Research Article

Fully Adaptive Clutter Suppression for Airborne Multichannel Phase Array Radar Using a Single A/D Converter

Dan Madurasinghe and Andrew P. Shaw

Electronic Warfare and Radar Division, Defence Science and Technology Organisation, Edinburgh, SA 5111, Australia

Correspondence should be addressed to Dan Madurasinghe, dan.madurasinghe@dsto.defence.gov.au

This study considers an airborne multichannel phase array radar consisting of an analog phase shifter on each channel, where the sum channel (output) is digitised using a single A/D converter. Generally for such a configuration, the array weights are predetermined for each transmit/receive direction and are nonadaptive to the clutter. In order to achieve any adaptivity to the environment, the convention is to split the array into at least two subgroups and implement two analogs to digital converters. A single A/D-based software solution (numerically stable, robust) is proposed to achieve the full sidelobe adaptation to clutter. The proposed algorithm avoids these engineering complications involved in implementing multiple A/Ds for radar applications while maintaining the same desired performance. As a large number of airborne radar platforms already exist worldwide, the possible applications of this proposed fully adaptive upgrade as a software solution can be huge.

1. Introduction

The objective of an adaptive array is to combine the elemental outputs, appropriately weighted so as to generate an output that is interference free. To achieve this we need to have observations from a sufficient number of channels of the array that we can use to calculate the adapted weights [1–3]. If a "traditional" analog beamformer is employed, then it is not usually possible to observe the individual channels. If multiple beamforming manifolds are used, it is possible to compute an adaptation in beamspace, but in most cases only a small number of beams are produced severely restricting the number of interfering sources that can be accommodated. In practice this is further complicated because "real" arrays, especially with near-field scatterers, do not have uniform elements.

There are a number of engineering advantages to employing an analog beamformer, particularly related to the number of digitisers employed and the consequential simplification in all those processes associated with digitisers (maintaining alignment, power consumption/cooling, and data management), but if low sidelobe performance is required, this is offset by the increased difficulty in calibration of the array, especially for active arrays, where effective impedance of path depends upon the frequency, power on/off, and phase status of adjacent elements. Current capabilities are such as to favour the use of analog beamforming to produce a small number of beams, typically a single sum, also known as a "sigma" beam, and additionally a number of difference beams, also known as "delta" beams, and then either (a) sacrifice low sidelobe performance; (b) require complex calibration; or (c) attempt to mitigate the sidelobes with limited adaptive processing, such as "sigma-delta" processing [4] or other forms of reduced-dimension adaptive processing.

This study considers a phased array wherein we can adjust the amplitude and phase of each element, but where we can only observe the output of a single "sum" channel, and introduces an algorithm on this channel to adaptively null any residual sidelobe clutter. The method described in this paper transmits $2 \times N_p$ pulses in each beam direction. Firstly coherent N_p burst of pulses are received using an initial set of antenna weights. Then, after allowing for a switching delay, a second burst of N_p pulses are received using a set of weights that are linearly independent, whilst satisfying certain requirements. The new algorithm

developed in this paper uses the properties of the data stream to adaptively null the ground clutter with N_p degrees of freedom. The procedure we have developed is tested using both simulated data and data from the MCARM system [5], suitably processed to represent a single "sum" beam, including the delay caused by the switching of the antenna weights. The results obtained are then compared with the fully adaptive solution available via mutlichannel data with the same number of degrees of freedom.

This paper is organised as follows. In Section 2, we formulate the standard multichannel problem and consider multichannel observation-based signal processing gains (full STAP, beamspace STAP, etc.) to provide a baseline for comparison. Section 3 formulates the proposed software solution using a single observation channel and derives the signal processing gain. Section 4 examines the theoretical performances and compares the algorithms using Monte Carlo simulation. Finally Section 5 uses MCARM data to validate the results.

