# Knowledge-Aided STAP Algorithm Using Affine Combination of Inverse Covariance Matrices for Heterogenous Clutter

Rui Fa, Rodrigo C. de Lamare
Communications Research Group, Department of Electronics,
University of York, YO10 5DD, United Kingdom.
Email: {rf533, rcdl500}@ohm.york.ac.uk

Vítor H. Nascimento
Electronic Systems Engineering Department,
Escola Politecnica, University of Sao Paulo.
Email: vitor@lps.usp.br

***Abstract*—By incorporating *a priori* knowledge into radar signal processing architectures, knowledge-aided space-time adaptive processing (KA-STAP) algorithms can offer the potential to substantially enhance detection performance and to combat heterogeneous clutter effects. In this paper, we develop a KA-STAP algorithm to estimate directly the interference covariance matrix inverse rather than the covariance matrix itself, by using a linear combination of inverse covariance matrices (LCICM), which leads to an equivalent expression of the combination of two filters. The computational load is greatly reduced due to the avoidance of the matrix inversion operation. The performance of the LCICM scheme can be further improved by applying a modification. Moreover, adaptive algorithms for the mixing parameters are developed using affine combinations (AC). Numerical examples show the potential of our proposed algorithms for substantial performance improvement.**

***Index Terms*—Space-time adaptive processing, Knowledge-aided, Airborne radar applications, Affine Combination.**

## I. Introduction

Following the landmark publication by Brennan and Reed [1]- [7], space-time adaptive processing (STAP) techniques have been well developed since 1973. A joint-domain optimization of the spatial and temporal degrees-of-freedom (DOFs) can potentially offer a significant increase in output signal-to-interference-plus-noise-ratio (SINR) for airborne radar applications. To estimate the covariance matrix $\mathbf{R}$ in the optimal detector [2], the STAP must employ secondary data, generally taken from range cells adjacent to the cell under test (CUT). Prior work ( [3] and the references therein) have focused on algorithms with the underlying assumption that the training samples are independent and identically distributed (i.i.d) with the same covariance matrix as the primary data (so called homogeneous training). However, it is widely understood that the clutter environments are often heterogeneous (or non-i.i.d) [4], [5], for example, clutter reflectivity varies spatially and target-like signals frequently reside within the training data. Thus, merely using the sample covariance matrix estimate $\hat{\mathbf{R}}$ results in significant output SINR performance degradation.

To combat the deleterious effects of the heterogeneity in the secondary data, knowledge-aided (KA) STAP techniques, which make use of an *a priori* knowledge of the clutter covariance matrix, have currently gained significant attention [6], [8]–[11]. In KA-STAP, two questions have to be answered: the first is how to derive the prior covariance matrix from the terrain knowledge of the clutter and how to estimate the real interference covariance matrix with the prior knowledge [6], [8], [9]; the second is how to apply the covariance matrix estimate in the filtering [10], [11]. A number of techniques have been shown to result in more robust detection performance when the limited sample support is used in highly non-stationary clutter environments. However, most of the KA-STAP techniques studied previously inevitably require matrix inversion which has complexity of $\mathcal{O}(M^3)$, where $M$ is the dimension of the matrix.

In this paper, we propose a KA-STAP algorithm which estimates directly the inverse interference covariance matrix instead of the covariance matrix itself, by using a linear combination of inverse covariance matrices (LCICM), say the prior inverse covariance matrix $\mathbf{R}_0^{-1}$ and the sample inverse covariance matrix estimate $\hat{\mathbf{R}}^{-1}$. The linear combination of $\mathbf{R}_0^{-1}$ and $\hat{\mathbf{R}}^{-1}$, more precisely, can be expressed as $\mathbf{R}^{-1} = \alpha\mathbf{R}_0^{-1} + \beta\hat{\mathbf{R}}^{-1}$. Because the computation of $\hat{\mathbf{R}}^{-1}$ can be simplified by using the matrix inversion lemma and may be obtained recursively, our proposed algorithm has considerably lower complexity compared with the scheme proposed in [9]. Futhermore, [9] needs to substitute $\mathbf{R}$ by $\hat{\mathbf{R}}$ under the assumption of homogeneous training when estimating the mixing parameter, while our algorithm does not require such assumption and can be represented as a combination of filters [12]–[16]. The proposed KA-STAP-LCICM can be further modified with a combination of two recursive least squares (RLS) filters with prior knowledge initialization and normal initialization, respectively, which we refer as the modified LCICM (MLCICM) scheme. The mixing parameter in the algorithm can be determined as an affine combination (AC), making $\beta = 1 - \alpha$, where $\alpha \in \mathbb{R}$ [12], [13]. The simulation results show significant performance improvement brought by our LCICM-AC and MLCICM-AC algorithms.

