EUSIPCO, Tampere, Finland, 5-8 Sept, 2000

ML ESTIMATION OF SSR SIGNALS, IDENTIFIABILITY, AND CRAMER-RAO BOUNDS

Nicolas Petrochilos [1,2] *and Pierre Comon* [1]
(1) I3S–CNRS, Algorithmes, 2000 route des Lucioles, BP.121, F-06903 Sophia-Antipolis
(2) Dept. EE, Delft University of Technology, mekelweg 4, NL-2628 CD Delft
petro@ieee.org comon@i3s.unice.fr

ABSTRACT

In this article, a Maximum Likelihood method is applied to Secondary Surveillance Radar (SSR) signals impinging on an M-element antenna. At base-band, a SSR signal consists of a binary signal with alphabet $\{0,1\}$ multiplied by a complex exponential. Identifiability is proved in this binary case, and specific Cramer-Rao Bounds are derived and assessed.

1 Introduction

Secondary Surveillance Radar (SSR) is essential for Air Traffic Control (ATC). Unlike other radar systems, it is a communication radar that informs the airport about the identity and the altitude of the aircrafts in the line of sight. A new protocol (the mode S, for Selective) will address selectively the aircrafts and permit small communications between the station and the aircrafts. This new mode will also allow Traffic Advisory and Collision Avoidance System (TCAS) by automated communication between the aircrafts. A base station interrogates an aircraft, which answers via a device called transponder, the so-called SSR signal. It consist of a bursty train of pulses that are modulated on a frequency carrier of 1090 MHz.

Because each transponder is different, the local oscillator of the aircraft has a jitter, which can attain up to ± 3 MHz. After complex demodulation the binary signal is multiplied by a complex exponential. Moreover, the signals impinging on the antenna are not synchronized, the signal dynamic range is around 50 dB due to the range of the radar, the pulse and the train durations are different for the old and the new modes (pulse duration of $0.45\mu s$ for mode A and $0.5\mu s$ for mode S, and train duration of 20.5 for mode A and 64 or $112\mu s$ for S mode), and the carrier residual is quite smaller than the bandwidth of the binary signal. On the other hand, the SNR varies much and can be quite high ($0 \leq$ SNR ≤ 50), and there is no multipath.

In our work, based on a multi-static configuration, we want to perform multi-localization and to guarantee a safe data communication (as described in [1] and [2]). At the base station level, we want to recover the binary signals, the directions of arrival, and the carrier residuals. To reach this goal, we take advantage of the parameterization of the data: the array manifold, the carrier residual, and the fact that in a first estimation the binary data are almost always well estimated.

While similar problems are addressed in [3], and [4], it is usually considered either signals with no bandwidth, or such that $\text{BW}_i \ll f_j$.

2 Joint estimation of the parameters

2.1 Modeling

Considering no multipath, the observation model is, for d signals impinging on a M-element antenna:

$$\mathbf{x}(t) = \sum_{j=1}^{d} g_j \cdot \mathbf{a}(\theta_j)\, b_j(t-t_j)\, \exp\left(-j2\pi f_j (t-t_j)\right) + \mathbf{n}(t)$$

where t is the time, $1 \leq t \leq N$, t_j is the time of arrival of the signals, $b_j(t)$ is a binary signal, f_j is its carrier residual, g_j is its complex gain, and $\mathbf{a}(\theta)$ is the steering vector (of squared norm M). After sampling, and stacking N samples, this can be rewritten:

$$\begin{aligned} \mathbf{X} &= \mathbf{S} + \mathbf{N} \\ \text{with} \quad \mathbf{S} &= \mathbf{A} \cdot \mathbf{G} \cdot [\mathbf{F} \odot \mathbf{B}] \stackrel{\text{def}}{=} \mathbf{A} \cdot \mathbf{G} \cdot \mathbf{S}_{\oplus} \end{aligned} \quad (1)$$