2. Formulation

2.1. General Formulation. Assume that the airborne platform travels in the positive y-direction at speed V_A (Figure 1), x is the array broadside direction, ϕ is the azimuth angle measured from the array broadside, and θ is the elevation angle where $\theta > 0$ corresponds to $z > 0$. Suppose that we have a planer array of N elements, which transmits and receives a burst of N_p coherent pulses. The measured $N \times 1$ signal vector $\mathbf{x}^{(m)}(r)$ due to the mth coherent pulse and rth range ring, which is also referred to as the fast time scale, can be expressed as

$$\mathbf{x}^{(m)}(r) = \sum_{k=1}^{N_s} \alpha_{k,r} \mathbf{s}_N(\phi_k, \theta_r) \exp\left(j2\pi f_{k,r}(m-1)T_p\right)$$
$$+ \delta(r - r_t)\alpha_t \exp\left(j2\pi f_t T_p(m-1)\right) \quad (1)$$
$$\times \mathbf{s}_N(\phi_t, \theta_{r_t}) + \mathbf{e}^{m,r},$$
$$m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2,$$

where $\mathbf{s}_N(\phi, \theta)$ is the $N \times 1$ array steering vector, $\alpha_{k,r}$ is the received complex clutter amplitude due to the kth scatterer also referred to as clutter discrete on the rth range ring, $(\phi_k, \theta_r, f_{k,r})$ is azimuth, elevation and Doppler frequency, respectively, of the kth scatterer on the rth range ring, $f_{k,r} = 2V_A \sin\phi_k \cos\theta_r/\lambda$ is the Doppler return due to a scatterer, λ is the wavelength of the carrier, N_s is the total number of scatterers on any range ring, T_p is the pulse repetition interval (PRI), r_t is the range ring index corresponding to the target range cell, α_t is the received signal amplitude due to the target, f_t is the Doppler frequency of the target, ϕ_t is the target azimuth, θ_{r_t} is the target elevation, $\delta(r)$ is the Kronecker delta function, and N_1, N_2 are the range indices corresponding to the nearest range ring on the ground and the furthermost ring on the ground, respectively. The

random component of the received signal vector has the following structure:

$$\mathbf{e}^{m,r} = (e_1^{m,r}, e_2^{m,r}, \ldots, e_N^{m,r})^T, \quad (2)$$

where $e_i^{m,r}(i = 1, 2, \ldots)$ represents a series of independent and identical Gaussian random variables and the superscript T denotes the vector transpose. In addition to that, we assume that

$$E\left\{|e_i^{m,r}|^2\right\} = \sigma_n^2, \quad i = 1, 2, \ldots N, \quad (3)$$

where σ_n^2 is the noise variance and $E\{\cdot\}$ denotes the expectation operator. The usual assumptions such as patch-to-patch statistical independence (zero-mean Gaussian) are made on the clutter as well as target. The data cube defined in (1) is of the size $N \times N_p \times (N_2 - N_1 + 1)$ that is generally known as a CPI data cube. The total clutter power on the ground before applying the transmit or receive tapering is

$$p(r) = \sum_{k=1}^{N_s} |\alpha_{k,r}|^2, \quad (4)$$

per range ring. It should be noted that traditionally the received data stream when observed via a single receiver after analog beamforming is represented by

$$\mathbf{w}^H \mathbf{x}^{(m)}(r) = \sum_{k=1}^{N_s} \alpha_{k,r} \mathbf{w}^H \mathbf{s}_N(\phi_k, \theta_r) u(f_{k,r})^{m-1} + \delta(r - r_t)$$
$$\times \alpha_t \mathbf{w}^H \mathbf{s}_N(\phi_t, \theta_{r_t}) u(f_t)^{m-1} + \mathbf{w}^H \mathbf{e}^{m,r},$$
$$m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2,$$
$$(5)$$

where $u(f) = \exp(j2\pi f T_p)$ and $\mathbf{w}$ represents the received weights vectors which are chosen to satisfy $\mathbf{w}^H \mathbf{s}_N(\phi_t, \theta_{r_t}) = 1$. The simplest beamforming choice is the uniform weights given by $\mathbf{w} = (1/N)\mathbf{s}_N(\phi_t, \theta_{r_t})$, and here we have ignored transmit pattern effect. The above data stream is then passed through a Fast Fourier Transform (FFT) processor to obtain the output for each Doppler bin of interest. In the presence of clutter the performance is reduced severely.