The rest of the paper is organized as follows. In Sec. II, we briefly describe the problem at hand. Sec. III focuses on the key principle of the proposed KA-STAP algorithms including LCICM-AC and MLCICM-AC algorithms and describes the methods to estimate the mixing parameter. Sec. IV analyses the complexity of the proposed algorithms in terms of the number of additions and multiplications. Sec. V presents numerical examples illustrating the performance improvement of the proposed algorithm. Finally, conclusions are drawn in Sec. VI.

[0]This work is funded by the Ministry of Defence (MoD), UK. Project MoD, Contract No. RT/COM/S/021. The work of Vítor H. Nascimento was partly funded by FAPESP (grant 2008/04828-5) and CNPq (grant 303361/2004-2).

## II. Problem Statement

Consider a pulsed Doppler radar residing on an airborne platform that consists of an N-element uniformly spaced linear array antenna. Radar returns are collected in a coherent processing interval (CPI), which is referred to as the 3-D radar data-cube shown in Fig. 1(a), where $K$ denotes the number of samples collected to cover the range interval. A slice of the CPI data-cube with $J \times N$ dimensionality consists of $N \times 1$ spatial snapshots for $J$ pulses at the range of interest. It is convenient to stack the matrix column-wise to form the $M \times 1, M = JN$ vector $\mathbf{r}(i)$, termed a space-time snapshot [1].

Detection of targets in clutter involves determining the correct hypothesis in the CUT,

$$\begin{aligned} \mathbf{H}_0 &: \mathbf{r} = \boldsymbol{\nu} \\ \mathbf{H}_1 &: \mathbf{r} = \alpha \mathbf{s}_t + \boldsymbol{\nu}, \end{aligned} \tag{1}$$

where $\alpha$ is a complex gain and $\mathbf{s}_t$ is the target space-time steering vector, which is the $M \times 1$ normalized space-time steering vector in the space-time look-direction. Vector $\boldsymbol{\nu}$ encompasses any undesired interference or noise component of the data including clutter $\mathbf{c}$, jamming $\mathbf{j}$ and thermal noise $\mathbf{n}$.

The general STAP schematic diagram is shown in Fig. 1(b). An optimal STAP, in the maximum SINR sense, is given by $\boldsymbol{\omega}_{opt} = \gamma \mathbf{R}^{-1}\mathbf{s}_t$, where $\gamma$ is an arbitrary scalar and $\mathbf{R}$ is the interference-plus-noise covariance matrix. Normally, since $\mathbf{R}$ is unknown, secondary data $\{\mathbf{r}(k)\}_{n=1}^{K}$, which is from $K$ range cells adjacent to the CUT, is employed to estimate the covariance matrix by means of the well known formula

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{n=1}^{K} \mathbf{r}(n)\mathbf{r}^H(n). \tag{2}$$

Such estimate can be sufficiently accurate when $K$ is at least twice as great as $M$ and the training samples are assumed i.i.d. However, it has been widely recognized that the clutter environments are often heterogeneous and the impact of the heterogeneity on the STAP performance loss has been investigated in [4]. Thus, KA signal processing is becoming an important technique to combat the heterogeneity [8]. In [6], [9], the covariance matrix is estimated by combining an initial guess of the covariance matrix $\mathbf{R}_0$, derived from the digital terrain database or the data probed by the radar in previous scans, and the sample average covariance matrix estimate in the present scan $\hat{\mathbf{R}}$, so that

$$\mathbf{R} = \alpha \mathbf{R}_0 + \beta \hat{\mathbf{R}}. \tag{3}$$

With the estimated covariance matrix, many STAP algorithms including data pre-whitening [6], [8] and knowledge-aided constraints [10] improve clutter mitigation performance. However, most of these KA-STAP algorithms require a matrix inversion operation with a complexity of $\mathcal{O}(M^3)$, which motivates us to develop a novel KA-STAP with low complexity.

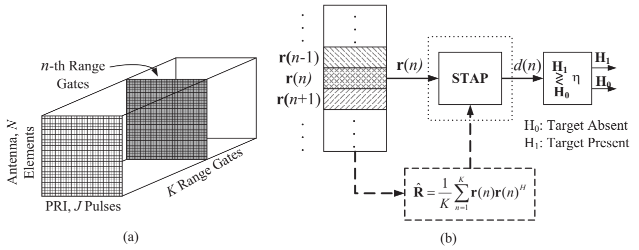


Fig. 1. (a) The Radar CPI datacube. (b) The STAP schematic.