where $\mathbf{X}$ is the $M \times N$ received signal matrix, $\mathbf{X} = [\mathbf{x}(1) \cdots \mathbf{x}(N)]$; $\mathbf{N}$ is the $M \times N$ noise matrix, $\mathbf{N} = [\mathbf{n}(1) \cdots \mathbf{n}(N)]$; $\mathbf{A}$ is the $M \times d$ steering matrix, $\mathbf{A} = [\mathbf{a}(\theta_1) \cdots \mathbf{a}(\theta_d)]$; $\odot$ is element-wise multiplication; $\mathbf{F}$ is the $d \times N$ frequency matrix, $\mathbf{F} = \left[\mathbf{f}_1^{\mathrm{T}} \cdots \mathbf{f}_d^{\mathrm{T}}\right]^{\mathrm{T}}$ with $\mathbf{f}_i = \left[1, \phi_i \cdots \phi_i^{N-1}\right]$ and $\phi_i = \exp(j2\pi f_i T_s)$; and $\mathbf{G} = |\mathbf{G}| \,.\, \Phi$ is the $d \times d$ diagonal gain matrix, whose diagonal entries are $g_i = |g_i| \exp(j\phi_i)$. For the sake of convenience, $\mathbf{S}_{\oplus}$ will denote the product $\mathbf{F} \odot \mathbf{B}$.

We have a spatially and temporally white Gaussian noise, so the total likelihood to maximize is:

$$p(\mathbf{X}/\Theta) = \left[\pi\sigma^2\right]^{-MN} \cdot \exp\{-\text{trace}\{\frac{[\mathbf{X} - \mathbf{S}(\Theta)]^{\mathrm{H}}[\mathbf{X} - \mathbf{S}(\Theta)]}{\sigma^2}\}\}$$

where $\Theta = \{\theta_i, g_i, f_i; 1 \leq i \leq d\}$ (because the t_i's are well approximated). Since the value of σ^2 is known (it origins from the thermal noise of the receiver), we may simplify the negative log-likelihood and summarize the problem as $\Theta = \arg\min_{\Theta\in\Omega} V$ where $V(\Theta) = \|\mathbf{X} - [\mathbf{A}\cdot\mathbf{G}\cdot[\mathbf{F}\odot\mathbf{G}]](\Theta)\|_{FRO}^2$ with $\|.\|_{FRO}$ denoting the Fröbenius norm.

2.2 Identifiability

In this section, the identifiability problem is addressed under two different assumptions: first DOA's are assumed to be different (space property), and then, alternatively, source signals are assumed statistically independent.

Proposition 2.1 *If the $\{\theta_i\}$ are all different, and if $M \geq 2\,d$, then parameters $\{\theta_i, g_i, f_i\}$ are uniquely identifiable.*

Proof: Assume there are two non-trivial solutions: Θ and $\tilde{\Theta}$. Then

$$E\{\mathbf{X}|\mathbf{B}\} = \mathbf{A}\cdot\mathbf{G}\cdot[\mathbf{F}\odot\mathbf{B}] = \tilde{\mathbf{A}}\cdot\tilde{\mathbf{G}}\cdot\left[\tilde{\mathbf{F}}\odot\tilde{\mathbf{B}}\right] \quad (2)$$

Suppose that the two sets of DOA's share some common θ_i for $\{1, d'\}$, thus we can form the following factors:

$$\left[\mathbf{A}_{nc},\ \mathbf{A}_c,\ \tilde{\mathbf{A}}_{nc}\right]\cdot\begin{bmatrix} \mathbf{G}_{nc}\cdot[F\odot B]_{nc} \\ \mathbf{G}_c\cdot\mathbf{S}_{\oplus c} - \mathbf{P}_1\cdot\tilde{\mathbf{G}}_c\cdot\tilde{\mathbf{S}}_{\oplus c} \\ -\tilde{\mathbf{G}}_{nc}\cdot\mathbf{P}_2\cdot\left[\tilde{F}\odot\tilde{B}\right]_{nc}\end{bmatrix} = 0 \quad (3)$$

where $\mathbf{P}_j$ is a permutation matrix that exchanges the rows of $\tilde{\mathbf{S}}_\oplus$ in order to to factorize equation 2. The first matrix in the previous equation is full column rank (tall or square Vandermonde matrix with different θ's, with $M \geq 2d$), therefore the second matrix is equal to zero. It is impossible that non trivial binary signals multiplied by a complex exponential are null, so there were only common DOA's. Thus for $i = 1, d$:

$$g_i\cdot(\mathbf{f}_i\odot\mathbf{b}_i) = \tilde{g}_{\beta i}\cdot\left(\tilde{\mathbf{f}}_{\beta i}\odot\tilde{\mathbf{b}}_{\beta i}\right) \quad (4)$$

where the β_i depends on the permutation matrix $\mathbf{P}_1$. By taking the absolute value of equation 4, it is clear that $|g_i| = |\tilde{g}_{\beta i}|$ and that $\tilde{\mathbf{b}}_{\beta i} = \mathbf{b}_i$, otherwise equation |4| would lead to the absurd result $0 = 1$.

