Computationally Efficient DOA Estimation of a Scattered Source*

Petre Stoica† and Olivier Besson‡

†Department of Systems and Control, Uppsala University, Uppsala, Sweden. ps@syscon.uu.se.
‡Department of Avionics and Systems, ENSICA, Toulouse, France. besson@ensica.fr.

ABSTRACT
The problem of Direction-Of-Arrival (DOA) estimation in the presence of local scatterers using a uniform linear array (ULA) of sensors is addressed. We consider two models depending on whether the form of the azimuthal power distribution is explicitly known or not. For both models, the block-diagonal structure of the associated Fisher Information Matrix (FIM) is exploited to decouple the estimation of the DOA from that of the other model parameters. An asymptotically efficient Maximum Likelihood (ML) DOA estimator is derived which entails solving a 1-D minimization problem only. Furthermore, the 1-D criterion can be expressed as a simple Fourier Transform. Numerical simulations illustrate the fact that our computationally very simple DOA estimators have an accuracy close to the Cramér-Rao bound in a wide range of scenarios.

1 PROBLEM STATEMENT
We consider a suburban or rural mobile communication scenario where the source is surrounded by scatterers, cf. Figure 1. The signal received at a base station antenna array is thus the superposition of all contributions due to scatterers in the vicinity of the mobile. The local scattering phenomenon was recently modeled from a mathematical point of view in [8] and evidenced from a campaign of measurements led at the University of Aalborg [6]. Briefly stated, as seen from the base station antenna array, the source is no longer perceived as a point source, but rather as a spatially dispersed source with a mean DOA and a spatial extent. The main difficulty with scattered sources stems from the fact that for such a source, in contrast to the point source case, the signal subspace is no longer of dimension one. Hence, the popular subspace-based techniques cannot handle this problem satisfactorily unless proper modifications are introduced, see e.g. [1, 3] and references therein. Many methods presented so far for DOA estimation of scattered sources do not rely on the concept of signal/noise subspaces. Most frequently, they utilize a model of the data covariance matrix in either Maximum Likelihood (ML) estimators or covariance matching estimators (COMET). In [7], the ML estimator was derived. Since it entails solving a computationally burdensome 4-D maximization problem, a simpler optimally weighted covariance matching estimator was advocated whose performance is close to that of the MLE, but still requires a 2-D search over the DOA and the angular spread. In this paper, we consider simplifying the ML estimator by exploiting some special structure of the Fisher Information Matrix (FIM) for the problem at hand. We consider two models depending on whether the form of the azimuthal power distribution is explicitly known or not. For both models, the block-diagonal structure of FIM enables to decouple the estimation of the DOA from that of the other model parameters. An asymptotically efficient Maximum Likelihood (ML) DOA estimator is derived which entails solving a 1-D minimization problem only. Furthermore, the 1-D criterion can be expressed as a simple Fourier Transform.

2 DATA MODEL
The model used is the GAAO model developed in [8]. The received signal is

$$\mathbf{x}(t) = \mathbf{s}(t) \sum_{n=1}^{L} \gamma_n(t) \mathbf{a}(\theta_0 + \tilde{\theta}_n(t)) + \mathbf{n}(t)$$

(1)
where \( a(\theta) \) is the array response for a source with DOA equal to \( \theta \); \( \theta_0 \) is the mean DOA of the source, \( s(t) \) is the transmitted signal, \( \gamma_n(t) \) is the complex gain of ray \( n \) and \( \tilde{\theta}_a(t) \) is a random angular deviation. The noise \( n(t) \) in (1) is assumed to be a zero-mean complex Gaussian random vector with covariance matrix \( \sigma^2_n I \). It has been established \([7, 8]\) that \( x(t) \) is a zero-mean complex Gaussian vector, with covariance matrix

\[
R = E \{ x(t)x^H(t) \} = \int p(\tilde{\theta}; \sigma_{\theta}) a(\theta_0 + \tilde{\theta})a^H(\theta_0 + \tilde{\theta}) \, d\tilde{\theta} + \sigma^2_n I
\]

