Applications of Space-Time Adaptive Processing (STAP) in real situations requires dimension-reducing methods. This is due to both the large computational cost involved in calculating the interference statistics and the smaller number of stationary training samples available to estimate the clutter covariance. Recently, auto-regressive (AR) filtering techniques have been used to help reduce computation and secondary sample support requirements in STAP scenarios. In this paper, we compare the detection performance of several AR-based algorithms with more standard GLRT-type approaches. In particular, we consider the Parametric Amplitude Matched Filter (PAMF) and the Space-Time Auto-Regressive Filter (STAR), and show that they outperform standard GLRT tests, especially in challenging situations with low sample support. Among the parametric methods considered, the STAR approach provides the most robust overall performance.

I. INTRODUCTION

The aim of an airborne surveillance radar system is to determine the presence of a target amidst noise, ground clutter, and jamming signals. These interfering signals are typically much more powerful than the target signal, and must be nulled in order for a reliable detection to be performed. Space-time adaptive processing (STAP) has been shown [1] to effectively reduce the strength of the interference through the use of both temporal and spatial information.

In reality, it is impractical to implement many STAP tests due to the large amount of computation required. Additionally, due to non-homogeneity of the data, it is difficult to obtain enough secondary samples to “train” the algorithms (e.g., estimate the interference covariance). This is especially true when working with circular antenna arrays at close range. Besides the loss in performance that results from using too little secondary data, numerical errors can arise since the sample covariance estimate will be rank deficient. Reduced-dimension STAP detectors have been proposed to help speed up computation and reduce the amount of required secondary data to accurately predict clutter statistics [2]. These methods generally rely on some form of temporal or spatial focusing (beamforming, Doppler filters, FFT etc.) [3], and offer some improvement, but still do not perform well in extreme cases where very little secondary data is available.

To ameliorate performance in such extreme cases, the use of parametric models have been proposed in [4], [5]. These approaches require a relatively small number of parameters to be estimated when compared to the number required for a sample covariance. This leads to a reduction in computation and required secondary sample support, and dramatically improved performance. In this paper, we will present several adaptations of a parametric method proposed in [4] and referred to as Space-Time Auto-Regressive (STAR) filtering. This method, which can be thought of as a parametric matched subspace detector [7], uses a vector AR model to determine the subspace orthogonal to that in which the clutter and interference reside. Once this subspace is found, it is used to project the interference out of the data prior to detection. We will investigate three specific implementations of the STAR filter for detection, referred to here as the Auto-Regressive Filtered Generalized Likelihood Ratio Test (ARGLRT), Diagonally Loaded STAR (STARDL) test, and the STAR Prediction Error (STARPE) test. The performance of these techniques will be compared with the Generalized Likelihood Ratio Test GLRT [8], the Loaded GLRT (LGLRT) [2], and the Parametric Adaptive Matched Filter (PAMF) [5]. The parametric methods are shown to dramatically improve performance over conventional STAP detectors in situations where there is little secondary data available. Among the parametric methods considered, the STARPE approach provides the most robust overall performance.

The paper is organized as follows: Section II introduces the data model and notation used throughout the paper. In Section III, the GLRT and LGLRT are described, and their detection behavior as a function of the amount of available secondary data is illustrated. The STAR method is presented in Section IV, along with definitions for the PAMF, ARGLRT, STARPE, and STARDL tests. Simulation results are described in Section V, and some conclusions are discussed in Section VI.

II. DATA MODEL

We will assume a radar system with $M$ active elements and $N$ pulses per coherent processing interval (CPI). The data is arranged in an $MN$ dimensional vector $z_k$. We will denote the primary data vector (i.e., the data from the range bin we are testing) as $z = z_0$ and let $z_k, k = 1, \ldots, K$ denote the secondary data vectors available for use in the adaptive detector. The presence of a target in the primary data vector

\[ z = \begin{bmatrix} z_0 & z_1 & \cdots & z_K \end{bmatrix} \]