## III. Proposed KA-STAP Algorithms

In this section, we describe the principle of the proposed KA-STAP algorithm using a linear combination of inverse covariance matrices (LCICM) and its modification (MLCICM). Note that in the context, the RLS algorithms under consideration are based on linearly constrained minimum variance (LCMV) criterion.

### A. LCICM Algorithms

The principle of our proposed algorithm is detailed in this subsection. We consider the inverse covariance matrix estimate by using an LCICM scheme instead of combining the covariance matrices themselves, more precisely,

$$\tilde{\mathbf{R}}^{-1}(n) = \alpha \mathbf{R}_0^{-1} + \beta \hat{\mathbf{R}}^{-1}(n). \tag{4}$$

Note that the covariance matrices with time index denote the time-dependent covariance matrix estimates, while those without time index denote the actual covariance matrix or fixed covariance matrix estimates.

The computation of the inverse sample average covariance matrix can be simplified by the matrix inversion lemma [20] and obtained recursively by

$$\hat{\mathbf{R}}^{-1}(n+1) = \lambda^{-1}\hat{\mathbf{R}}^{-1}(n) - \frac{\lambda^{-2}\hat{\mathbf{R}}^{-1}(n)\mathbf{r}(n)\mathbf{r}^H(n)\hat{\mathbf{R}}^{-1}(n)}{1 + \lambda^{-1}\mathbf{r}^H(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{r}(n)}, \tag{5}$$

where $\lambda$ is a forgetting factor and $\hat{\mathbf{R}}^{-1}(0)$ is initialized with $\delta \mathbf{I}_M$, $\delta$ is a small positive constant and $\mathbf{I}_M$ is an $M \times M$ identity matrix. Thus, the inverse covariance matrix estimate can be recursively calculated, which brings a significant reduction in the computational complexity. The remaining work is to effectively estimate the mixing parameter $\eta$.

If we multiply the target space-time steering vector $\mathbf{s}_t$ on both sides of (4), the formula will lead to a combination of two filters given by

$$\tilde{\mathbf{R}}^{-1}(n)\mathbf{s}_t = \alpha \mathbf{R}_0^{-1}\mathbf{s}_t + \beta \hat{\mathbf{R}}^{-1}(n)\mathbf{s}_t, \tag{6}$$

where for the time being we restrict $\alpha, \beta$ to be positive as in [9].

Recall that multiplication by a nonzero constant does not affect the performance of the algorithm. Let us define then $\boldsymbol{\omega}(n) = \gamma \tilde{\mathbf{R}}^{-1}(n)\mathbf{s}_t$, $\boldsymbol{\omega}_0 = \gamma \mathbf{R}_0^{-1}\mathbf{s}_t$ and $\boldsymbol{\omega}_1(n) = \gamma \hat{\mathbf{R}}^{-1}(n)\mathbf{s}_t$, with $\gamma = (\alpha + \beta)^{-1}$. With this choice, we obtain from (6)

$$\boldsymbol{\omega}(n) = \alpha' \boldsymbol{\omega}_0 + \beta' \boldsymbol{\omega}_1(n), \tag{7}$$

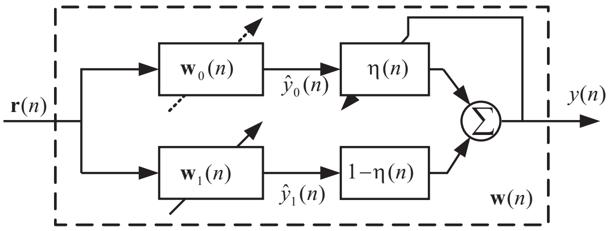


Fig. 2. Diagram of the proposed KA-STAP algorithms. The upper-branch filter $\boldsymbol{\omega}_0(n)$ is fixed in the LCICM scheme and adaptive in the MLCICM scheme. The lower-branch filter, $\boldsymbol{\omega}_1(n)$, is adaptive in both schemes.

where $\alpha' = \frac{\alpha}{\alpha+\beta}$ and $\beta' = \frac{\beta}{\alpha+\beta}$. Noting that $\alpha' + \beta' = 1$, we can reduce the number of parameters to estimate by defining $\eta = \alpha'$, so that

$$\boldsymbol{\omega}(n) = \eta\boldsymbol{\omega}_0 + (1-\eta)\boldsymbol{\omega}_1(n). \tag{8}$$

