EFFICIENT DOA TRACKING FOR TDMA-BASED SDMA MOBILE COMMUNICATIONS

Rolf Weber, Josef A. Nossek*

Institute for Network Theory and Signal Processing
Munich University of Technology
D-80290 Munich, Germany

Abstract - In this paper, we discuss subspace tracking algorithms for DOA estimation that take a burst-wise data flow - as it naturally occurs in systems using a TDMA component - into account. This leads to a matrix model rather than to the well known snapshot vector model for the sensor array outputs. Unfortunately, many existing algorithms take advantage of the vector model and do either not support the matrix model or are not efficient with it . Therefore, we derive three efficient algorithms based on the matrix model which are extensions of the known ISU and PAST methods. Furthermore, we show how changes in the number of present signals can be detected and how to incorporate this step. Simulation results are presented.

I. INTRODUCTION

Recently, space-division multiple access (SDMA) has received much attention as a method of reducing losses due to multipath and interference and thus increasing the capacity of a cellular mobile communication system. This is achieved by using spatial diversity introduced by an antenna array in addition to time- and/or frequency-division multiple access (TDMA/FDMA). For example, different users transmit in the same frequency band and in the same time slot and are only separated by the different directions of arrival (DOA's) of their signals received at the base station. The prerequisite, however, is exact knowledge of the DOA's of the mobile users which is a heavy computational burden and thus the need for efficient computational methods arises.

II. DATA MODEL

Let us first assume that information is transmitted by means of a continuous data flow. Assume further that there are r narrow-band signals having the same center-frequency f_0 and being characterized by their complex envelopes $s_i(t)$, $1 \leq i \leq r$, impinging on an antenna array consisting of n, $n \geq r$, identical sensors under the directions of arrival θ_i.

*Now with Signal Theory Group, Ruhr University Bochum, D-44780 Bochum, Germany (E-mail: wbr@sth.ruhr-uni-bochum.de)

Let $\mathbf{x}(k) \in \mathbb{C}^n$ be the data vector observed at the sensor array during the kth snapshot. The previous stated assumptions then lead to the model [5, 10]

$$\begin{aligned} \mathbf{x}(k) &= \sum_{i=1}^{r} \mathbf{a}(\theta_i)\, s_i(k) + \mathbf{n}(k) \\ &= \mathbf{A}(\boldsymbol{\theta})\, \mathbf{s}(k) + \mathbf{n}(k) \qquad (1) \end{aligned}$$

with $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_i) \cdots \mathbf{a}(\theta_r)] \in \mathbb{C}^{n \times r}$, the array steering matrix, depending on the vector $\boldsymbol{\theta}$ of the directions of arrival, and the complex valued noise vector $\mathbf{n}(k) \in \mathbb{C}^n$ which is assumed to be spatially white with equal variance σ^2 and uncorrelated with the signal vector $\mathbf{s}(k)$. This yields the following expression for the spatial correlation matrix

$$\mathbf{C} = \mathcal{E}\left\{\mathbf{x}(k)\,\mathbf{x}^H(k)\right\} = \mathbf{A}(\boldsymbol{\theta})\,\mathbf{C}_S\,\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2\,\mathbf{I} \qquad (2)$$

where $\mathcal{E}\{\circ\}$ denotes expectation, $(\circ)^H$ denotes the conjugate transpose and $\mathbf{I}$ is the identity matrix. Based on an eigendecomposition of (2) or a SVD (singular value decomposition) of a data matrix composed of l consecutive vectors (1) the lower dimensional signal and noise subspaces can be identified which may be used to calculate the DOA's via high resolution methods like MUSIC, Unitary-ESPRIT or Weighted Subspace Fitting [3, 5]. But finding the subspaces is a high numerical burden, especially if they change with time and therefore have to be computed recurringly which shows the necessity to track them efficiently.

Let us now assume a communication system like GSM that uses a TDMA component. Then there is no continuous data flow but the mobile user rather transmits a whole data burst consisting of m symbols during its time slot which is periodically recurring within a TDMA frame. This leads to the matrix model

$$\mathbf{X}(p) = \mathbf{A}(\boldsymbol{\theta})\,\mathbf{S}(p) + \mathbf{N}(p) \qquad (3)$$

with the data matrix $\mathbf{X} \in \mathbb{C}^{n \times m}$, the signal matrix $\mathbf{S} \in \mathbb{C}^{r \times m}$, the noise matrix $\mathbf{N} \in \mathbb{C}^{n \times m}$ and p, the time index of TDMA frames. Because of the different time scales a sequential use of known subspace tracking algorithms for

0-7803-5565-2/99/$10.00 © 1999 IEEE

snapshot vectors is not useful in the matrix case. On the other hand, as the time between consecutive bursts of one user is usually small, e.g. 4.616 ms in the GSM system, the whole scenario will not change considerably so that an update of the interesting subspace from burst to burst is sufficient. We therefore derive efficient algorithms for the subspace tracking problem taking the burst-wise data flow into account.