2.2. Adaptive Solutions (STAP). In order to achieve full adaptivity to the clutter, generally the radar system has to undergo a multiple-A/D (hardware) upgrade where a number of sampled data streams are made available. However, for practical implementation, typically one would apply some of the degrees of freedom nonadaptively via Pre Doppler STAP, Post Doppler STAP, or Beamspace STAP, in order to simplify the computations and inversion of the covariance matrix. This will not lower the performance significantly of the system providing the number of adaptive degrees of freedom sufficient to null the number of interference signals present in the system due to clutter-related arrivals, and the results are well documented in the literature[1, 2].

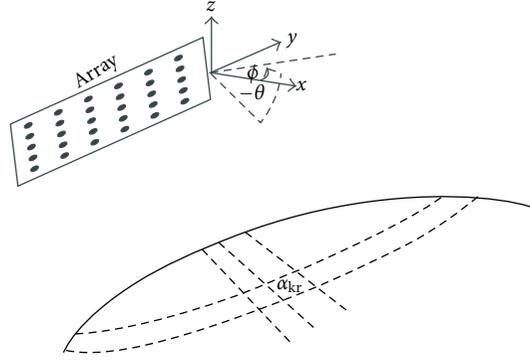In order to compare systems we will develop the necessary formulas for at least one multiple-A/D-based reduced

FIGURE 1: Airborne array with axis system.

STAP solution referred to as Beamspace STAP, where the number of adaptive channels is reduced to a manageable size, and then apply STAP on the reduced system using all available coherent pulses, giving us sufficient adaptive degrees of freedom. Suppose that N_B is the number of digitised channels we would like the system to be reduced to; then we apply $N_R \times 1$ ($N_R = (N - N_B + 1)$) weights vectors $\mathbf{w}_t^{(j)}$ ($j = 1, 2, \ldots, N_B$), to subarrays consisting of elements $1, 2, \ldots, N_R$, as the first subarray ($j = 1$), the elements $2, 3, 4, \ldots, (N_R + 1)$, as the second subarray ($j = 2$), and so forth, and finally the elements $(N - N_R + 1)$ to N as the last subarray ($j = N_B$). One obvious choice is $\mathbf{w}_t^{(j)} = (1/N_R)\mathbf{s}_{N_R}^{(j)}(\phi_t, \theta_{r_t})$ representing uniform array weights suitable for the jth subarray, where $\mathbf{s}_{N_R}^{(j)}(\phi_t, \theta_{r_t})$ denotes the $N_R \times 1$ steering vector to represent the jth subarray which consists of the entries taken from $\mathbf{s}_N(\phi, \theta)$ starting from jth to $(N_R - 1 + j)$th positions. This reduces the original $N \times 1$ data vector in (1) to the $N_B \times 1$ data vector $\tilde{\mathbf{x}}^{(m)}(r) = (\tilde{x}_1^{(m)}(r), \tilde{x}_2^{(m)}(r), \ldots, \tilde{x}_{N_B}^{(m)}(r))^T$, requiring only N_B digital receivers, where

$$\tilde{x}_j^{(m)}(r) = \sum_{k=1}^{N_s} \alpha_{k,r}\left[\mathbf{w}_t^{(j)H}\mathbf{s}_{N_R}^{(j)}(\phi_k, \theta_r)\right]\exp\left(j2\pi f_{k,r}(m-1)T_p\right)$$

$$+ \delta(r - r_t)\alpha_t \exp\left(j2\pi f_t T_p(m-1)\right)$$

$$\times \left[\mathbf{w}_t^{(j)H}\mathbf{s}_{N_R}^{(j)}(\phi_t, \theta_{r_t})\right] + \hat{\mathbf{w}}_t^{(j)H}\mathbf{e}^{m,r},$$

$$m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2,$$

$$j = 1, 2, \ldots, N_B,$$

(6)

where $\mathbf{w}_t^H \mathbf{s}_{N_R}^{(j)}(\phi_t, \theta_{r_t}) = 1$, $j = 1, 2, \ldots, N_B$, and $\hat{\mathbf{w}}_t^{(j)H}$ is the jth row of

$$\mathbf{W}_t^H = \begin{pmatrix} \mathbf{w}_t^{(1)H} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{w}_t^{(2)H} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{w}_t^{(N_B)H} \end{pmatrix}_{N_B \times N}.$$

(7)