Let us consider the AC case, where $\eta$ can be any real number [12], [13]. A diagram of the proposed KA-STAP algorithm is shown in Fig. 2, where for LCICM the filter $\boldsymbol{\omega}_0$ is fixed. Thus, the mixing parameter $\eta$ can be optimized in the sense of minimizing the output power. The optimum mixing parameter which minimizes the cost function $\mathcal{L}(\eta)$ is given by

$$\eta_o = \arg\min_{\eta} \mathcal{L}(\eta) = \arg\min_{\eta} \mathbf{E}\left\{|\boldsymbol{\omega}^H(n)\mathbf{r}(n)|^2\right\}. \tag{9}$$

By equating to zero the gradient of the cost function $\mathcal{L}(\eta)$ with respect to $\eta$, we get

$$\eta_o = \frac{\Re\left\{\mathbf{s}_t^H \,\mathrm{E}\left[(\hat{\mathbf{R}}^{-1}(n) - \mathbf{R}_0^{-1})\mathbf{r}(n)\mathbf{r}^H(n)\hat{\mathbf{R}}^{-1}(n)\right]\mathbf{s}_t\right\}}{\mathbf{s}_t^H \,\mathrm{E}\left[(\mathbf{R}_0^{-1} - \hat{\mathbf{R}}^{-1}(n))\mathbf{r}(n)\mathbf{r}^H(n)(\mathbf{R}_0^{-1} - \hat{\mathbf{R}}^{-1}(n))\right]\mathbf{s}_t}. \tag{10}$$

where $\Re\{\cdot\}$ denotes the real part of a complex value. To simplify this expression, we need the following assumption.

*A-1:* If $\lambda \approx 1$, in steady-state it holds that $\hat{\mathbf{R}}(n)^{-1} \approx (\mathrm{E}\{\hat{\mathbf{R}}(n)\})^{-1}$.

This assumption is widely used in adaptive filtering [18], [19], and holds because for large $\lambda$ the variance of $\hat{\mathbf{R}}(n)$ is small in steady-state. Under A-1, we can write

$$\lim_{n\to\infty} \eta_o \approx \frac{\Re\left\{\mathbf{s}_t^H(\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1} - \mathbf{R}_0^{-1})\mathbf{R}\,\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1}\mathbf{s}_t\right\}}{\mathbf{s}_t^H(\mathbf{R}_0^{-1} - \mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1})\mathbf{R}(\mathbf{R}_0^{-1} - \mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1})\mathbf{s}_t}. \tag{11}$$

However, because $\mathbf{R}$ in (11) is unknown, we need an adaptive algorithm to estimate $\eta$ in real time, which is presented in the following section.

## B. MLCICM Algorithm

Motivated by the idea of combining two adaptive filters, we can modify the LCICM scheme for further improvement. The only difference between the LCICM and the MLCICM schemes is that $\boldsymbol{\omega}_0(n)$ is fixed over time in the LCICM case while it is adaptive using the RLS algorithm in the MLCICM scheme, but with a different initial condition. As shown in Fig. 2, the two component filters are both adaptive in the MLCICM scheme. In this case, we initialize the first component filter using the RLS algorithm with the inverse prior covariance matrix $\mathbf{R}_0^{-1}$ and the second component filter using the RLS algorithm with a scaled identity matrix.

The combination of two adaptive filters $\boldsymbol{\omega}_0(n)$ and $\boldsymbol{\omega}_1(n)$ is given mathematically as

$$\boldsymbol{\omega}(n) = \eta(n)\boldsymbol{\omega}_0(n) + (1-\eta(n))\boldsymbol{\omega}_1(n), \tag{12}$$

where $\boldsymbol{\omega}_1(n)$ is obtained exactly as in (5) in the LCICM case and $\boldsymbol{\omega}_0(n)$ can be calculated by using

$$\begin{aligned}\boldsymbol{\omega}_0(n) &= \hat{\mathbf{R}}_0^{-1}(n)\mathbf{s}_t,\\ \hat{\mathbf{R}}_0^{-1}(n+1) &= \lambda_0^{-1}\hat{\mathbf{R}}_0^{-1}(n) - \frac{\lambda_0^{-2}\hat{\mathbf{R}}_0^{-1}(n)\mathbf{r}(n)\mathbf{r}^H(n)\hat{\mathbf{R}}_0^{-1}(n)}{1+\lambda_0^{-1}\mathbf{r}^H(n)\hat{\mathbf{R}}_0^{-1}(n)\mathbf{r}(n)},\end{aligned} \tag{13}$$

where $\lambda_0$ is the forgetting factor for the first component filter using the RLS algorithm, $\hat{\mathbf{R}}_0^{-1}(0)$ is initialized with $\mathbf{R}_0^{-1}$. Similar to the LCICM scheme, we also can get the optimum mixing parameter by equating to zero the gradient of the cost function $\mathcal{L}(\eta)$ with respect to $\eta$. The solution is given by

$$\eta_o = \frac{\Re\left\{\mathbf{s}_t^H\left[\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1} - \mathrm{E}\{\hat{\mathbf{R}}_0(n)\}^{-1}\right]\mathbf{R}\,\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1}\mathbf{s}_t\right\}}{\mathbf{s}_t^H\left[\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1} - \mathrm{E}\{\hat{\mathbf{R}}_0(n)\}^{-1}\right]\mathbf{R}\left[\mathrm{E}\{\hat{\mathbf{R}}(n)\}^{-1} - \mathrm{E}\{\hat{\mathbf{R}}_0(n)\}^{-1}\right]\mathbf{s}_t}. \tag{14}$$

