APPROXIMATE ML DIRECTION FINDING IN SPATIALLY CORRELATED NOISE USING OBLIQUE PROJECTIONS

Mustapha Djeddou, Adel Belouchrani* and Saïd Aouada***

*Electrical Engineering Department, École Nationale Polytechnique (ENP),
PO Box 182, El Harrach 16200, Algiers, Algeria
**Signal Processing Group, Darmstadt University of Technology
Merckstrasse 25, D-64283, Darmstadt, Germany.

ABSTRACT

We consider the problem of Maximum Likelihood estimation of Directions of Arrival of multiple source signals in the presence of unknown spatially correlated Gaussian noise. Oblique Projections are used to separate the structured noise from the signal and an Approximate Maximum Likelihood solution is derived. The estimates are obtained by maximizing the modified cost function using a nonlinear optimization technique. Numerical simulations are provided to assess the performance of the proposed approach. Simulations include comparison to the Stochastic Maximum Likelihood and to the Weighted Subspace Fitting, as well as to the Cramér-Rao Bound.

1. INTRODUCTION

Direction Of Arrival (DOA) estimation of multiple narrow-band sources is well addressed in array signal processing. In the presence of spatially white Gaussian noise, many estimation techniques have been developed [1, 2]. However, when the noise is not spatially white, the classical techniques can be applied only if the data is prewhitened by measuring the spatial noise covariance matrix, under the conditions of large data sizes, high SNR and a stationary noise covariance matrix. Failure to fulfill these conditions results in highly biased estimates. When prewhitening is not possible, the noise is modeled as a spatially dependent process with an unknown covariance matrix. In this paper, we consider the problem of DOA estimation in the presence of unknown spatially correlated noise, modeled as a combination of a structured and an unstructured process [3]. An Approximate Maximum Likelihood (AML) solution is formulated where the dimension of the optimization problem is reduced through the estimation of the unknown noise parameters using Oblique Projections (OP). It is shown in Section 3 that the range and null space of the considered OP are the unstructured noise and signal with structured noise subspaces respectively.

2. DATA MODEL

Consider an array with K sensors, receiving d impinging source signals from respective directions of arrival $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_d]^T$, where $(.)^T$ denotes matrix transpose. The sensor data output is written in vector form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t); \quad t = 1, \cdots, T \tag{1}$$

where

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{a}(\theta_1) & \cdots & \mathbf{a}(\theta_d) \end{bmatrix}$$

is the array spatial steering matrix, with $\mathbf{a}(\theta_i)$ being the array response to a path impinging from direction θ_i. It is assumed that the parameterization of the array response vectors is known and that no ambiguities are introduced in the manifold.

The signal waveform $\mathbf{s}(t)$ is modeled as a stationary stochastic process, independent of $\mathbf{n}(t)$, with zero mean and covariance matrix $\mathbf{S}$. The noise $\mathbf{n}(t)$ is modeled as a combination of two parts. The first part, $\mathbf{w}(t)$ is an unstructured noise generated internally by the electronics in the receiver. The second part is an external structured noise $\mathbf{z}(t)$, resulting from processing an unknown process $\mathbf{v(t)}$ through a known linear system $\mathbf{B} \in \mathbb{C}^N$. This linear process can be regarded as a set of N base matrices spanning the noise subspace (structured). The dimension of $\mathbf{B}$ is therefore $K \times N$, with $K \geq N$. Hence we can model the additive noise as

$$\begin{aligned} \mathbf{n}(t) &= \mathbf{z}(t) + \mathbf{w}(t) && (2) \\ &= \mathbf{B}\mathbf{v}(t) + \mathbf{w}(t) && (3) \end{aligned}$$