III. INVARIANT SUBSPACE UPDATE (ISU)

Let $\mathbf{C} \in \mathbb{C}^{n\times n}$ and let $\mathbf{X}_1$ be an orthonormal basis for an invariant subspace [1] of $\mathbf{C}$ which means that the space spanned by the columns of $(\mathbf{C}\,\mathbf{X}_1)$ is a subspace of the one spanned by the columns of $\mathbf{X}_1$. Choose the unitary matrix $[\mathbf{X}_1\,\mathbf{X}_2]$ to be a basis for the space spanned by the columns of $\mathbf{C}$.
Now consider $\mathbf{C}'$, a perturbed version of $\mathbf{C}$ and its orthonormal basis $[\mathbf{X}_1'\,\mathbf{X}_2']$, where $\mathbf{X}_1'$ is a basis for the invariant subspace that corresponds to $\mathbf{X}_1$. Then there exists a unitary matrix $\mathbf{U}$ such that $[\mathbf{X}_1'\,\mathbf{X}_2'] = [\mathbf{X}_1\,\mathbf{X}_2]\,\mathbf{U}$ and $\mathbf{U}$ can be factored in the form [7]

$$\mathbf{U} = \begin{bmatrix} \mathbf{I} & -\mathbf{P}^H \\ \mathbf{P} & \mathbf{I} \end{bmatrix} \begin{bmatrix} (\mathbf{I}+\mathbf{P}^H\mathbf{P})^{-1/2} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}+\mathbf{P}\mathbf{P}^H)^{-1/2} \end{bmatrix}$$

with $\mathbf{P}$ chosen appropriately. For $\mathbf{X}_1'$ to be a basis for an invariant subspace of $\mathbf{C}'$, $\mathbf{X}_2'^H(\mathbf{C}'\,\mathbf{X}_1') = \mathbf{0}$ must be satisfied. Using the partitioned matrix

$$[\mathbf{X}_1\,\mathbf{X}_2]^H\,\mathbf{C}'\,[\mathbf{X}_1\,\mathbf{X}_2] = \begin{bmatrix} \mathbf{C}_{11}' & \mathbf{C}_{12}' \\ \mathbf{C}_{21}' & \mathbf{C}_{22}' \end{bmatrix}$$

then leads to the algebraic Riccati equation

$$\mathbf{P}\,\mathbf{C}_{11}' - \mathbf{C}_{22}'\,\mathbf{P} = \mathbf{C}_{21}' - \mathbf{P}\,\mathbf{C}_{12}'\,\mathbf{P}. \qquad (4)$$

In [6] MacInnes and Vaccaro introduce an efficient iterative procedure to solve (4) for $\mathbf{P}$ and show how to use the above theory to track signal and noise subspaces based on the vector model (1). This approach can easily be modified to work with the matrix model (3) as is shown in the sequel.
Choose $\mathbf{C}$ to be the estimate

$$\mathbf{C}(p) = \sum_{t=1}^{p} \beta^{p-t}\,\mathbf{X}(t)\,\mathbf{X}^H(t)$$

of the correlation matrix at time index p, where β is the forgetting factor, and $\mathbf{X}_1 \in \mathbb{C}^{n\times r}$, $\mathbf{X}_2 \in \mathbb{C}^{n\times(n-r)}$ to be orthonormal eigenbases for the signal and noise subspaces respectively. If $\mathbf{C}'$ is $\mathbf{C}(p+1)$ we get the new bases of the subspace of interest at time $p+1$ by means of $\mathbf{U}$ which can be computed from $\mathbf{P}$. To ensure that the new bases contain orthogonal columns a QR-decomposition of $\mathbf{X}_1'(p+1)$ has to be done [6].