Now it remains to prove that, for any fixed i, f_i and ϕ_i coincide with $\tilde{f}_i$ and $\tilde{\phi}_i$. Let $g_i = |g_i|\,.\,e^{\phi_i}$ and denote n and $\tilde{n}$ two time values belonging to the support of $b_i(n)$. Then $\exp[j(2\pi n(f-\tilde{f})T + (\phi-\tilde{\phi}))] = 1$ yields that $(f-\tilde{f}) = (k-\tilde{k})/(n-\tilde{n})T$, for some pair of integers $(k,\tilde{k})$, and $\phi - \tilde{\phi} = (k+\tilde{k})\pi$. □

Lemma 2.2 *If frequencies f_i are all equal, the matrix $\mathbf{S}_\oplus = \mathbf{F}\odot\mathbf{B}$ is full row rank, for sufficiently large N and for ergodic and uncorrelated sequences $b_i(n)$.*

Proof: We shall prove that the matrix $R_{ss} = \frac{1}{N}\mathbf{S}_\oplus\mathbf{S}_\oplus^{\mathrm{H}}$ is full rank. Its generic term is:

$$r_{k\ell} = \frac{1}{N}\sum_{n=0}^{N} b_k(n)b_\ell(n)\exp[(2\pi j(f_k - f_\ell)nT_s]. \quad (5)$$

For equal frequencies, the entries reduce to $\frac{1}{N}\sum_{n=0}^{N} b_k(n)b_\ell(n)$, which converge to the product of the sample means, namely $\frac{1}{4}$ for $k \neq \ell$ or $\frac{1}{2}$ for $k = \ell$. As a consequence, R_{ss} converges to $\frac{1}{4}\mathbf{1}\mathbf{1}^{\mathrm{T}} + \frac{1}{4}\mathbf{I}$, which is full rank. □

Lemma 2.3 *If frequencies f_i are all distinct, the matrix $\mathbf{S}_\oplus = \mathbf{F}\odot\mathbf{B}$ is full row rank, for sufficiently large N and for stationary $\mathbf{b}(n)$.*

Proof: Consider the generic term (5) with $k \neq \ell$, and denote $\rho_{k\ell} = \mathrm{E}\{b_k(n)b_\ell(n)\}$. Taking the expectation leads to $\mathrm{E}\{r_{k\ell}\} = \frac{1}{N}\rho_{k\ell}\sum_n \exp[(2\pi j(f_k - f_\ell)nT_s]$. Provided that $(f_k - f_\ell)T_s$ is not integer, this quantity tends to zero as N tends to infinity. As a conclusion, only diagonal terms of R_{ss} are non zero, which proves that $\mathbf{S}_\oplus$ tends in the mean square sense to a full rank matrix. □

Proposition 2.4 *If time series $b_i(n)$ are ergodic and uncorrelated, then for sufficiently large N, the matrix $\mathbf{S}_\oplus$ is full row rank.*

Proof: It suffices to partition the matrix R_{ss} defined in the previous lemmas into sets, each associated with distinct frequencies. Then it takes asymptotically a block-diagonal form because of lemma 2.3, and its diagonal blocks are given by lemma 2.2. Thus, matrix $\mathbf{S}_\oplus$ tends in the mean square sense to a full rank matrix. □

Proposition 2.5 *For sufficiently large observation duration N, if the source time series b_i are stationary ergodic and uncorrelated, then the parameters $\{\theta_i, g_i, f_i\}$ are uniquely identifiable.*