(2)

where \( P = \sigma^2 E \{ |s(t)|^2 \} \) denotes the source power (including the path gain factor) and \( p(\tilde{\theta}; \sigma_{\theta}) \) is the azimuthal power distribution. Generally, a model is assumed for \( p(\tilde{\theta}; \sigma_{\theta}) \) that is expressed in terms of a scalar parameter \( \sigma_{\theta} \), referred to as the angular spread. For instance, a Gaussian distribution was assumed in \([7]\) whereas a Laplacian distribution seemed to fit well the experimental measurements obtained in \([6]\). Characterization of \( p(\tilde{\theta}; \sigma_{\theta}) \) is of primary importance as the elements of \( R \) depend on it. The assumptions made have a strong impact as they must be taken into account in the estimation procedure. However, exact knowledge of \( p(\tilde{\theta}; \sigma_{\theta}) \) may not be available. Moreover, in order to improve robustness to mismodelling \( p(\tilde{\theta}; \sigma_{\theta}) \), an appealing feature would be that the DOA estimates do not rely on this knowledge. This is the route we pursue now.

Let us consider a uniform linear array with element separation \( \Delta \) in wavelength and introduce the variables \( \omega_0 = 2\pi \Delta \sin(\theta_0) \) and \( \sigma_{\omega} = 2\pi \Delta \cos(\theta_0) \sigma_{\theta} \). Assuming small angular spreads, the covariance matrix can be rewritten as \([3]\)

\[
R = P \Phi(\omega_0) B(\sigma_{\omega}) \Phi^H(\omega_0) + \sigma^2_n I
\]

(3)

where \( \Phi(\omega_0) = \text{diag}(1, e^{i\omega_0}, \ldots, e^{i(m-1)\omega_0}) \). If \( p(\tilde{\theta}; \sigma_{\theta}) \) is a symmetric distribution (which is a mild condition), \( B \) and hence \( \tilde{B} \) will be a real-valued symmetric Toeplitz (RST) matrix, independently of the form of the distribution. Consequently, we consider the two following models:

**Model 1** We assume a specific form for \( B \), e.g. \( B(\sigma_{\omega}) \), and the parameter vector is thus \( \theta = [\sigma^2_n \quad P \quad \sigma_{\omega} \quad \omega_0]^T \triangleq [\eta^T \quad \omega_0]^T \)

**Model 2** We do not attach any specific structure for \( B \), except that it is a RST matrix. The model becomes

\[
R = \Phi(\omega_0) \tilde{B} \Phi^H(\omega_0)
\]

(4)

where \( \tilde{B} \) is characterized by its first column \( \gamma = [\gamma_0, \ldots, \gamma_{m-1}]^T \) and hence \( \theta = [\gamma^T \quad \omega_0]^T \).

The second model does not require any assumption on \( p(\tilde{\theta}; \sigma_{\theta}) \). Hence, any estimate of \( \omega_0 \) based on this model will be robust to mismodelling the spatial distribution of the scatterers.

### 3 A COMPUTATIONALLY EFFICIENT APPROACH

The approaches proposed herein rely on the following result.

**Proposition 1.** Consider the problem of estimating a vector \( \theta \triangleq [\theta_1^T \quad \theta_2^T]^T \) and assume that the associated FIM is block-diagonal, i.e.

\[
F = \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix}
\]

where the partition of \( F \) corresponds to that of \( \theta \). Let \( \hat{\theta}_2 \) be a root-\( N \) consistent estimate of \( \theta_2 \) and let \( \Lambda(\hat{\theta}_1; \hat{\theta}_2) \) denote the Likelihood Function with \( \theta_2 \) replaced by \( \hat{\theta}_2 \).

Then,

\[
\hat{\theta}_1 = \arg \max_{\theta_1} \Lambda(\theta_1; \hat{\theta}_2)
\]

is an asymptotically (in \( N \)) efficient estimate of \( \theta_2 \).