IV. TRACKING BY OPTIMIZATION

In [10] Yang demonstrated that the signal subspace can be found by minimizing a scalar cost function. Unfortunately, both the derivation of the theoretical results and the resulting numerically efficient tracking algorithms heavily depend on the snapshot vector model. We therefore introduce a modified cost function whose global minimum again yields the signal subspace.
Consider the cost function

$$\begin{aligned} J'(\mathbf{W}) &= \mathcal{E}\left\{\|\mathbf{X} - \mathbf{W}\mathbf{W}^H\mathbf{X}\|_F^2\right\} \\ &= \sum_{j=1}^{m} \mathcal{E}\left\{\|\mathbf{x}_j - \mathbf{W}\mathbf{W}^H\mathbf{x}_j\|_2^2\right\} \qquad (5) \end{aligned}$$

of $\mathbf{W} \in \mathbb{C}^{n\times r}$, where $\|\circ\|_F$ denotes the Frobenius norm of a matrix and $\mathbf{X} = [\mathbf{x}_1\,\mathbf{x}_2\,\cdots\,\mathbf{x}_m] \in \mathbb{C}^{n\times m}$ is the data matrix. Without loss of generality, we assume $\mathbf{W}$ to have full rank r. Assuming at least short time stationary signals $\mathbf{x}_j$ we get

$$J'(\mathbf{W}) = m\cdot\mathcal{E}\left\{\|\mathbf{x} - \mathbf{W}\mathbf{W}^H\mathbf{x}\|_2^2\right\} \qquad (6)$$

which is m times the original cost function introduced by Yang. Therefore, the two theorems concerned with the stationary points of the cost function which were proved in [10] are still valid. Their main result is that all stationary points of $J'(\mathbf{W})$ are saddle points except when $\mathbf{W} = \mathbf{U}_r\,\mathbf{Q}$, where $\mathbf{U}_r$ contains the r dominant eigenvectors of $\mathbf{C}$ and $\mathbf{Q}$ is an arbitrary unitary matrix. In that case, $\mathbf{W}$ is an orthonormal basis of the signal subspace and the minimization can be done without additional orthogonalization steps by iterative algorithms that will always converge to the global minimum of $J'(\mathbf{W})$.

Gradient-Based Algorithm

Since (5) describes an unconstrained cost function to be minimized, a steepest descent algorithm can be used to calculate $\mathbf{W}(p)$ recursively. The gradient of $J'(\mathbf{W})$ is [10]

$$\nabla J'(\mathbf{W}) = m\left[-2\mathbf{C} + \mathbf{C}\mathbf{W}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H\mathbf{C}\right]\mathbf{W}.$$

Choosing the instantaneous estimate

$$\hat{\mathbf{C}}(p) = \frac{1}{m}\,\mathbf{X}(p)\,\mathbf{X}^H(p)$$

as an appropriate estimate of the correlation matrix at time index p and observing that $\mathbf{W}(p)$ will converge to a matrix with orthonormal ($\mu \to 0$) or nearly orthonormal ($\mu = \text{const}$ but small) columns [10] and thus justifying the use of the approximation $\mathbf{W}^H(p-1)\,\mathbf{W}(p-1) \approx \mathbf{I}$ yields the well known LMS algorithm [4]

$$\mathbf{W}(p) = \mathbf{W}(p-1) + \mu\left[\mathbf{X}(p) - \mathbf{W}(p-1)\,\mathbf{Y}(p)\right]\mathbf{Y}^H(p)$$

with the step size μ and $\mathbf{Y}(p) = \mathbf{W}^H(p-1)\,\mathbf{X}(p)$. The algorithm is initialized by $\mathbf{W}(0) = \mathbf{0}$, see [4].

QR-RLS-Based Algorithm

We now replace the expectation in (5) through the exponentially weighted sum

$$J'(\mathbf{W}(p)) = \sum_{t=1}^{p} \beta^{p-t} \, ||\mathbf{X}(t) - \mathbf{W}(p)\,\mathbf{W}^H(p)\,\mathbf{X}(t)||_F^2$$

with the forgetting factor $0 < \beta \leq 1$. Using the approximation $\mathbf{Y}(t) = \mathbf{W}^H(t-1)\,\mathbf{X}(t)$ for $\mathbf{W}^H(p)\,\mathbf{X}(t)$ which can be instantaneously calculated at time p yields a cost function which is a good approximation of the original one [10] and whose derivative [8] must be zero at the global minimum (normal equations):

$$\sum_{t=1}^{p} \sum_{j=1}^{m} \beta^{p-t} \left[\mathbf{x}_j(t) - \mathbf{W}(p)\,\mathbf{y}_j(t)\right] \mathbf{y}_j^H(t) = \mathbf{0}\,.$$