Proof: Assume two different non-trivial solutions Θ and $\tilde{\Theta}$. Then using equation (2) in a similar fashion as equation (3), we get:

$$\left[(\mathbf{AG})_{nc},\ (\mathbf{AG})_c - (\tilde{\mathbf{A}}\tilde{\mathbf{G}})_c,\ -(\tilde{\mathbf{A}}\tilde{\mathbf{G}})_{nc}\right]\cdot\begin{bmatrix}(\mathbf{F}\odot\mathbf{B})_{nc}\\(\mathbf{F}\odot\mathbf{B})_c\\(\tilde{\mathbf{F}}\odot\tilde{\mathbf{B}})_{nc}\end{bmatrix} = \mathcal{A}_{\mathcal{G}}\cdot\mathcal{S}_\oplus = 0 \quad (6)$$

where subscript $(.)_c$ denotes the rows that share the same frequencies and sequences ($\mathbf{F}$ and $\mathbf{B}$). Matrix $\mathcal{S}_\oplus$ is a wide matrix, which fulfills the requirements of proposition 2.4; then, as the time of integration tends to infinity, it becomes full row rank, which means that $\mathcal{A}_{\mathcal{G}}$ is equal to the zero matrix. Since the non-common part of $\mathcal{A}_{\mathcal{G}}$ has entries with a non zero norm, there were only common $\mathbf{S}_\oplus$. Additionally, we have $(\mathbf{A}\cdot\mathbf{G})_c - (\tilde{\mathbf{A}}\cdot\tilde{\mathbf{G}})_c = 0$; the first row proves the equality of the gain matrices $\mathbf{G}$ and $\tilde{\mathbf{G}}$. Then the equality of $\mathbf{A}$ and $\tilde{\mathbf{A}}$ leads, because of non ambiguity of the array considered, to equality of the DOA's θ_i. The proposition is proved. □

The previous results are of asymptotic nature, and may not hold valid for limited observation lengths. Therefore, we propose another result:

Lemma 2.6 *The probability that a column of a $d \times N$ matrix B with iid binary entries takes the form $[* \cdots * \; 1 \; \mathbf{0}_{q-1}]^{\mathrm{T}}$ is: $p(q,N) = 1-(1-2^{-q})^N$*

Proof: Only the last q entries of the column of size d are imposed, so that the probability to draw such a vector is 2^{-q}. After N realizations, the probability not to draw such a vector is thus $(1-2^{-q})^N$. □

Proposition 2.7 *The probability $P(d,N)$ that the $d \times N$ binary matrix B is full rank is bounded by:*

$$[1-(1-2^{-d})^N]^d \leq P(d,N) \tag{7}$$

Proof: A sufficient condition for matrix B to be full rank is that there exist d columns building together an upper triangular matrix with ones along the diagonal. From lemma 2.6, the probability that such columns exist is: $\prod_{q=1}^{d} p(q,N)$. Note that this probability is already a lower bound; in fact, this probability is related to the event that an upper triangular submatrix exists, without even permuting the rows. The advantage of still keeping this loose bound lies in its simplicity. The result of the proposition is eventually obtained by bounding every term of the product. □

Following the same reasoning than in the two first propositions of the paper and using the above proposition, identifiability with finite time of integration can be derived.

2.3 Cramer-Rao Bounds

The CRB is derived for the specific case of known binary signals (matrix $\mathbf{B}$ is known), or for the case of a stochastic $\mathbf{B}$ matrix. In [4], one can find the CRB for similar problems, which differ only by the absence of the binary matrix.

For the sake of simplicity, we consider the real parameter $\Theta' = \{\theta_i, |g_i|, \phi_i, f_i; 1 \leq i \leq d\}$. Define:

$$\mathbf{D}_k(\Theta) = \left[\frac{\partial \mathbf{s}(k,\Theta)}{\partial \Theta_1} \cdots \frac{\partial \mathbf{s}(k,\Theta)}{\partial \Theta_{4d}}\right]$$

then the CRB is the inverse of the Fisher Information matrix (FIM), which is:

$$\mathcal{FIM}(\Theta) = \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{k=1}^{N} \mathbf{D}_k^H(\Theta)\mathbf{D}_k(\Theta)\right\}$$