Note that if we try to evaluate the limit of $\eta_o$ as $n \to \infty$ using assumption A-1, we would get a $0/0$ indetermination, since in steady-state both filters will converge to the same steady-state value (the initial conditions will be forgotten), and the any value of $\eta$ would give the same performance. This is not a problem for our algorithm — in practice, $\eta$ will simply stop adapting when the filters reach steady-state. Compared with the LCICM scheme, which can guarantee that the initial performance of the scheme is not worse than that of the fixed filter $\mathbf{R}_0^{-1}\mathbf{s}_t$, the benefit of the modified scheme is that the first component filter is adapted starting with $\mathbf{R}_0^{-1}$ so that the initial performance is further improved. However, the conventional RLS algorithm converges to the steady-state much faster than the RLS algorithm with special initialization. The overall performance of the scheme will combine the benefits of both filters and outperform both of them.

## C. Mixing Parameter Estimation

In this subsection, an adaptive algorithm is developed to estimate the mixing parameter $\eta(n)$. Here, we borrow the ideas from [13] which deal with the adaptation of $\eta(n)$ for an AC of two adaptive filters. It should be remarked that our adaptation of $\eta$ is derived based on the minimum variance (MV) criterion for complex systems, which is an extension of previous results that dealt only with real variables.

For the adaptation of the mixing parameter $\eta(n)$, we use a gradient descent method to minimize the output power of the overall filter, say $p(n) = |y(n)|^2$, where $y(n)$ is the output of the overall filter which is a function of $\eta(n)$, given by

$$y(n) = \eta(n)y_0(n) + [1-\eta(n)]\hat{y}_1(n), \tag{15}$$

where $y_0(n) = \boldsymbol{\omega}_0^H\mathbf{r}(n)$ and $\hat{y}_1(n) = \hat{\boldsymbol{\omega}}_1^H(n)\mathbf{r}(n)$ are the outputs of the two component filters at time $n$. Here, we develop the adaptive algorithm for AC of filters where the mixing parameter $\eta$ is any real number. We can directly update $\eta(n)$ via the equation

$$\eta(n+1) = \eta(n) - \frac{\mu_\eta}{2[\sigma_\eta + q(n)]}\frac{\partial p(n)}{\partial \eta(n)}, \quad (16)$$

where $\mu_\eta$ is a step size, $\sigma_\eta$ is a small positive constant and $q(n)$ can be expressed by

$$q(n+1) = (1-\lambda_q)|y_0(n) - \hat{y}_1(n)|^2 + \lambda_q q(n), \quad (17)$$

where $\lambda_q$ is a forgetting factor. It was shown that better behavior is obtained by the normalized adaptation of the mixing parameter and the selection of $\lambda_q$ is simple [13]. Because $\eta$ should be a real number, we have to modify the recursion of [13]. Expanding the cost function, we obtain

$$\begin{aligned} |y(n)|^2 =& \Big|\eta(n)[y_0(n) - \hat{y}_1(n)] + \hat{y}_1(n)\Big|^2 \\ =& \eta(n)^2|y_0(n) - \hat{y}_1(n)|^2 \\ &+ 2\eta(n)\Re\Big\{[y_0(n) - \hat{y}_1(n)]\hat{y}_1(n)^*\Big\} + |\hat{y}_1(n)|^2, \end{aligned} \quad (18)$$

where $(\cdot)^*$ denotes the complex conjugate. Differentiating (18) with respect to $\eta$, we obtain

$$\begin{aligned} \frac{\partial p(n)}{\partial \eta(n)} =& \frac{\partial |y(n)|^2}{\partial \eta(n)} \\ =& 2\eta(n)|y_0(n) - \hat{y}_1(n)|^2 \\ &+ 2\Re\Big\{[y_0(n) - \hat{y}_1(n)]\hat{y}_1(n)^*\Big\}. \end{aligned} \quad (19)$$