Also, the following assumptions on the noise are considered

$$\begin{aligned} \mathsf{E}[\mathbf{n}(t)\mathbf{n}^H(t)] &= \mathbf{Q}\delta_{tk} \\ \mathsf{E}[\mathbf{n}(t)\mathbf{n}^T(t)] &= 0 \end{aligned} \tag{4}$$

where $\mathsf{E}(.)$ denotes expectation, $(.)^H$ denotes Hermitian transpose and $\mathbf{Q}$ is the spatial noise covariance matrix which, in the general case, can be modeled as

$$\mathbf{Q} = \begin{bmatrix} q_0 & q_1 & \cdots & q_{K-1} \\ q_1^H & q_0 & \cdots & q_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K-1}^H & q_{K-2}^H & \cdots & q_0 \end{bmatrix} \tag{5}$$

where $q_0 = \sigma^2$ is the unstructured noise power. The $(2K-1)$-dimensional vector of unknown noise parameters is therefore given as $\mathbf{q} = [q_0, \Re(q_1), \Im(q_1), \ldots, \Re(q_{K-1}), \Im(q_{K-1})]^T$.

The data covariance matrix can be written as

$$\mathbf{R} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \mathbf{Q} \tag{6}$$

Note that the dependence of $\mathbf{A}$ on $\boldsymbol{\theta}$ is omitted for simplicity. The problem at hand is to estimate the parameters $\boldsymbol{\theta}$ from the collected data, using the known structure of the noise.

3. APPROXIMATE ML ESTIMATION

In (1), the collected data $\mathbf{x}(t)$, $t = 1, \ldots, T$ is modeled as a zero mean random processes with covariance matrix (6). Thus, the joint density function of the data is given by [4]

$$f_{\boldsymbol{\eta}} = (2\pi)^{-\frac{T}{2}} \det\left\{\mathbf{R}^{-\frac{T}{2}}(\boldsymbol{\eta})\right\} \exp\left\{-\frac{T}{2}\text{trace}\left[\mathbf{R}^{-1}(\boldsymbol{\eta})\hat{\mathbf{R}}\right]\right\} \quad (7)$$

where $\boldsymbol{\eta} = [\boldsymbol{\theta}^T, \mathbf{p}^T, \mathbf{q}^T]^T$ is the vector of unknown parameters, with $\mathbf{p} = [\Re(\mathbf{S}), \Im(\mathbf{S})]$ and $\hat{\mathbf{R}}$ is the sample covariance matrix of the data, given by

$$\hat{\mathbf{R}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}(t)\mathbf{x}^H(t) \quad (8)$$

After normalization and omitting constant terms, it can be easily shown that the stochastic negative Log-Likelihood (LL) function of the observed data is [4]

$$\mathcal{L}(\boldsymbol{\eta}) = \ln\{\det[\mathbf{R}(\boldsymbol{\eta})]\} + \text{trace}\left\{\mathbf{R}^{-1}(\boldsymbol{\eta})\hat{\mathbf{R}}\right\} \quad (9)$$

After some straightforward manipulations, the ML estimate of the covariance matrix $\mathbf{S}$ is given by [5]

$$\hat{\mathbf{S}}(\boldsymbol{\theta}, \mathbf{q}) = (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}(\hat{\mathbf{R}} - \mathbf{Q})\mathbf{Q}^{-1}\mathbf{A}(\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1} \quad (10)$$

and the insertion of (10) in (6) leads to the following concentrated expression

$$\mathbf{R}(\boldsymbol{\theta}, \mathbf{q}) = \boldsymbol{\Pi}(\boldsymbol{\theta}, \mathbf{q})(\mathbf{R} - \mathbf{Q})\boldsymbol{\Pi}^H(\boldsymbol{\theta}, \mathbf{q}) + \mathbf{Q} \quad (11)$$

where $\boldsymbol{\Pi}(\boldsymbol{\theta}, \mathbf{q}) = \mathbf{A}(\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{Q}^{-1}$. Note that using the concentrated expression of the data covariance matrix (11) in the LL function (9) reduces the dimension of the optimization search. However this direct optimization remains unattractive as the dependence between the unknown parameters involves high nonlinearities. At this stage, we seek a further simplified cost function by replacing $\mathbf{Q}$ by a favorable estimate $\hat{\mathbf{Q}}$. Note that since the structured noise part is uniquely parametrized by a given process $\mathbf{B}$, matrix $\mathbf{Q}$ exhibits a unique dependence on $\mathbf{B}$. As mentioned previously, this process can be seen as the result of a linear expansion of the noise subspace and truncation after the N most significant terms, with the truncation error being taken care of in the unstructured part whose asymptotic properties are known.