Proposition 2.8 *In case of known binary signal, the FIM is:*

$$\mathcal{FIM}_\alpha(\Theta) = \frac{2}{\sigma^2}\begin{bmatrix} \mathbf{I}_{\theta\theta} & \mathbf{I}_{\theta f} & \mathbf{I}_{\theta|g|} & \mathbf{I}_{\theta\phi} \\ \mathbf{I}_{f\theta}^T & \mathbf{I}_{ff} & \mathbf{I}_{f|g|} & \mathbf{I}_{f\phi} \\ \mathbf{I}_{|g|\theta}^T & \mathbf{I}_{|g|f}^T & \mathbf{I}_{|g||g|} & \mathbf{I}_{|g|\phi} \\ \mathbf{I}_{\phi\theta}^T & \mathbf{I}_{\phi f}^T & \mathbf{I}_{\phi|g|}^T & \mathbf{I}_{\phi\phi} \end{bmatrix} \tag{8}$$

where:

$$\mathbf{I}_{\theta\theta} = \mathrm{Re}\left\{(\mathbf{G}.\mathbf{S}_\oplus.\mathbf{S}_\oplus^H.\mathbf{G}^H) \odot (\mathbf{D}_A^H.\mathbf{D}_A)^*\right\}$$
$$\mathbf{I}_{ff} = \mathrm{Re}\left\{(\mathbf{G}.(C_T \odot \mathbf{S}_\oplus).(C_T \odot \mathbf{S}_\oplus)^H\mathbf{G}^H) \odot \mathbf{R_A}^*\right\}$$
$$\mathbf{I}_{\theta f} = \mathrm{Re}\left\{(\mathbf{G}.\mathbf{S}_\oplus.(C_T \odot \mathbf{S}_\oplus)^H\mathbf{G}^H) \odot (\mathbf{D}_A^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{|g||g|} = \mathrm{Re}\left\{(\Phi.\mathbf{S}_\oplus.\mathbf{S}_\oplus^H.\Phi^H) \odot (\mathbf{A}^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{\theta|g|} = \mathrm{Re}\left\{(\mathbf{G}.\mathbf{S}_\oplus.\mathbf{S}_\oplus^H.\Phi^H) \odot (\mathbf{D}_A^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{f|g|} = \mathrm{Re}\left\{\left[(\mathbf{G}.(C_T \odot \mathbf{S}_\oplus)).(\Phi.\mathbf{S}_\oplus)^H\right] \odot (\mathbf{A}^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{\phi\phi} = \mathrm{Re}\left\{\left[(\mathbf{G}.\mathbf{S}_\oplus).(\mathbf{G}.\mathbf{S}_\oplus)^H\right] \odot (\mathbf{A}^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{\theta\phi} = \mathrm{Re}\left\{j\left[(\mathbf{G}.\mathbf{S}_\oplus).(\mathbf{G}.\mathbf{S}_\oplus)^H\right] \odot (\mathbf{D}_A^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{f\phi} = \mathrm{Re}\left\{j\left[(\mathbf{G}.(C_T \odot \mathbf{S}_\oplus)).(\mathbf{G}.\mathbf{S}_\oplus)^H\right] \odot (\mathbf{A}^H.\mathbf{A})^*\right\}$$
$$\mathbf{I}_{|g|\phi} = \mathrm{Re}\left\{j\left[(\Phi.\mathbf{S}_\oplus).(\mathbf{G}.\mathbf{S}_\oplus)^H\right] \odot (\mathbf{A}^H.\mathbf{A})^*\right\}$$

with $\mathbf{D}_A = \left[\frac{\partial \mathbf{a}}{\partial \theta_1}(\theta_1) \cdots \frac{\partial \mathbf{a}}{\partial \theta_d}(\theta_d)\right]$, $\mathbf{G} = \mathbf{G}'.\Phi$, $\mathbf{R_A} = \mathbf{A}^H.\mathbf{A}$, $C_T = j2\pi T_S\,[1 \cdots 1]^T.[0\,1 \cdots N-1]$, *and* $\mathbf{S}_\oplus = \mathbf{F} \odot \mathbf{B}$.