Thus, the adaptation of $\eta(n)$ in (16) can be rewritten as

$$\begin{aligned} \eta(n+1) =& \eta(n) - \frac{\mu_\eta}{\sigma_\eta + q(n)}\Big\{\eta(n)|y_0(n) - \hat{y}_1(n)|^2 \\ &+ \Re\big\{[y_0(n) - \hat{y}_1(n)]\hat{y}_1(n)^*\big\}\Big\}. \end{aligned} \quad (20)$$

To obtain an analytical expression for the optimum mixing parameter $\eta_o(n)$ at steady-state, we apply the usual approximation of slow learning, i.e., we assume that $\eta(n)$ varies much slower than $\hat{y}_0(n) - \hat{y}_1(n)$, so that these variables may be assumed independent of each other [13], [17]. Using this approximation, we obtain

$$\lim_{n\to\infty} \mathrm{E}\{\eta(n)\} \approx -\frac{\mathrm{E}\{\Re\{(\hat{y}_0(n) - \hat{y}_1(n))\hat{y}_1^*(n)\}\}}{\mathrm{E}\{|\hat{y}_0(n) - \hat{y}_1(n)|^2\}}. \quad (21)$$

We expand the right hand side (RHS) of (21) in terms of the LCICM scheme as follows

$$\begin{aligned} &-\frac{\mathrm{E}\{\Re\{[\boldsymbol{\omega}_0^H\mathbf{r}(n) - \boldsymbol{\omega}_1^H(n)\mathbf{r}(n)]\mathbf{r}^H(n)\boldsymbol{\omega}_1(n)\}\}}{\mathrm{E}\{|\boldsymbol{\omega}_0^H\mathbf{r}(n) - \boldsymbol{\omega}_1^H(n)\mathbf{r}(n)|^2\}} \\ =& -\frac{\mathrm{E}\left\{\Re\{[\mathbf{s}^H\mathbf{R}_0^{-1} - \mathbf{s}^H\hat{\mathbf{R}}^{-1}(n)]\mathbf{r}(n)\mathbf{r}^H(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{s}^H\}\right\}}{\mathrm{E}\left\{[\mathbf{s}^H\mathbf{R}_0^{-1} - \mathbf{s}^H\hat{\mathbf{R}}^{-1}(n)]\mathbf{r}(n)\mathbf{r}^H(n)[\mathbf{R}_0^{-1}\mathbf{s} - \hat{\mathbf{R}}^{-1}(n)\mathbf{s}]\right\}} \\ =& \frac{\Re\left\{\mathbf{s}_t^H\,\mathrm{E}\left[(\hat{\mathbf{R}}^{-1}(n) - \mathbf{R}_0^{-1})\mathbf{r}(n)\mathbf{r}^H(n)\hat{\mathbf{R}}^{-1}(n)\right]\mathbf{s}_t\right\}}{\mathbf{s}_t^H\,\mathrm{E}\left[(\mathbf{R}_0^{-1} - \hat{\mathbf{R}}^{-1}(n))\mathbf{r}(n)\mathbf{r}^H(n)(\mathbf{R}_0^{-1} - \hat{\mathbf{R}}^{-1}(n))\right]\mathbf{s}_t} \\ =& \eta_o, \end{aligned} \quad (22)$$

which is identical to the optimum mixing parameter that we derived in (10). Similarly, we can expand the RHS of (21) in terms of the MLCICM scheme and obtain again (22), where $\mathbf{R}_0$ is replaced by $\hat{\mathbf{R}}_0(n)$.

TABLE I
COMPUTATIONAL COMPLEXITY OF ALGORITHMS.

| Algorithm | Number of operations per snapshot | |
|---|---|---|
| | Additions | Multiplications |
| Conventional RLS | $6M^2 - 8M + 3$ | $6M^2 + 2M + 2$ |
| Stoica et al.'s scheme | $M^3 + 3M$ | $M^3 + 4M^2 + 3M + 3$ |
| LCICM-AC | $6M^2 - 8M + 10$ | $6M^2 + 2M + 10$ |
| MLCICM-AC | $12M^2 - 16M + 13$ | $12M^2 + 4M + 12$ |

TABLE II
RADAR SYSTEM PARAMETERS

| Parameter | Value |
|---|---|
| Antenna array | Sideway-looking array (SLA) |
| Carrier frequency ($f_c$) | 1 GHz |
| Transmit pattern | Uniform |
| PRF ($f_r$) | 1000 Hz |
| Platform velocity ($v$) | 75 m/s |
| Platform height ($h$) | 9000 m |
| Clutter-to-Noise ratio (CNR) | 40 dB |
| Elements of sensors ($N$) | 10 |
| Number of Pulses ($J$) | 8 |