Let the data vector $\mathbf{x}(t)$ defined in (1) be written in the following form

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{A}\mathbf{s}(t) + \mathbf{B}\mathbf{v}(t) + \mathbf{w}(t) \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{s}(t) \\ \mathbf{v}(t) \end{bmatrix} + \mathbf{w}(t) \end{aligned} \quad (12)$$

Provided that the composite matrix $\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}$ has full rank, and $d + N \leq K$, the zero forcing solution of equation (12) is given by

$$\begin{bmatrix} \hat{\mathbf{s}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}^{\dagger} \mathbf{x}(t) \quad (13)$$

where $(.)^{\dagger}$ stands for Moore-Penrose pseudo-inverse. Using matrix properties, equation (13) can be rewritten as

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{s}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix} &= \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{B} \\ \mathbf{B}^T\mathbf{A} & \mathbf{B}^T\mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} \mathbf{x}(t) \\ &= \mathbf{H}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} \mathbf{x}(t) \end{aligned} \quad (14)$$

Recalling that the LL function is to be concentrated with respect to the noise parameters, we focus on the estimation of $\hat{\mathbf{v}}(t)$. Applying the inversion formula of 2×2-block matrices, $\mathbf{H}^{-1}$ can be expressed as

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_3 & \mathbf{H}_4 \end{bmatrix} \quad (15)$$

with

$$\begin{aligned} \mathbf{H}_1 &= \mathcal{A}^{-1} - \mathcal{A}^{-1}\mathcal{C}\left[\mathcal{C}^T\mathcal{A}^{-1}\mathcal{C} - \mathcal{B}\right]^{-1\,T}\mathcal{A}^{-1} \\ \mathbf{H}_2 &= \mathcal{A}^{-1}\mathcal{C}\left[\mathcal{C}^T\mathcal{A}^{-1}\mathcal{C} - \mathcal{B}\right]^{-1} \\ \mathbf{H}_3 &= \left[\mathcal{C}^T\mathcal{A}^{-1}\mathcal{C} - \mathcal{B}^T\right]^{-1}\mathcal{C}^T\mathcal{A}^{-1} \\ \mathbf{H}_4 &= -\left[\mathcal{C}^T\mathcal{A}^{-1}\mathcal{C} - \mathcal{B}\right]^{-1} \end{aligned}$$

where $\mathcal{A} = \mathbf{A}^T\mathbf{A}$, $\mathcal{B} = \mathbf{B}^T\mathbf{B}$ and $\mathcal{C} = \mathbf{A}^T\mathbf{B}$. Thus, $\hat{\mathbf{v}}(t)$ is obtained as

$$\hat{\mathbf{v}}(t) = (\mathbf{H}_3\mathbf{A}^T + \mathbf{H}_4\mathbf{B}^T)\mathbf{x}(t) \quad (16)$$

or equivalently, using the expressions of $\mathbf{H}_3$ and $\mathbf{H}_4$ and after applying some algebraic manipulations,

$$\hat{\mathbf{v}}(t) = (P_{\mathbf{A}}^{\perp}\mathbf{B})^{\dagger}\mathbf{x}(t) \quad (17)$$

where $P_{\mathbf{A}}^{\perp}$ is the orthogonal projector onto $\mathbf{A}$.

Using the known structure $\mathbf{B}$, the estimated structured noise is

$$\begin{aligned} \hat{\mathbf{z}}(t) &= \mathbf{B}\hat{\mathbf{v}}(t) \\ &= \mathbf{B}(P_{\mathbf{A}}^{\perp}\mathbf{B})^{\dagger}\mathbf{x}(t) \\ &= E_{\mathbf{BA}}\mathbf{x}(t) \end{aligned} \quad (18)$$

where $E_{\mathbf{BA}}$ is the Oblique Projector (OP) with range $\langle\mathbf{B}\rangle$ and null space $\langle\mathbf{A}\rangle$, [6]. Similarly, the OP with range $\langle\mathbf{A}\rangle$ and null space $\langle\mathbf{B}\rangle$ is given as

$$E_{\mathbf{AB}} = \mathbf{A}(P_{\mathbf{B}}^{\perp}\mathbf{A})^{\dagger} \quad (19)$$