This work was supported by the Office of Naval Research under grant N00014-99-1-0692.
is denoted by the equation
\[ z = b_0 s + \epsilon \]  
with complex amplitude \( b_0 \), space-time steering vector \( s \), and interference \( \epsilon \). The interference is modeled by
\[ \epsilon = \epsilon_c + \epsilon_j + \epsilon_n \]
consisting of ground clutter \( \epsilon_c \), jammer interference \( \epsilon_j \), and complex zero-mean multivariate white Gaussian noise \( \epsilon_n \) with variance \( \sigma_n^2 \). The \( K \) secondary data vectors, written as
\[ z_k = \epsilon_k \text{ for } k = 1, \ldots, K, \]
are assumed to be target free. The interference vector is broken down into \( N \) individual \( M \)-element vectors received from each pulse:
\[ \epsilon_k = \begin{bmatrix} \epsilon_k(1) \\ \vdots \\ \epsilon_k(N) \end{bmatrix}. \]

According to Brennan’s rule \([1]\), the rank of the clutter for a uniform linear array (ULA) is approximated by \( \rho = M + (N - 1)\beta \), where \( \beta \) is a function of array platform velocity and pulse repetition frequency (PRF). \( \rho \) is generally much less than \( MN \) for a ULA. When dealing with circular arrays, the clutter rank increases and has been observed to be roughly two times that predicted by Brennan’s rule for a ULA \([9]\). The rank of the interference is important because it is used to estimate how many secondary samples will be needed to estimate the sample covariance.

### III. GLRT AND LGLRT

A well-known solution to the radar target detection problem is Kelly’s GLRT \([8]\), given by
\[ \frac{|s^* \hat{R}^{-1} z|^2}{s^* \hat{R}^{-1} s (1 + z^* \hat{R}^{-1} z)} \leq \eta, \]
where \( * \) denotes the conjugate transpose, \( s \) denotes the target steering vector, \( \eta \) is the threshold, and the sample covariance matrix is:
\[ \hat{R} = \frac{1}{K} \sum_{k=1}^{K} z_k z_k^* . \]

To achieve good performance with this detector, a reasonable estimate for \( \hat{R} \) is required, which in turn requires at least \( K \approx 2MN \). In many situations, especially when \( MN \) is large, this condition cannot be satisfied. In fact, \( K \) is often not large enough to form a full-rank estimate of \( \hat{R} \), or even one that would account for the clutter rank \( \rho \). In such cases, an alternative to the matrix inverse in (4) is necessary.

The Loaded GLRT (LGLRT) \([2]\) is one method that has been proposed for situations where the sample covariance is rank deficient. The LGLRT has the same form as (3), except \( \hat{R} \) is replaced with the diagonally-loaded covariance estimate
\[ \hat{\Sigma} = \frac{1}{K} \left( \sum_{k=1}^{K} z_k z_k^* + \kappa I \right), \]
where \( I \) is an \( MN \)-dimension identity matrix, and \( \kappa \) denotes the scalar diagonal loading factor. A criteria for choosing \( \kappa \) can be found in \([2]\). To illustrate the benefit of the LGLRT, we have plotted in Figure 1 Receiver-Operator Characteristic (ROC) curves with \( MN = 360 \) and \( K = 400 \) and 720 for both the GLRT and LGLRT. In the figure the higher curve for the GLRT is for \( K = 720 \) while the LGLRT is higher when \( K = 400 \). The sensitivity of the GLRT to small values of \( K \) is readily apparent. The fact that the LGLRT degrades in performance with large \( K \) is due to the fact that there are enough samples to estimate an accurate covariance, and by loading we are introducing noise rather than decreasing the computational error. In extreme cases, such as with a circular array at close range, assuming the availability of even 400 secondary data vectors is unreasonable. In such situations, the performance of even the LGLRT will be very poor, as illustrated later in the paper.