The proof is skipped to the appendix. In [3, 4], the signals are only complex sinusoids, thus it is a simplification of our model with $\mathbf{B}$ as a matrix of ones. Then for this case, $\mathbf{B}$ can be removed from the above equations, and $\mathbf{F}$ can replace by $\mathbf{S}_\oplus$, let us denominate the resulting Fisher Information Matrix as $\mathcal{FIM}_\delta$, and state:

Proposition 2.9 *Considering unknown independent binary signals as unwanted parameters, the Fisher Information Matrix is:*

$$\mathcal{FIM}_\beta(\Theta) = \mathcal{FIM}_\delta \odot C_B^D \tag{9}$$

with $C_B^D = (\mathbf{1}_{4\times 1}.\mathbf{1}_{1\times 4}^T) \otimes C_B^{(d)}$.

The proof is skipped for the journal article, the only assumption of this proposition is that $\mathbf{A}.\mathbf{G}$ has all its singular values above several times the noise standard deviation, which suit our data model. One can notice that $E\{\mathcal{FIM}_\alpha(\Theta)\} = \mathcal{FIM}_\beta(\Theta)$. For both CRB, the complexity is of the order of $d^2 \times N$ (given $d^2 \leq N$).

3 Simulations

3.1 Identifiability

In order to check the validity of proposition 2.6, we estimated for different d the required N for identifiability with various probability $P_{id.}$ (resp. 0.9, 0.99, and 0.999). For each d, we ran one million independent realizations of $\mathbf{B} \odot \mathbf{F}$. Figure 1 display the results, there is also the average number of samples, and the number of sources. As supposed, the bound (7) is too large; for low value of d, N seems to be bounded by a constant (resp. 10, 12, and 15), then N seems to tends to a multiple of d as d get larger. For the number of samples we have to deal (from 100 to 4096), it is almost assured that we have identifiability.

3.2 ML and CRB

In this article, we are using a gradient-descent algorithm:

$$\mathbf{\Theta}_{k+1} = \mathbf{\Theta}_k - \mu_k \cdot \frac{\partial V(\mathbf{\Theta}_k)}{\partial \mathbf{\Theta}_k} \tag{10}$$

The initial point, Θ_0 is chosen on the basis of an ESPRIT, then after filtering, a phase-shift estimation over the time gives the f_i's, and a thresholding gives the $\mathbf{b}_i$'s.

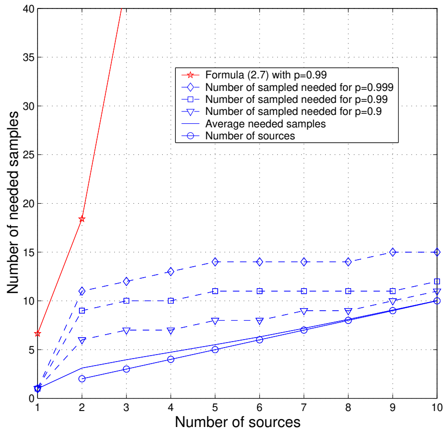

Figure 1: *Approximated and simulated $P_{id.}$.*

We used in the simulations a 8 elements ULA with half wavelength spacing, the basic scenario involve 3 sources with power $[200, 50, 125]$ ($[46, 34, 42]$ dB SNR), coming from the directions $[90 - \delta\theta, 90, 90 + \delta\theta]$, and having the remaining frequency $[\delta f, 0, -\delta f]$. The sampling frequency is 20 MHz, and the number of samples is 100. In one scenario we change $\delta\theta$, in the other δf; For both simulations, we present the CRB, the mean square error (MSE) for the ML with the solution as an initial point (ML1), and the MSE for the ML with the initial point described above (ML2). We run 1000 independent realizations to get the MSE.

In the first simulation, we fix $\delta f = 10^4$, while $\delta\theta$ varies. As Figure 2 displays, ML1 follows quite well the CRB, while ML2 has one order of difference. This comes from a wrong initial detection of parameters, or bits; one error leading to a local minimum, the other to non-optimal minimization. At low angle, ML2 is definitively wrong due to the first estimation by ESPRIT. Notice that ML1 can endure it, it demonstrates that indeed, identifiability holds! So, if we know the solution as initial point, ML1 can find it !!