By neglecting the component of the unstructured noise belonging to the subspace $\langle\mathbf{A}\rangle$ [1]. $\mathbf{w(t)}$ can be considered to belong to the subspace $\langle\mathbf{N}\rangle$ that is orthogonal to both subspaces $\langle\mathbf{A}\rangle$ and $\langle\mathbf{B}\rangle$. This noise subspace can be estimated using the following algebraic property $\mathbf{I} = P_{\mathbf{N}} + P_{\mathbf{AB}}$, where $P_{\mathbf{AB}} = E_{\mathbf{AB}} + E_{\mathbf{BA}}$, is the projector onto the space shared by $\langle\mathbf{A}\rangle$ and $\langle\mathbf{B}\rangle$ and $P_{\mathbf{N}}$ is the projector onto the subspace $\langle\mathbf{N}\rangle$. Thus, exploiting this orthogonality, an estimate of $\mathbf{w}(t)$ can be obtained as

$$\begin{aligned} \hat{\mathbf{w}}(t) &= P_{\mathbf{N}}\mathbf{x}(t) \\ &= (\mathbf{I} - P_{\mathbf{AB}})\mathbf{x}(t) \\ &= P_{\mathbf{AB}}^{\perp}\mathbf{x}(t) \end{aligned} \quad (20)$$

and the estimated noise covariance matrix is

$$\begin{aligned} \hat{\mathbf{Q}}(\boldsymbol{\theta}) &= \mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} + \mathbf{R}_{\hat{\mathbf{w}}\hat{\mathbf{w}}} \\ &= E_{\mathbf{BA}}\mathbf{R}E_{\mathbf{BA}}^H + P_{\mathbf{AB}}^{\perp}\mathbf{R}P_{\mathbf{AB}}^{\perp H} \end{aligned} \quad (21)$$

Under asymptotic conditions, replacing the covariance data matrix $\mathbf{R}$ by the sample covariance matrix $\hat{\mathbf{R}}$ and inserting (21) in (10), we obtain the following approximation

$$\tilde{\mathbf{R}}(\boldsymbol{\theta}) = \hat{\boldsymbol{\Pi}}(\boldsymbol{\theta})(\hat{\mathbf{R}} - \hat{\mathbf{Q}})\hat{\boldsymbol{\Pi}}^H(\boldsymbol{\theta}) + \hat{\mathbf{Q}} \quad (22)$$

[1] Note that the unstructured noise can be assumed orthogonal to subspace $\langle\mathbf{B}\rangle$ without loss of generality since the component of $\mathbf{w}(t)$ that would belong to $\langle\mathbf{B}\rangle$ can be included in the model of $\mathbf{v}(t)$

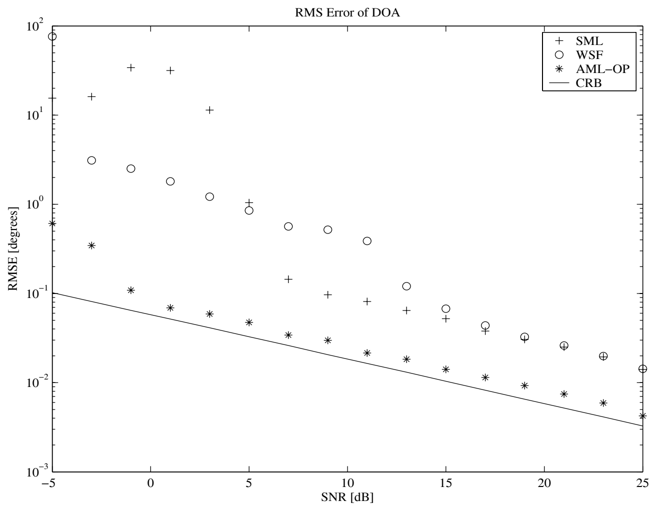

Fig. 1. Comparison of AML-OP, SML and WSF vs SNR. $T = 200, K = 10, \rho = 0.98$.