## IV. COMPLEXITY ANALYSIS

Here, we compare the computational complexity of our proposed algorithms with that of Stoica et al.'s [9] in terms of the number of additions and multiplications per snapshot. The comparison of the complexity is detailed in Table I. Note that the complexity of LCICM-AC algorithm is slightly lower than that of a conventional RLS algorithm since the computation of $\boldsymbol{\omega}_0$ and $\boldsymbol{\omega}_1$ is simpler than what would be required for RLS, given that $\mathbf{s}_t$ is fixed) and the complexity of MLCICM-AC algorithm is twice as much. Since computing $\mathbf{R}^{-1}$ with a Gauss-Jordan technique requires $M^3$ multiplications and $M^3 - 2M^2 + M$ additions [20], the complexity of Stoica et al.'s scheme is $\mathcal{O}(M^3)$ additions and multiplications.

## V. NUMERICAL EXAMPLES

In this section we assess the proposed KA-STAP-LCICM algorithms in an airborne radar application under certain heterogenous clutter conditions. The parameters of the radar platform are shown in Table II. We assume that the clutter-to-noise-ratio (CNR) is fixed at 40 dB and there is no jammer. Assuming that the calibration-on-clutter is known, the prior clutter covariance matrix can be calculated using these radar parameters. To model the heterogenous clutter, the spectral variation is introduced and target-like signals are Poisson-seeded over 300 training snapshots [4], [8]. We investigate the SINR performance against the number of snapshots for our proposed algorithm and the behavior of the mixing parameter $\eta$. In the simulation, the filter $\hat{\omega}$ is implemented by using the recursive least squares (RLS) algorithm with $\lambda = 0.998$. The parameters for the adaptive algorithm adjusting $\eta(n)$ are set as $\lambda_q = 0.5$, $\sigma_\eta = 0.1$ and $\mu_\eta = 0.1$. All presented results are averages over 1000 independent Monte-Carlo runs.

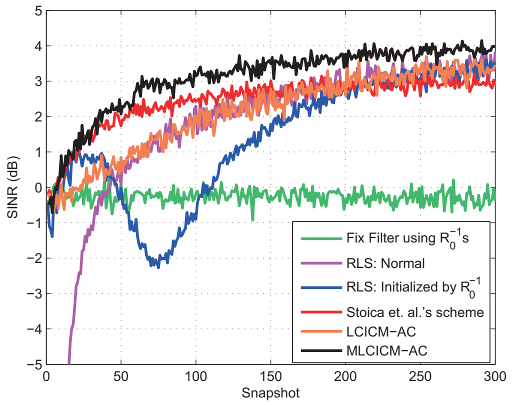


Fig. 3. The SINR performance against the number of snapshots for the proposed algorithm with $\lambda = 0.998$, $\lambda_q = 0.5$, $\sigma_\eta = 0.1$, $\mu_\eta = 0.1$.

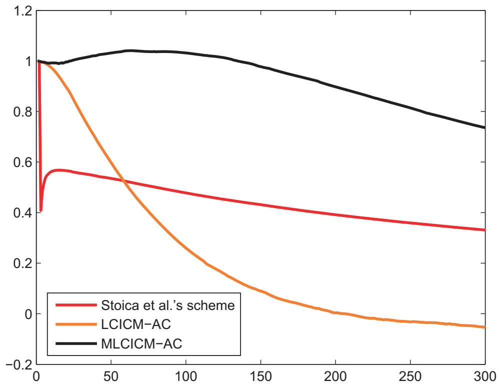


Fig. 4. Ensemble-average of $\eta(n)$ and $\alpha(n)$ for our proposed schemes and Stoica et al.'s scheme, respectively.

Fig. 3 shows the SINR performance against the number of snapshots for our proposed algorithms. We compare them with component filters and Stoica et al.'s scheme [9]. The curves show that our proposed MLCICM-AC algorithm outperforms Stoica et al.'s scheme almost 1.5dB at the steady-state and converges as quickly as Stoica et al.'s scheme. Note that the proposed MLCICM-AC algorithm has twice the complexity of a conventional RLS filter, which is still much less that that of Stoica et al.'s scheme. We also note that the proposed LCICM-AC scheme performs better than each component filters with computational complexity as low as a single RLS filter. In Fig. 4, we plot ensemble-average estimates of $\mathrm{E}\{\eta(n)\}$ and $\mathrm{E}\{\alpha(n)\}$ for our proposed schemes and Stoica et al.'s scheme. Note however that these parameters have different meanings, since $\eta(n)$ is the mixing parameter for the affine combination of inverse matrices and $\alpha(n)$ is the mixing parameter for the convex combination of matrices. We also note that the mixing parameter $\eta$ for the LCICM scheme converges to a steady-state value close to 0 and behaves quite differently than $\eta$ for the MLCICM scheme, although they are both adapted with the normalized algorithm based on the MV criterion.