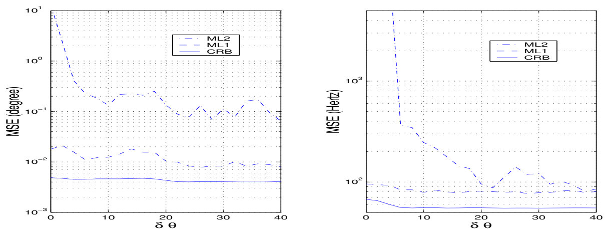

Figure 2: *MSE and CRB of the second signal.*

In the second simulation, we fix $\delta\theta = 20$ and δ varies (see Figure 3). In this case, while ML1 and (ML2,CRB) have an order of difference for the DOA MSE, it is interesting to notice that now for the frequency MSE, all the curve are together. Notice also that as δf decrease, the ML2 performance does not degrade as in the previous simulation, it is because the ESPRIT step perform well, then the frequency estimation is less sensible to other signals.

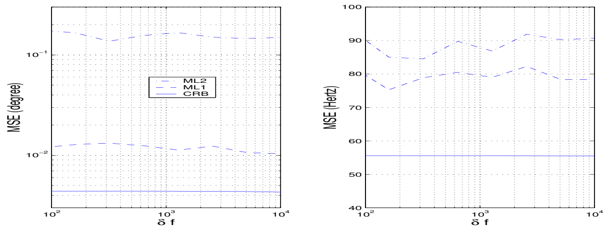

Figure 3: *MSE and CRB of the second signal.*

4 Conclusions

The problem of identifying source amplitudes, directions of arrival, and carrier residuals, has been addressed via a Maximum Likelihood approach, when sources are *binary*. Identifiability conditions as well as Cramér-Rao bounds have been derived. The Maximum Likelihood algorithm is not optimum, but sufficient to assess the CRB, and demonstrate the identifiability.

References

[1] N. Petrochilos, P. van Genderen, "A New Approach to Handle SSR Replies", in *Radar'99*, Brest, France, 17-21 May 1999.

[2] P. Bezousek, "A passive radar surveillance system VERA for ATC", in *IRS'98*, Munich, Germany, 1998.

[3] A. K. Shaw, R. Kumaresan, "Frequency-Wavenumber Estimation By Structured Matrix Approximation", in *Spectrum Estimation workshop*, Boston, MA, 1987.

[4] Aweke N. Lemma, A. J. van der Veen, E.F. Deprettre, "Analysis of Joint Angle-Frequency Estimation Using ESPRIT", *IEEE trans. on Acoustics, speech, and signal processing*, 1999, to appears.

Appendix

Proof of Proposition 2.8 In order to save space, and since the same mathematical properties apply for all submatrix $\mathbf{I}_{**}$, the derivations are given only for one, $\mathbf{I}_{\theta|g|}$, the other submatrices have equivalent transformations.

$$\begin{aligned}
&\sum_{k=1}^{N} \left(\frac{\partial \mathbf{s}(k)}{\partial \theta_i}\right)^H \left(\frac{\partial \mathbf{s}(k)}{\partial |g_i|}\right) \\
&= \left(\frac{\partial \mathbf{a}(\theta_i)}{\partial \theta_i}\right)^H \mathbf{a}(\theta_j) \,.\, \sum_{k=1}^{N} \left(g_i \phi_i^k b_i(k)\right)^* \psi_j \phi_j^k b_j(k) \\
&= \left(\frac{\partial \mathbf{a}(\theta_i)}{\partial \theta_i}\right)^H \mathbf{a}(\theta_j) \,.\, g_i^* [\mathbf{F} \odot \mathbf{B}]_i^* [\mathbf{F} \odot \mathbf{B}]_j^T \psi_j \\
&= \left(\frac{\partial \mathbf{A}}{\partial \theta}^H \mathbf{A}(\theta)\right)_{ij} \,.\, \left(\mathbf{G}^* [\mathbf{F} \odot \mathbf{B}]^* [\mathbf{F} \odot \mathbf{B}]^T \Psi\right)_{ij}
\end{aligned}$$

which demonstrates:

$$\mathbf{I}_{\theta|g|} = \frac{2}{\sigma^2} \mathrm{Re} \left\{ \left(\frac{\partial \mathbf{A}}{\partial \theta}^H \mathbf{A}(\theta) \right)^* \odot \left(\mathbf{G} \mathbf{S}_{\oplus} \mathbf{S}_{\oplus}^H \Psi^H \right) \right\} . \square$$