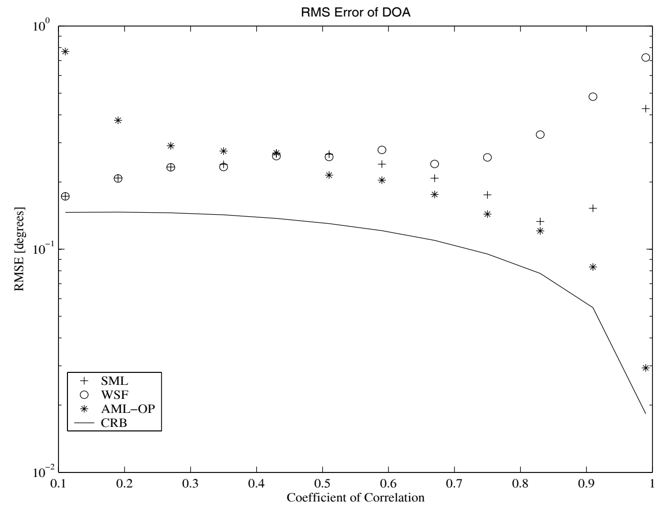

Fig. 2. Comparison of AML-OP, SML and WSF vs correlation coefficient. $T = 200, K = 8, SNR = 0dB$.

where $\hat{\mathbf{\Pi}}(\boldsymbol{\theta}) = \mathbf{A}(\mathbf{A}^H\hat{\mathbf{Q}}^{-1}\mathbf{A})^{-1}\mathbf{A}^H\hat{\mathbf{Q}}^{-1}$. Thus, the modified LL cost function becomes

$$\mathcal{L}(\boldsymbol{\theta}) = \ln\left\{\det\left[\tilde{\mathbf{R}}(\boldsymbol{\theta})\right]\right\} + \text{trace}\left\{\tilde{\mathbf{R}}^{-1}(\boldsymbol{\theta})\hat{\mathbf{R}}\right\} \tag{23}$$

Finally, estimation of the parameters reduces to solving the following optimization problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}}\{\mathcal{L}(\boldsymbol{\theta})\} \tag{24}$$

In order to achieve convergence, the algorithm requires favorable initial parameters. One way among others to initialize the algorithm is to use a standard estimator such as ROOT-MUSIC [7]. We use a quasi-Newton algorithm to solve the optimization problem in (24) using the scoring technique [8] with a mixed quadratic and cubic line search procedure.

4. SIMULATION RESULTS

We use the following simulation scenario: assume a radar working in passive mode (only one set of snapshots is provided). The array is simulated as a Uniform Linear Array (ULA). Two impinging signals corresponding to two targets situated respectively at $\theta_1 = -3^\circ$ and $\theta_2 = 1^\circ$ with different Doppler frequencies and T samples are collected at each sensor. We illustrate the global performance of the proposed approach in terms of the Root Mean Square Error (RMSE) versus the Signal to Noise Ratio (SNR), the spatial correlation coefficient ρ, the number of collected snapshots T and the number of sensors K. We compare the performance of the proposed AML-OP approach versus SML and WSF [9] techniques where the noise is assumed white Gaussian and the Cramér-Rao Bound (CRB). The parameter set is indicated in the figure captions. The results are averaged after 200 Monte Carlo runs. As the performance is similar for the two DOAs, only the first DOA results are shown.

Figure 1 illustrates the performance of AML-OP, WSF and SML in terms of RMSE. As expected, the AML-OP performs better than the other techniques, especially at low SNR.

Figure 2 shows the variation of the RMSE with the coefficient of correlation ρ. As the correlation coefficient increases, AML-OP outperforms the other methods as the correlation in the noise is taken into account. For low value of ρ, the other techniques suit better the uncorrelated model. The AML-OP clearly outperforms the two other approaches only at ρ greater than to 0.4.

Figure 3 illustrates the performance as we increase the number of antenna sensors. As the structured model is accounted for, AML-OP exhibits high performance.