## VI. Conclusions

In this paper, we have developed a KA-STAP algorithm to estimate the inverse interference covariance matrix rather than the covariance matrix itself by using the proposed LCICM scheme and combat the heterogeneous clutter effects. The computational load has been greatly reduced due to the avoidance of the matrix inversion operation. The LCICM scheme has been modified to improve convergence and steady-state performance. Moreover, adaptive algorithms for the mixing parameters have been developed for affine combinations in complex systems based on minimum variance criterion. Numerical examples have shown the potential of our proposed LCICM-AC and MLCICM-AC algorithms for substantial performance improvement.

## References


[1] L. E. Brennan and L. S. Reed, "Theory of adaptive radar," *IEEE Trans. Aero. Elec. Syst.*, vol. AES-9, no. 2, pp. 237–252, 1973.

[2] J. Ward, "Space-Time Adaptive Processing for Airborne Radar,", *Tech. Rep. 1015, MIT Lincoln lab., Lexington,MA*, Dec. 1994.

[3] R. Klemm, *Principle of Space-Time Adaptive Processing*, IEE Press, Bodmin, UK, 2002.

[4] W. L. Melvin, "Space-time adaptive radar performance in heterogeneous clutter," *IEEE Trans. Aero. Elec. Syst.*, vol. 36, no. 2, pp. 621–633, 2000.

[5] M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-time adaptive processing: a knowledge-based perspective for airborne radar," *IEEE Sig. Proc. Mag.*, vol. 23, no. 1, pp. 51–65, 2006, 1053-5888.

[6] W. L. Melvin and J. R. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, pp. 983–996, 2006, 0018-9251.

[7] R. Fa., R. C. de Lamare, and L. Wang, "Reduced-Rank STAP Schemes for Airborne Radar Based on Switched Joint Interpolation, Decimation and Filtering Algorithm", *IEEE Trans. on Signal Processing*, vol. 58 , No. 8, 2010, pp. 4182 - 4194.

[8] W. L. Melvin and G. A. Showman, "An approach to knowledge-aided covariance estimation," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, pp. 1021–1042, 2006, 0018-9251.

[9] P. Stoica, Li Jian, Zhu Xumin, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Trans. Sig. Proc.*, vol. 56, no. 6, pp. 2598–2602, 2008, 1053-587X.

[10] J. S. Bergin, C. M. Teixeira, P. M. Techau, and J. R. Guerci, "Improved clutter mitigation performance using knowledge-aided space-time adaptive processing," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, pp. 997–1009, 2006, 0018-9251.

[11] S. D. Blunt, K. Gerlach, and M. Rangaswamy, "STAP using knowledge-aided covariance estimation and the FRACTA algorithm," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, pp. 1043–1057, 2006, 0018-9251.

[12] N. J. Bershad, J. C. M. Bermudez, and J.-Y. Tourneret, "An affine combination of two LMS adaptive filters - transient mean-square analysis," *IEEE Trans. Signal Process.*, vol. 56, pp. 1853–1864, May 2008.

[13] R. Candido, M. T. M. Silva, and V. H. Nascimento, "Affine Combinations of Adaptive Filters," in *2008. Conf. Record of the Fourty-Second Asilomar Conf. Sig. Syst. Comp.*, 2008.

[14] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank MMSE Filtering with Interpolated FIR Filters and Adaptive Interpolators," *IEEE Signal Processing Letters*, vol. 12, no. 3, March 2005, pp. 177 - 180.

[15] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters", *IEEE Sig. Process. Lett.*, vol. 14, no. 12, pp. 980–983, 2007.

[16] R. C. de Lamare and R, Sampaio-Neto, "Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation and Filtering", *IEEE Trans. on Signal Processing*, vol. 57, no. 7, July 2009, pp. 2503 - 2514.

[17] R. Candido, M. T. M. Silva, and V. H. Nascimento, "Analysis of the affine combination of two adaptive filters: a unified approach," Submitted to *IEEE Trans. Signal Process.*, 2009.

[18] M. T. M. Silva, V. H. Nascimento, "Improving the Tracking Capability of Adaptive Filters via Convex Combination," *IEEE Trans. Signal Process.*,v.56, no.7, pp.3137-3149, July 2008.

[19] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, NJ, 2003.

[20] C. Meyer, "Matrix Analysis and Applied Linear Algebra," John Wiley, 2000, Chapter 3. 0018-9251.