### IV. PARAMETRIC METHODS

In recent years, the use of AR filters has been extended to the vector case for use in multichannel radar detection \([4]\), \([5]\), \([10]\). In these papers, it is assumed that the clutter and interference follow a vector AR model over the multichannel vectors received from each pulse:
\[ \epsilon_k(t) + \hat{\epsilon}_k(t) = \sum_{i=1}^{L-1} A_i \epsilon_k(t-i), \quad t = 1, ..., N_e \]
where \( \hat{\epsilon}_k(t) \) is the residual error vector, \( N_e = N - L + 1 \), \( \epsilon_k(t) \) is an \( M \) dimensional vector sampled from range bin \( k \) at pulse \( t \), and \( A_i \) is the \( i \)th matrix filter tap. Because the parametric model depends on relatively few parameters \( (L \) and the elements of \( A_i \)), fewer secondary data vectors are required to estimate them. This is especially advantageous in situations where few homogeneous secondary samples are available.

Assuming the parameter \( L \) has been estimated, the matrix filter coefficients \( A_i \) can be determined by solving (6) via
a standard least-squares procedure. $K$ secondary data vectors are chosen from an interval surrounding the primary range bin. For example, if we let $k = 0$ denote the primary bin, then the indices of the secondary data vectors could be chosen as

$$k = [-K/2, \ldots, -1, 1, \ldots, K/2].$$

(7)

For notational convenience, we rewrite (6) as

$$\varepsilon_k(t) = \sum_{i=0}^{L-1} A_i \varepsilon_k(t-i) \quad t = 1, \ldots, N_k,$$

(8)

where $A_0$ is the identity matrix $I_M$, and we define the following:

$$A^* = \begin{bmatrix} A_0 & \cdots & A_{L-1} \end{bmatrix}$$

(9)

of size $M \times ML$.

$$\mathcal{E}_k = \begin{bmatrix} \varepsilon_k(1) & \cdots & \varepsilon_k(N_k) \\ \vdots & \ddots & \vdots \\ \varepsilon_k(L) & \cdots & \varepsilon_k(N) \end{bmatrix}$$

(10)

of size $ML \times (N_k)$, and

$$\mathcal{E} = \begin{bmatrix} \mathcal{E}_1 & \cdots & \mathcal{E}_K \end{bmatrix}$$

(11)

of size $ML \times (KN_c)$.

With these definitions, equation (8) may now be rewritten in vector notation using (9) and (11) as:

$$\varepsilon \triangleq A^* \mathcal{E} = \begin{bmatrix} \varepsilon_1 & \cdots & \varepsilon_N \end{bmatrix}$$

(12)

where $\varepsilon_k$ is the matrix

$$\varepsilon_k = \begin{bmatrix} \varepsilon_k(1) & \cdots & \varepsilon_k(N_k) \end{bmatrix}$$

With this notation defined, we may now proceed to the description of the PAMF and STAR autoregressive filtering methods.

A. Parametric Adaptive Matched Filter

After solving for the coefficients $A_i$, which presumably provide temporal whitening of the clutter and jamming, the PAMF approach forms the residual sample covariance

$$\mathcal{D}_A = \frac{1}{KN_c} A^* \mathcal{E} \mathcal{E}^* A^*$$

(13)

for use in spatially whitening the AR-filtered data:

$$\mathcal{D}_A^{-1/2} A^* = \begin{bmatrix} A_0^w & \cdots & A_{L-1}^w \end{bmatrix},$$

(14)

where the superscript $w$ is added to differentiate the whitened filter taps, $A_i^w = \mathcal{D}_A^{-1/2} A_i$ from those in (9). Defining the matrix

$$A^* = \begin{bmatrix} A_0^w & \cdots & A_{L-1}^w \\ A_0^w & \cdots & A_{L-1}^w \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_0^w \end{bmatrix}$$

(15)

allows us to write the PAMF test:

$$\frac{|s^* A^* A^* z|^2}{s^* A^* s(z^* A^* z)} = \frac{H_1}{H_0} \theta_{pamf}.$$  

(16)

We see that, in comparison with the GLRT and LGLRT, the matrix $A^* A^*$ takes the place of the inverse of the sample covariance in the detection test. The addition of 1 in the denominator is dropped as the PAMF was derived as an extension of the standard adaptive matched filter in [1], [11] rather than the GLRT.