**Proof.** see \([4]\) \( \square \)

In the case considered herein, it is possible to show \([3, 4]\) that the FIM for both models is block-diagonal so that the problem at hand falls in the framework of Result 1. Therefore, an asymptotically efficient estimate of \( \omega_0 \) can be obtained provided that a consistent estimate of \( \eta \) or \( \gamma \) can be derived. This is the aim of next result.

**Result 1.** Let \( \xi = [\xi_0 \ldots \xi_{m-1}]^T \triangleq [\xi_0 \quad P\xi]^T \) be the vector obtained by averaging the moduli of the elements on the sub-diagonals of \( R \), i.e.

\[
\xi_k = \frac{1}{m-k} \sum_{\ell=1}^{m-k} |R(k+\ell, \ell)| \quad (5)
\]

Also, let \( \tilde{R} = \frac{1}{N} \sum_{i=1}^{N} x(t)x^H(t) \) denote the sample covariance matrix and let \( \tilde{\xi} \) be constructed from the sample covariance matrix the way \( \xi \) is constructed from \( R \).

**Model 1** Consistent estimates of \( P, \sigma_{\omega} \) and \( \sigma^2_n \) can be obtained as

\[
\hat{P}, \hat{\sigma}_\omega = \arg \min_{P, \sigma_{\omega}} ||\tilde{\xi} - P\tilde{\xi}(\sigma_{\omega})||^2 \quad (6)
\]

\[
\hat{\sigma}^2_n = \frac{1}{m} \text{Tr} \{ \tilde{R} \} - \hat{P} \quad (7)
\]

**Model 2** A consistent estimate of \( \gamma \) is simply \( \hat{\gamma} = \tilde{\xi} \).
It follows from Proposition 1 and the block-diagonal structure of the FIM that an asymptotically efficient estimate of \( \omega \) can be obtained as the maximizer of the Likelihood Function with \( \eta \) (or \( \gamma \)) replaced by \( \hat{\eta} \) (or \( \hat{\gamma} \)), i.e.

\[
\hat{\omega}_0 = \begin{cases} 
\arg \max_{\omega} \Lambda(\omega; \hat{\eta}) & \text{for model 1} \\
\arg \max_{\omega} \Lambda(\omega; \hat{\gamma}) & \text{for model 2}
\end{cases}
\] (8)

Let \( \hat{B} \) denote the consistent estimate of \( B \) obtained from the estimates introduced in Result 1, for either model. Then, regardless of the model, one has (we omit the dependence on \( \hat{\eta} \) or \( \hat{\gamma} \))

\[
\Lambda(\omega) = -\ln \det \left\{ \Phi(\omega) \hat{B}^{-1} \Phi^H(\omega) \right\} \\
- \text{Tr} \left\{ \Phi(\omega) \hat{B} \Phi^H(\omega) \hat{R} \right\}
\] (9)

The first term in the right-hand side of the above equation does not depend on \( \omega \) since \( \Phi(\omega) \) is a unitary matrix. Hence, only the maximization of the second term is necessary. Moreover, the latter term can be expressed as

\[
\text{Tr} \left\{ \Phi(\omega) \hat{B}^{-1} \Phi^H(\omega) \hat{R} \right\} = \text{Tr} \left\{ \hat{B}^{-1} \otimes \hat{R} \right\} \]

\[
+ 2 \text{Re} \left\{ \sum_{n=1}^{m-1} \alpha_n e^{-in\omega} \right\}
\] (10)

where \( \otimes \) denotes the Hadamard matrix product and \( \alpha_n \) is obtained by summing along the \( n \)th sub-diagonal of the Hermitian matrix \( \hat{B} \otimes \hat{R} \):

\[
\alpha_n = \sum_{k=1}^{m-n} \hat{B}^{-1} (k + n, k) \hat{R}(k + n, k)
\] (11)

Therefore, we end-up with a very computationally efficient estimator which consists of minimizing the real part of the Fourier transform of an \( m \) length sequence, i.e.