Figure 4 shows the improvement of the performance as the number of collected data is increased while keeping the remaining parameters constant during simulation. Similar remarks as previously are noted.

5. CONCLUSION

In this paper, an AML estimator for DOA retrieving in the presence of structured and unstructured noise is proposed. The proposed approach uses an Oblique Projection to provide an estimate of the noise covariance matrix. Simulated data examples are provided to assess the performance of the AML-OP and to illustrate its relative superiority over two other techniques (SML, WSF) where the structure of noise is not taken into account.

6. APPENDIX
CRAMÉR-RAO BOUND

Considering the covariance matrix of the received data

$$\mathbf{R}(\boldsymbol{\eta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q}$$

then the (i, j)-th element of the Fisher Information Matrix (FIM) is

$$\begin{aligned}\boldsymbol{\mathcal{F}}_{i,j} &= \text{trace}\left\{\mathbf{R}^{-1}\left[\frac{\partial\mathbf{R}(\boldsymbol{\eta})}{\partial\boldsymbol{\eta}_i}\right]\mathbf{R}^{-1}[\frac{\partial\mathbf{R}(\boldsymbol{\eta})}{\partial\boldsymbol{\eta}_j}]\right\} && (25)\\ i,j &= 1, 2, \cdots, d^2 + d + 2K - 1 && (26)\end{aligned}$$

Let $\mathbf{G} = \mathbf{A}(\boldsymbol{\theta})'$, where $(.)'$ stands for differentiation with respect to the individual parameters. Using the results in [10], [11], the

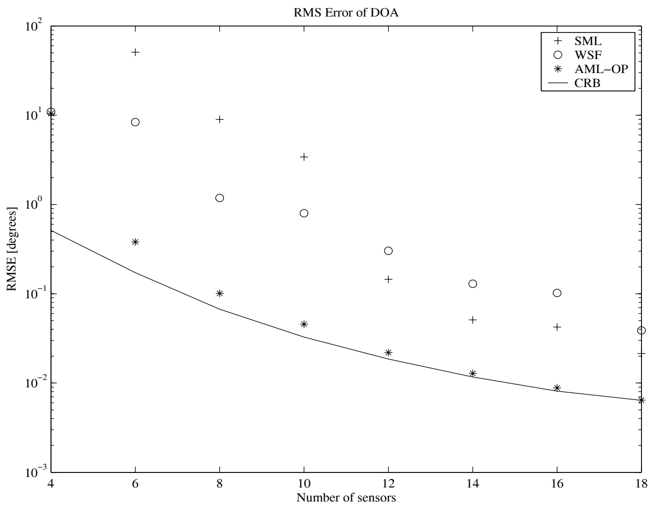

Fig. 3. Comparison of AML-OP, SML and WSF vs number of antennas. $T = 200, \rho = 0.9, SNR = 5dB$.

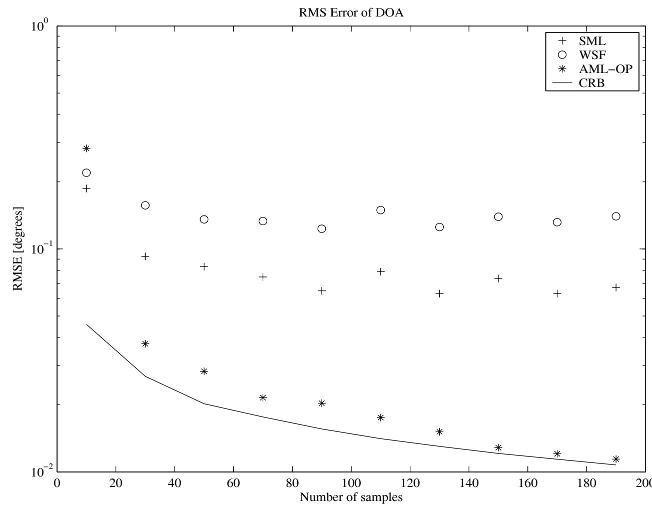

Fig. 4. Comparison of AML-OP, SML and WSF vs number of snapshots. $K = 10, \rho = 0.98, SNR = 5dB$.