B. Space-Time Auto-Regressive Filter

The STAR filtering approach is similar to that assumed for the PAMF, except that the coefficients in (8) are allowed to be non-square, and the zeroth tap is not constrained to be the identity. To avoid confusion with the PAMF filter coefficients, $A_i$, we will denote the STAR matrix filter taps as $H_i$. The coefficients $H_i$ are assumed to be $M' \times M$ matrices and $H_0$ is not constrained to be the identity $I_M$. Criteria for choosing $M'$ and the model order $L$ are given in [4].

In order to understand the approach taken by the STAR method, we use the definition of $\mathcal{E}$ in (11) and

$$H^* = \begin{bmatrix} H_0 & \cdots & H_{L-1} \end{bmatrix}$$

(17)

to write:

$$H^* \mathcal{E} = \varepsilon.$$  

(18)

The coefficients are chosen to minimize $\varepsilon$, noting that as this term approaches zero, $H$ can be thought of as approximating the left nullspace of $\mathcal{E}$. To guarantee a solution to (18) we must ensure enough secondary data vectors for $\mathcal{E}$ to have a nullspace of appropriate dimension (a similar constraint on $K$ is required by the PAMF). Typically, this constraint on $K$ is easily met, even in the most extreme cases.

The linear constraint $A_0 = I_M$ in the PAMF model leads to a solution involving the pseudo-inverse of $\mathcal{E}$. For the STAR method, the least-squares minimization of (18) also requires a constraint to avoid the trivial solution $H = 0$. Typically, the quadratic constraint $H^* H = I$ is used, which means that the solution to

$$\tilde{H} = \arg\min_H \|H^* \mathcal{E}\|^2_F$$

(19)

is simply the $M'$ left singular vectors of $\mathcal{E}$ corresponding to the smallest singular values.

In principle, the STAR method eliminates the second whitening step (14) required by PAMF. Since $H_0$ is not constrained to be the identity, both temporal and spatial whitening occur in one step. To describe the detection tests based on the STAR filter, we form a matrix $\mathcal{H}^*$ similar to (15):

$$\mathcal{H}^* = \begin{bmatrix} H_0 & \cdots & H_{L-1} \\ 0 & \cdots & 0 \end{bmatrix}.$$  

(20)

As shown in [4], $\mathcal{H}^*$ is approximately orthogonal to the subspace spanned by the clutter and interference. Three uses of $\mathcal{H}^*$ to null clutter and jamming prior to detection are now described.
1) **AR-filtered GLRT**: Our first use of $\mathcal{H}$ in deriving a new detection strategy comes from adapting the LGLRT to the STAR filter. In particular, the STAR filter is used to first eliminate the clutter and interference in the secondary data prior to formation of the sample covariance:

$$\hat{z}_k = \mathcal{H}^* y_k .$$

(21)

As with the LGLRT, the sample covariance is diagonally loaded

$$\hat{R}_H = \sum_{k=1}^{K} \hat{z}_k \hat{z}_k^* + \kappa I ,$$

(22)

where, in the simulations, the loading factor $\kappa$ is chosen according to the method shown in [2]. This leads to the following test statistic that we refer to as the AR-filtered GLRT (ARGLRT):

$$\frac{\hat{s}^* \hat{R}_H^{-1} \hat{s}^2}{\hat{s}^* \hat{R}_H^{-1} \hat{s}(1 + \hat{z}^* \hat{R}_H^{-1} \hat{z})} \frac{\gamma}{\eta_{ar}} .$$

(23)

In our simulation studies, the ARGLRT is shown to perform best in situations where there are ample stationary secondary data vectors available, but it significantly outperforms the GLRT and LGLRT as $K$ decreases.

2) **Prediction Error STAR**: The motivation for deriving a prediction error filter based on STAR comes from [12], [13]. The STARPE is written as

$$\frac{|s^* \mathcal{H}^* \mathcal{H} s|^2}{s^* \mathcal{H}^* \mathcal{H} s(1 + s^* \mathcal{H}^* \mathcal{H} s)} \frac{\gamma_{ar}}{\gamma_{pe}} .$$

(24)

where we see that $\mathcal{H}^* \mathcal{H}$ takes the place of $\mathbf{R}^{-1}$ in the GLRT. While the STARPE approach has slightly poorer performance than the other parametric methods we consider, it has the considerable advantage of not requiring the inverse of a (typically large) matrix. The STARPE method does significantly outperform the GLRT and LGLRT in situations with extremely low secondary sample support.