This gives the matrix equation

$$\mathbf{W}(p)\,\mathbf{C}_{YY}(p) = \mathbf{C}_{XY}(p)$$

with

$$\begin{aligned} \mathbf{C}_{YY}(p) &= \sum_{t=1}^{p} \beta^{p-t}\,\mathbf{Y}(t)\,\mathbf{Y}^H(t) \\ \mathbf{C}_{XY}(p) &= \sum_{t=1}^{p} \beta^{p-t}\,\mathbf{X}(t)\,\mathbf{Y}^H(t)\,. \end{aligned}$$

Using the vector model as in [10] leads to a rank one update of $\mathbf{C}_{YY}$ from p to $p+1$ and therefore $\mathbf{W}(p)$ can be determined without a matrix inversion by using the inversion lemma [4] resulting in the highly efficient PAST algorithm. Unfortunately, this lemma cannot be gainfully used in the matrix model case.

Assuming that $\mathbf{C}_{YY}(p)$ is positive definite we can use the Cholesky factorization $\mathbf{C}_{YY}(p) = \mathbf{R}^H(p)\mathbf{R}(p)$ with the unique upper triangular matrix $\mathbf{R}(p) \in \mathbb{C}^{r\times r}$ to obtain

$$\mathbf{R}(p)\,\mathbf{W}^H(p) = \mathbf{\Gamma}(p) \qquad (7)$$

with $\mathbf{\Gamma}(p) = \mathbf{R}^{-H}(p)\mathbf{C}_{XY}^H(p)$. If $\mathbf{R}(p-1)$ and $\mathbf{\Gamma}(p-1)$ are known and at time p new data $\mathbf{X}(p)$ and $\mathbf{Y}(p)$ become available, the following pre-array can be formed:

$$\begin{bmatrix} \sqrt{\beta}\,\mathbf{R}(p-1) & \sqrt{\beta}\,\mathbf{\Gamma}(p-1) \\ \mathbf{Y}^H(p) & \mathbf{X}^H(p) \end{bmatrix}.$$

Multiplying this pre-array from the left by a unitary matrix $\mathbf{Q}(p) \in \mathbb{C}^{(r+m)\times(r+m)}$ to create a block zero in the post-array (which can be done by mr Givens rotations)

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

one can identify $\mathbf{A}_{11} = \mathbf{R}(p)$ and $\mathbf{A}_{12} = \mathbf{\Gamma}(p)$ [8]. $\mathbf{W}(p)$ can be efficiently and numerically stable computed by back-substitution in (7). The QR-RLS algorithm can be exactly initialized by $\mathbf{R}(0) = \sqrt{\delta}\,\mathbf{I}$ and $\mathbf{\Gamma}(0) = \mathbf{0}$, where δ is a non-negative constant, has good numerical properties and suits for implementation on systolic arrays [4, 9].

V. SIGNAL DETECTION

Usually r, the number of signals, is unknown and has to be estimated. But many of the well known algorithms like AIC and MDL fail to determine r whenever an exponentially weighted window is used to estimate the covariance matrix, because its rank does not change immediately. In [6] a new method of immediately detecting changes in the number of signals is proposed which is extended to a burst-wise data flow in the following.

Assume the vector of the most recent DOA's is $\hat{\boldsymbol{\theta}}$ and let $\hat{\mathbf{A}} = \mathbf{A}(\hat{\boldsymbol{\theta}})$ denote the array steering matrix computed using these DOA's. Denote with $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_m]$ the modified data matrix whose columns have a Euclidean norm of one and let $\mathbf{U}_S \in \mathbb{C}^{n\times r}$ be the most recently computed orthonormal basis for the signal subspace. Then, the orthogonal projections of the m columns of $\tilde{\mathbf{X}}$ onto the subspace spanned by the r columns of $\hat{\mathbf{A}}$ are given by

$$\mathbf{X}' := [\mathbf{x}'_1 \cdots \mathbf{x}'_m] = \mathbf{P}_{\hat{\mathbf{A}}}\,\tilde{\mathbf{X}} \in \mathbb{C}^{n\times m}\,,$$

with $\mathbf{P}_{\hat{\mathbf{A}}} = \hat{\mathbf{A}}\left(\hat{\mathbf{A}}^H\hat{\mathbf{A}}\right)^{-1}\hat{\mathbf{A}}^H$, the projector onto $\hat{\mathbf{A}}$.