The digitised $N_B \times 1$ data stream can be expressed as

$$\tilde{\mathbf{x}}^{(m)}(r) = \sum_{k=1}^{N_s} \alpha_{k,r_t}\tilde{\mathbf{s}}_{N_B}(\phi_k, \theta_r)\exp\left(j2\pi f_{k,r}(m-1)T_p\right)$$

$$+ \delta(r - r_t)\alpha_t \exp\left(j2\pi f_t T_p(m-1)\right)$$

(8)

$$\times \tilde{\mathbf{s}}_{N_B}(\phi_t, \theta_{r_t}) + \mathbf{W}_t^H \mathbf{e}^{m,r},$$

$$m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2,$$

where $\tilde{\mathbf{s}}_{N_B}(\phi_k, \theta_r) = \mathbf{W}_t^H \mathbf{s}_N(\phi_k, \theta_r)$ is the equivalent ($N_B \times 1$) spatial steering manifold for the new data vector. After stacking all N_p pulse returns to form a new reduced $N_B N_p \times 1$ space-time snapshot, we have

$$\tilde{\mathbf{Y}}(r) = \left(\tilde{\mathbf{x}}^{(1)}(r)^T, \tilde{\mathbf{x}}^{(2)}(r)^T, \tilde{\mathbf{x}}^{(3)}(r)^T, \ldots, \tilde{\mathbf{x}}^{(N_p)}(r)^T\right)^T$$

$$= \delta(r - r_t)\alpha_t \mathbf{v}_1(\phi_t, \theta_{r_t}, f_t)$$

(9)

$$+ \sum_{k=1}^{N_s} \alpha_{k,r}\mathbf{v}_1(\phi_k, \theta_r, f_{k,r}) + \left(\mathbf{I}_{N_p} \otimes \mathbf{W}_t^H\right)\tilde{\mathbf{e}},$$

where $\mathbf{v}_1(\phi, \theta, f) = \mathbf{t}_s(f) \otimes \tilde{\mathbf{s}}_{N_B}(\phi, \theta)$ is the space-time steering vector (manifold) [1–3], $\mathbf{t}_s(f) = (1, u(f), u(f)^2, \ldots, u^{N_p-1}(f))^T$, and $\tilde{\mathbf{e}} = ((\mathbf{e}^{1,r})^T, (\mathbf{e}^{2,r})^T, \ldots, (\mathbf{e}^{N_p,r})^T)^T$. This data stream allows us to apply $N_B N_p$ degrees of freedom adaptively to form the STAP output. When sample matrix inversion-based solution is used, the output signal to clutter plus noise ratio is given by [1]

$$\text{SCN}(\phi_t, \theta_{r_t}, f_t) = |\alpha_t|^2 \mathbf{v}_1(\phi_t, \theta_{r_t}, f_t)^H \tilde{\mathbf{R}}_I^{-1}\mathbf{v}_1(\phi_t, \theta_{r_t}, f_t),$$

(10)

where the covariance matrix is defined as [1, 6]

$$\tilde{\mathbf{R}}_I = E\left\{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^H\right\} = E\left\{|\alpha_{k,r}|^2\right\}\mathbf{v}_1(\phi_k, \theta_r, f_{k,r})\mathbf{v}_1(\phi_k, \theta_r, f_{k,r})^H$$

$$+ \sigma_n^2\left(\mathbf{I}_{N_p} \otimes \mathbf{W}_t^H\right)\left(\mathbf{I}_{N_p} \otimes \mathbf{W}_t^H\right)^H.$$

(11)

This is estimated by the formula

$$\tilde{\mathbf{R}}_I \approx \sum_{r=N_1}^{r=N_2} \frac{\tilde{\mathbf{Y}}(r)\tilde{\mathbf{Y}}(r)^H}{N_T},$$

(12)

where $N_T = (N_2 - N_1 + 1)$ is the number of range cells used for averaging. It should be noted that $N_B = N$ is equivalent to full STAP solution requiring an A/D for each channel, which allows us to use NN_p adaptive degrees of freedom.