3) **Diagonally Loaded STAR**: Since $\mathcal{H}^*$ is formed from the subspace orthogonal to clutter, [4] took the matched subspace approach of [7] and formed the projection matrix $\mathcal{H}((\mathcal{H}^* \mathcal{H})^{-1} \mathcal{H}^*)$ to filter the data. The filter was observed to have excellent SINR performance in [4] when $M'$ was chosen appropriately, but its performance was observed to be quite sensitive to the value of this parameter. In particular, if $M'$ is chosen to be too large, the inverse in the projection becomes numerically ill conditioned. Therefore, [4] proposed a compromise between the matched subspace and prediction error approaches by introducing a diagonal loading factor $\alpha$. This loading factor is chosen to be small enough that it does not affect the larger eigenvalues of the matrix $\mathcal{H}^* \mathcal{H}$ which correspond to the clutter and interference, but large enough to improve the condition number of the matrix. The STARDL test normalized to correspond to the GLRT is defined as follows:

$$\frac{\gamma^* (\mathcal{H}^* \mathcal{H} + \alpha I)^{-1} \gamma^2}{\gamma^* (\mathcal{H}^* \mathcal{H} + \alpha I)^{-1} \gamma(1 + \gamma^* (\mathcal{H}^* \mathcal{H} + \alpha I)^{-1} \gamma)} \frac{\gamma_{ar}}{\gamma_{dl}} .$$

(25)

The diagonally loaded STAR filter is robust to estimation of the model order $L$ as well the choice of $M'$. As discussed in the next section, it performs well in extreme situations where small amounts of secondary data are available and also in normal situations where adequate amounts of secondary data are available.

V. RESULTS

Each of the algorithms discussed in this paper were compared to one another using ROC curves obtained through numerical simulations. The algorithms were tested on simulated airborne radar data produced at MIT Lincoln Labs. The data simulates a circular array with $M = 20$ antennas and $N = 15$ pulses per CPI. We tested the algorithms at short range (30 km) and long range (300 km), and we also tested each with both extremely low sample support ($K = 10$) and with larger sample support ($K = 200$). The results are plotted in Figures 2, 3, 4, and 5.

For each of the simulations, the model order $L$ was selected using the selection criteria given in [4]. The loading factor for STARDL method was chosen to be either 1/10 or 1/20 the size of the largest diagonal value of $\mathcal{H}^* \mathcal{H}$, with the smaller value used when $K$ was large. The loading factor for the LGLRT was chosen according to the approach described in [2]. The simulations show that the STARDL method generally outperforms the other tests in each of the four cases, at close or long range, with small or larger sample support. It is clear from the figures that tests based on the sample covariance, filtered (ARGLRT) or not (LGLRT), are very dependent on the amount of available sample support. However, the ARGLRT is considerably more robust to small values for $K$ than the LGLRT.

VI. CONCLUSION

We have presented three new STAP detection criteria based on parametric clutter models. The methods differ in how the resulting AR filter is applied to the secondary data prior to forming the test statistic. These parametric techniques offer several distinct improvements over traditional detection tests. Because the models provide a relatively parsimonious parameterization of the clutter, only a small amount of secondary data is required to estimate them, much smaller than that required by traditional methods. The STARDL approach offers the best overall performance, slightly better than that achieved by the PAMF. The STARPE method performs less well, but requires significantly less computation to implement. Implementing the LGLRT after AR filtering provides a significant gain over the standard LGLRT implementation, but it still requires significant secondary data in order to form an accurate sample covariance estimate.

544
ROC Curves (Range: 30 Km)

Fig. 2. ROC curves for short range and small sample support

ROC Curves (Range: 300 Km)

Fig. 4. ROC curves for long range and small sample support

Fig. 3. ROC curves for short range and larger sample support

Fig. 5. ROC curves for long range and larger sample support

REFERENCES