Increasing Number of Signals

Now, if the number of signals in the new data burst $\mathbf{X}$ increases from r to $r+1$, the data burst contains a component that is not in the range of $\hat{\mathbf{A}}$ and therefore at least one of the plots of $1 - ||\mathbf{x}'_j||_2$, $1 \leq j \leq m$, versus the iteration number will exceed a threshold close to zero. In that case, let j' be

$$j' = \arg \max_{1\leq j\leq m} \{1 - ||\mathbf{x}'_j||_2\}\,.$$

Then we set $\mathbf{U}_S = \mathrm{orth}([\mathbf{U}_S\,\mathbf{x}_{j'}])$, where $\mathrm{orth}(\circ)$ means that the columns of $[\mathbf{U}_S\,\mathbf{x}_{j'}]$ have to be orthonormalized.

Decreasing Number of Signals

If the number of signals decreases from r to $r-1$, the new data burst $\mathbf{X} \in \mathbb{C}^{n\times m}$ will be in the range of only $r-1$ of the r estimated steering vectors. Assuming that the ith signal source disappeared, $\mathbf{X}$ will be in the range of

$$\hat{\mathbf{A}}_i := [\hat{\mathbf{a}}_1 \cdots \hat{\mathbf{a}}_{i-1}\,\hat{\mathbf{a}}_{i+1} \cdots \hat{\mathbf{a}}_r]\,.$$

Then, as $\tilde{\mathbf{X}}$ lies in the range of $\hat{\mathbf{A}}_i$, the projection of each vector $\tilde{\mathbf{x}}_j$, $1 \leq j \leq m$, on the columns of $\hat{\mathbf{A}}_i$ will be approximately of norm one:

$$||\mathbf{P}_{\hat{\mathbf{A}}_i}\tilde{\mathbf{x}}_j||_2 \approx 1\,, \qquad \forall j : 1 \leq j \leq m\,.$$

Therefore, a decrease in the number of signals is detected when for some $1 \leq j \leq m$ the norm of the left-hand side of the above equation exceeds a threshold close to one for all $0 \leq j \leq m$, and we set $\mathbf{U}_S = \mathrm{orth}(\hat{\mathbf{A}}_i)$.
In order to ensure that the correlation matrix $\mathbf{C}$ contains only components in the range of the remaining steering vectors, the modified correlation function

$$\mathbf{C} = \hat{\mathbf{A}}_i \hat{\mathbf{A}}_i^H$$

is suggested [6].

VI. SIMULATIONS

Consider a uniform linear array consisting of 9 equal sensors spaced half of the wavelength apart. In this case the array steering vector is known to be

$$\mathbf{a}(\theta_i) = \left[1, e^{j(2\pi f_i)}, \ldots, e^{j8(2\pi f_i)}\right]^T, \; f_i = \frac{1}{2}\sin(\theta_i).$$

A TLS-ESPRIT algorithm [5] is applied to calculate the directions of arrival. We consider a GSM-like system ($m = 156$, 8 time slots of $576.9\,\mu s$) except that we assume the transmitted symbols to be Gaussian random variables uncorrelated from each other. In all cases β is set to 0.97.
In a first experiment, there may be two signals positioned in the same frequency range and time slot (the number of signals is known). The first one has a fixed spatial frequency of $f_1 = -0.2$ and its signal to noise ratio (SNR) is 0 dB. The spatial frequency of the second one increases linearly from 0.2 to 0.3 within 6500 bursts ($\approx 30\,s$) at an SNR of 5 dB. The result for the QR-RLS algorithm is shown in Figure 1(a).