expressions of the (i, j)-th element of the FIM for the parameters of interest $\boldsymbol{\theta}$ is given by

$$\begin{aligned} \boldsymbol{\mathcal{F}}_{i,j} &= \Big[2\Re \Big\{ \text{trace} \Big\{ (\tilde{\mathbf{G}}^H P_{\tilde{\mathbf{A}}}^{\perp} \tilde{\mathbf{G}})_{i,j} (\mathbf{S}\tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}}\mathbf{S})_{i,j} \Big\} \Big\} \\ &\quad - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{H}}^T \Big] \end{aligned} \tag{27}$$

where $\tilde{\mathbf{A}} = \mathbf{Q}^{-1/2}\mathbf{A}$, $\tilde{\mathbf{G}} = \mathbf{Q}^{-1/2}\mathbf{G}$, $\tilde{\mathbf{R}} = \mathbf{Q}^{-1/2}\mathbf{R}\mathbf{Q}^{-1/2}$, and the real matrices $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{I}}$ are defined as follows

$$\boldsymbol{\mathcal{H}} = 2\Re \Big\{ (\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{A}}\mathbf{S})^T \odot (\tilde{\mathbf{G}}^H P_{\tilde{\mathbf{A}}}^{\perp}) \Big\}$$

and

$$\boldsymbol{\mathcal{I}} = \Big\{ (\tilde{\mathbf{R}}^{-1})^c \odot \tilde{\mathbf{R}}^{-1} - (P_{\tilde{\mathbf{A}}}\tilde{\mathbf{R}}^{-1})^c \odot (P_{\tilde{\mathbf{A}}}\tilde{\mathbf{R}}^{-1}) \Big\}^{-1}$$

respectively, with $(.)^c$ denoting complex conjugate.

The closed-form expression of the CRB for the parameters

$$\begin{aligned} CRB_\theta &= \Big[2\Re \Big\{ \text{trace} \Big\{ (\tilde{\mathbf{G}}^H P_{\tilde{\mathbf{A}}}^{\perp} \tilde{\mathbf{G}}) \odot (\mathbf{S}\tilde{\mathbf{A}}^H \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{A}}\mathbf{S}) \Big\} \Big\} \\ &\quad - \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{H}}^T \Big]^{-1} \end{aligned} \tag{28}$$

7. REFERENCES

[1] R.O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *Proc. RADC Spectrum Estimation Workshop*, pp. 243-258, October 1987.

[2] P. Stoica and K.C. Sharman, "Maximum Likelihood Methods for Direction-Of-Arrival Estimation", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 7, pp. 1132-1143, July 1990.

[3] S. Haykin, *Advances in Spectrum Analysis and Array Processing*. Englwood Cliffs, NJ: Prentice-Hall, 1995, vol. 34 no. 2, pp. 207-222, November 1993.

[4] L. L. Scharf, *Statistical Signal Processing : Detection, Estimation and Time Series Analysis*, Addison Wesley, 1990.

[5] J. Böhme and D. Kraus,"On Least Squares Methods for Direction of Arrival Estimation in the Presence of Unknown Noise Fields,"*In Proc. ICASSP 88*, April 1988.

[6] R. T. Behrens and L. L. Scharf, "Signal Processing Applications of Oblique Projection Operators," *IEEE Transactions on Signal Processing*, vol. 42, no. 6 pp. 1413- 1424, June 1994.

[7] B. Rao, K. V. S. Hari, "Performance Analysis of Root-Music," *IEEE Transactions on Signal Processing*, vol. 37, no. 12, December 1989.

[8] Dennis J. and Schnabel R., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall. 1983.

[9] M. Viberg, B. Ottersten, and T. Kailath, "Detection and Estimation in Sensor Arrays using Weighted Subspace Fitting," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2439-2449, November 1991.

[10] P. Stoica and A. Nehorai, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1783-1795, Oct. 1990.

[11] M. Pesavento and A. B. Gershman, "Maximum-Likelihood Direction-Of-Arrival Estimation in the Presence of Unknown Nonuniform Noise," *IEEE Transactions on Signal Processing*, vol. 49, no. 7, pp. 1310-11324, July 2001.