R}_a \mathbf{R}_b] = \beta = \sum_{k=1}^L \sum_{l=1}^L [\mathbf{B}]_{kl}$

Table 1. Mean squared error (deg^2) in the DOA ($=10 \text{ deg}$) vs L for $\sigma_\theta = 1$ for ML, LS, WLS criteria.

L	ML	LS	WLS
4	$1.2560e-02$	$1.2512e-02$	$1.3056e-02$
6	$4.4170e-03$	$4.3335e-03$	$5.6491e-03$
8	$3.7610e-03$	$3.7726e-03$	$4.3883e-03$
10	$2.3106e-03$	$2.2610e-03$	$2.9380e-03$

7. Numerical study

An experiment to study the performance of the algorithms based on the ML, LS and WLS criteria, is presented next.

Experiment. A scenario with $\theta = 10^\circ$, $\sigma_s = 10$, $\sigma_n = 1$, $\sigma_\gamma = 2$, $N = 100$ for a ULA is considered. Various values of σ_θ , L are considered as $\sigma_\theta = 1, 2, 3, 4, 5$ and $L = 4, 6, 8, 10, 12$.

The estimates of θ and σ_θ are obtained for each of the above combinations of L and σ_θ for the ML, LS and WLS criteria using the same data vectors. In this report a Gauss–Newton search method is used to obtain the parameters. The updated vector at each iteration is given by

$$\hat{\eta}(k+1) = \hat{\eta}(k) - \mu(k) \mathbf{H}^{-1} \mathbf{g},$$

where k denotes the iteration, \mathbf{H} is the Hessian and \mathbf{g} is the gradient of the cost function considered and $\mu(k)$ is the step-size at the k th iteration. The initial estimate of θ is obtained by using the ESPRIT algorithm and the initial estimate of $\sigma_\theta = 0$ for all the criteria. The sample statistics of the estimates are obtained from 200 independent trials.

Effect of number of sensors: Table 1 presents the MSE in the DOA as a function of L for a particular value of the angular spread, σ_θ .

- The MSE decreases as L increases for all methods which agrees with intuition,
- For any value of L , the performance of the LS method is very close to that of the ML method while the WLS method performs poorly in comparison with the other methods. This could be due to use of the estimate of the correlation matrix instead of the true weighting matrix.

Table 2 presents the MSE in the DOA as a function of the angular spread of the scatterers for a particular value of L .

Table 2 Mean squared error(deg^2) in the DOA ($=10 \text{ deg}$) vs σ_θ for $L = 10$ for ML, LS, WLS criteria.

σ_θ	ML	LS	WLS
1	$2.3106e-03$	$2.2610e-03$	$2.9380e-03$
2	$3.2862e-03$	$3.4789e-03$	$4.1946e-03$
3	$4.4316e-03$	$4.8979e-03$	$5.6216e-03$
4	$5.4543e-03$	$6.2753e-03$	$6.9071e-03$
5	$6.3826e-03$	$7.6113e-03$	$8.1168e-03$

- It is clear that the performance deteriorates as the angular spread increases, for all methods.
- As observed before, WLS performs poorer than LS and ML while the performance of ML is the best among the methods.

8. Conclusions

Estimation of parameters, the DOA and the variance of the angular spread, using an array of sensors in the case of a Ricean channel is considered, using the maximum-likelihood, least-squares and weighted least squares criteria. The Cramér–Rao bound is also obtained for the problem of interest. Due to the quadratic nature of the least-squares criteria, simplification of the cost functions to reduce the dimension of the problem has been carried out. The performance of the methods (in terms of the mean-squared error in the estimates of the parameters) has been studied based on numerical experiments which show that the maximum-likelihood and least-squares methods perform comparably while the weighted-least squares method is slightly poorer than the other methods. This could be due to the use of an estimated correlation matrix as the weighting matrix instead of the true one.

References

- Adachi F, Feeny M T, Williamson A G, Parsons J D 1986 Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site. *Inst. Elec. Eng. Proc.* 133: 506–512
- Anderson S, Millnert M, Viberg M, Wahlberg B 1991 An adaptive array for mobile communication systems. *IEEE Trans. Vehicular Technol.* 40: 230–236
- Balaban P, Salz J 1992 Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio. *IEEE Trans. Commun.* 40: 865–907
- Braun W R, Dersch U 1991 A physical mobile radio channel model. *IEEE Trans. Vehicular Technol.* 40: 472–482
- Goransson B 1995 Parametric methods for source localization in the presence of spatially correlated noise. Technical report TRITA-S3-SB-9503, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm
- Mendel J M 1989 *Lessons in digital estimation theory* (Englewood Cliffs, NJ: Prentice-Hall)
- Ohgane T, Shimura T, Matsuzawa N, Sasaoka H 1993 An implementation of a CMA adaptive array for high speed GMSK transmission in mobile communications. *IEEE Trans. Vehicular Technol.* 42: 282–288
- Parsons J D, Turkmani A M D 1991 Characterization of mobile radio signals; model description. *Inst. Elec. Eng. Proc.* 138: 549–555
- Proakis J G 1989 *Digital communications* 2nd edn (Singapore: McGraw-Hill)
- Trump T, Ottersten B 1996 Estimation of nominal direction of arrival and angular spread using an array of sensors. *Signal Process.* 50: 57–69
- Yeh Y S, Reudink D 1982 Efficient spectrum utilization for mobile radio systems using space diversity. *IEEE Trans. Commun.* 30: 447–455
- Zetterberg P, Ottersten B 1995 The spectrum efficiency of a base-station antenna array system for spatially selective transmission. *IEEE Trans. Vehicular Technol.* 44: 651–660