\[
\hat{\omega}_0 = \arg \min_{\omega} \text{Re} \left\{ \sum_{n=1}^{m-1} \alpha_n e^{-in\omega} \right\}
\] (12)

It should be pointed out that the estimate of \( \omega_0 \) is given by (12) for both models. Only the way \( \hat{B} \) and hence \( \{\alpha_n\} \) are computed depends on the model. Finally, an estimate of \( \theta_0 \) is given by

\[
\hat{\theta}_0 = \arcsin \left( \frac{\hat{\omega}_0}{2\pi \Delta} \right)
\] (13)

Since the mapping between \( \hat{\theta}_0 \) and \( \hat{\omega}_0 \) is continuous and differentiable, it follows that \( \hat{\theta}_0 \) is also asymptotically efficient [5]. We stress, once again, the fact that the computational burden associated with the above estimation of \( \theta_0 \) is very low.

4 NUMERICAL EXAMPLES AND CONCLUSIONS

In this section the performance of the proposed estimators (referred to as ML-1D in what follows) is illustrated by means of Monte-Carlo simulations and also compared with the Cramér-Rao Bound (CRB) for model 1 [7]. We assume a uniform linear array with \( m = 8 \) elements separated by half wavelength. The source is located at \( \theta_0 = 10^\circ \), and the emitted signal \( s(t) \) is a QPSK sequence. The power distribution is assumed to be Gaussian with \( \sigma_\theta = 5^\circ \). The signal was generated according to (1) with \( L = 50 \) independent scatterers. The signal to Noise Ratio is defined as \( SNR = P/\sigma_n^2 \).

200 Monte-Carlo simulations were used to compute the root mean-square error (rmse) of the DOA estimates. All values are given in degrees (\(^\circ\)).

First, the performance of \( \theta_0 \) versus the number of snapshots is examined in Figure 2 for \( SNR = 5dB \). It can be observed that our simple estimators have a performance very close to the CRB, even in small sample. We note that the estimator based on model 2 has a performance quite close to that corresponding to model 1, except when the number of snapshots is small, typically \( N = 20, 40 \). This means that lack of knowledge on the exact form of \( p(\theta; \sigma_\theta) \) does not penalize too heavily the DOA estimation. This is a very important property which is worth re-iterating. We remind the reader that the estimate based on model 2 is fully robust to miss-modelling the azimuthal power distribution, which is an additional feature compared to using model 1. Since the performances are similar, the former should be preferred in any application where there is some uncertainty on the ray distribution. Note, however, that the use of

![Figure 2: Cramér-Rao Bounds and RMSE of the ML-1D estimators of \( \theta_0 \) versus the number of snapshots. \( m = 8 \), \( \theta_0 = 10^\circ \), \( \sigma_\theta = 5^\circ \) and \( SNR = 5dB \).](image-url)
model 2 does not provide directly an estimate of the angular spread, which may sometimes be useful. However, this is not a real drawback. Indeed, we can first obtain the model 2-based estimate of $\theta_0$ to guarantee robustness to possible mismodelling of the ray distribution. Then, if model 1 is deemed to be a “reasonable description” of the data and an estimate of $\sigma_\omega$ is required, one can obtain an estimate of $\sigma_\omega$ by non-linear least-squares fit over $\hat{\gamma}$. Alternatively, one can make use of the simple method suggested in [2] which relies on DOA estimates only.

Next, the influence on the performance of the number of sensors and angular spread is examined, see Figures 3 and 4 respectively. Again, the performance remains close to the CRB over a wide range of $m$ and angular spread values. Observe however that the estimate based on model 2 has a performance inferior to that obtained using model 1 whenever $m$ or $\sigma_\theta$ are large. Evidently, model 2 is less parsimonious as $m$ increases, which explains the difference in performance between the two models for large values of $m$. Of course, this is true when the data has been generated using model 1, as in the present case. In the perhaps more practical cases in which the data does not satisfy model 1 exactly, the estimate based on this model may be biased and it will most likely be outperformed by the model 2 based estimate.

5 References