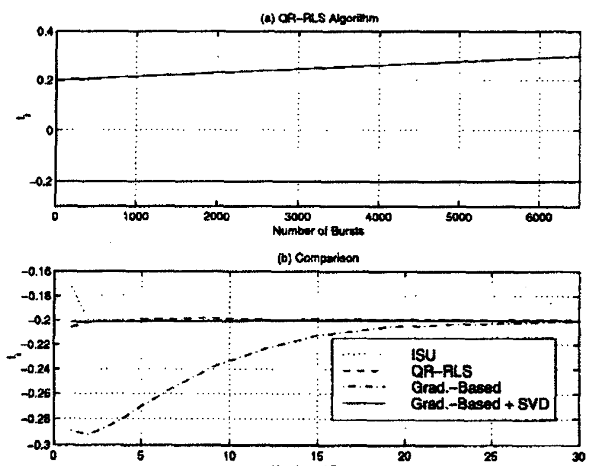

Figure 1: *Tracking behaviour of the different algorithms*

The other algorithms show a similar steady-state behaviour but differ significantly in the transient phase. Especially the gradient-based algorithm ($\mu = 8 \cdot 10^{-5}$) converges slowly, see Figure 1(b), where the first 30 bursts ($\approx 0.14\,s$) of signal 2 for the various algorithms are depicted.
The reason for this is that unlike the QR-RLS algorithm the gradient-based algorithm cannot be initialized exactly why $\mathbf{W}(0) = \mathbf{0}$ is used. Therefore, we suggest to perform one SVD of the first burst to find an estimate of a basis for the signal subspace. This basis can be used to initialize the gradient-based algorithm leading to a much better performance, see Figure 1(b).
In a second experiment, the detection of the current number of signals is tested. We simulate the transmission of 500 bursts ($\approx 2.3\,s$) during which a first signal is always present at $f_1 = -0.2$. A second signal is present from burst 100 to 149, 200 to 399 and 420 to 500 at $f_2 = 0.1$. The third signal is present from burst 100 to 399 and its spatial frequency f_3 increases linearly from 0.2 to 0.3. All signals have an SNR of 20 dB. We use the ISU algorithm and the thresholds for detecting increase or decrease are chosen to be 0.15 and 0.98, respectively. The results are shown in Figure 2 and

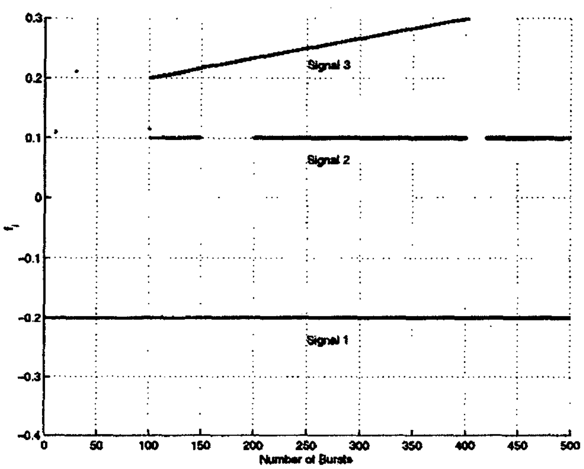

Figure 2: *Detecting changes in the number of signals*

demonstrate the good performance of the algorithm even if the time between changes is relatively short and several changes occur at once. In the later case the changes in the number of signals are detected in consecutive time steps.

VII. CONCLUDING REMARKS

We gave reason that a matrix model rather than a vector model for the senor array outputs should be used when there is a TDMA component inherent in the system. Then, we extended two known subspace tracking algorithms for the vector model to work with a burst-wise data flow leading to three new algorithms. Furthermore, we showed how changes in the number of present signals can be detected again taking the burst-wise data flow into account. Simulation results demonstrated the good performance of all proposed methods.

VIII. REFERENCES

[1] Gohberg, I., Lancaster, P., Rodman, L.: *Invariant Subspaces of Matrices with Applications*, John Wiley & Sons, New York, 1986

[2] Graham, A.: *Kronecker Products and Matrix Calculus: With Applications*, Ellis Horwood Ltd., Chichester, 1981

[3] Haardt, M., Nossek, J.A.: Unitary ESPRIT: How to Obtain Increased Estimation Accuracy with a Reduced Computational Burden, *IEEE Transactions on Signal Processing*, Vol. 43, pp. 1232-1242, May 1995

[4] Haykin, S.: *Adaptive Filter Theory*, 3rd ed., Prentice Hall, Upper Saddle River, 1996

[5] Krim, H., Viberg, M.: Two Decades of Signal Processing: The Parametric Approach, *IEEE Signal Processing Magazine*, Vol. 13, No. 4, pp. 67-94, July 1996

[6] MacInnes, C.S., Vaccaro, R.J.: Tracking directions-of-arrival with invariant subspace updating, *Signal Processing*, Vol. 50, pp.137-150, 1996

[7] Stewart, G.W.: Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problems, *SIAM Review*, Vol. 15, No. 4, pp. 727-764, October 1973

[8] Weber, R.: *Subspace Tracking for Mobile Communications*, Technical Report, Munich University of Technology, Munich, 1997

[9] Yang, B., Böhme, J.F.: Rotation-Based RLS Algorithms: Unified Derivations, Numerical Properties and Parallel Implementations, *IEEE Transactions on Signal Processing*, Vol. 40, No. 5, pp. 1151-1166, May 1992

[10] Yang, B.: Projection Approximation Subspace Tracking, *IEEE Transactions On Signal Processing*, Vol. 43, No. 1, pp. 95-107, January 1995