3. Multi-Transmit Receive STAP (MTR-STAP)

3.1. Proposed Software Solution (MTR-STAP). We now consider a system where only one digitised sum channel is available. Assume that the radar transmits and receives a burst of N_p coherent pulses with a certain set of array receiver weights and a second burst is transmitted and received with a different set of receiver weights. Both transmissions are aimed in the same direction; hence clutter return is related to the same patch on the ground, and transmission weights are not relevant as long as the desired direction is sufficiently illuminated (Figure 2). The $N \times 1$ receiver weights vectors $\mathbf{w}_A$ and $\mathbf{w}_B$ are different and to be determined later. The aim is to look at the changes we need to accommodate in order to represent two consecutive data streams, where the transmission of the second burst begins after t_0 (seconds) time delay. This delay time is the switching time allowed to change the received array weights (phase shifters). The second coherent burst is $T_p N_p$ seconds long. The total pulse length for two bursts is $2T_p N_p + t_0$. As seen later, t_0 is selected to be a multiple of T_p. This way we can maintain the transmission as a single train of $2N_p + 1$ pulses for $t_0 = T_p$. In this case the receiver simply changes the phase weights during the switching period and resumes colleting data for the second stream. Noting that r represents the digitised version of the time axis, let us represent the return signal due to any of the clutter patches for the first data stream for the N-element array as $\alpha \mathbf{s}_N(\phi, \theta) \exp[j2\pi(f_c + f_d)t]$, where α is a complex constant to describe the reflective properties of the target or the ground patch, (ϕ, θ) represents the angle of arrival pair, f_c is the radar carrier frequency, and f_d is the Doppler component of this ground patch. After down converting to baseband (i.e., $\times e^{-j2\pi f_c t}$), we have the received $N \times 1$ signal as $\mathbf{x}_1(t)$, where

$$\mathbf{x}_1(t) = \alpha \mathbf{s}_N(\phi, \theta) \exp[j2\pi f_d t]. \tag{13}$$

After applying the analog beamformer, the data stream will be digitised with two time scales generally known as the slow time scale (pulse to pulse) and the fast time scale (range index). This is represented by writing $t = t_s + r\Delta + (m-1)T_p$, where $r = 1, 2, 3, \ldots, N_2$, N_2 is the total possible number of range gates for each value of m, $m = 1, 2, \ldots, N_p$ represent the slow time scale (mth pulse), Δ is the time resolution of the digitizer, and t_s is an unknown reference time point or the starting point. On the other hand, the data points of the second stream is measured by $t = t_s + (t_0 + N_p T_p) + r\Delta + (m-1)T_p$ where $r = 1, 2, 3, \ldots, N_2$, $m = 1, 2, \ldots, N_p$, and $t_s + (t_0 + N_p T_p)$ is replaced as the starting point with $(t_0 + N_p T_p)$ being the total delay. This is the time it took to complete the first burst plus the switching time. Applying the time scales to (13), we have the patch contribution $N \times 1$ data vector which

is the received signal for the mth pulse rth range gate, before combining to form a single stream as

$$\begin{aligned} \mathbf{x}_1^{(m)}(r) &= \alpha \mathbf{s}_N(\phi, \theta) \exp\Big[j2\pi f_d\big(t_s + r\Delta + (m-1)T_p\big)\Big] \\ &= \big[\alpha \exp(j2\pi f_d(t_s + r\Delta))\big]\mathbf{s}_N(\phi, \theta) \\ &\quad \times \exp\Big[j2\pi f_d(m-1)T_p\Big] \\ &= \alpha_{k,r}\mathbf{s}_N(\phi_k, \theta_r) \exp\Big[j2\pi f_{k,r}(m-1)T_p\Big], \end{aligned} \tag{14}$$

where we have made the comparison with the patch return in (1) by analogy $\alpha_{k,r} = \alpha \exp(j2\pi f_d(t_s + r\Delta))$, $(\phi, \theta) = (\phi_k, \theta_r)$, and $f_d = f_{k,r}$ for any such patch denoted by indices k, r, and the contribution due to the same patch but for the second data stream is given by

$$\begin{aligned} \mathbf{x}_2^{(m)}(r) &= \alpha \mathbf{s}_N(\phi, \theta) \\ &\quad \times \exp\Big[j2\pi f_d\big(t_s + t_0 + N_p T_p + r\Delta + (m-1)T_p\big)\Big] \\ &= \alpha_{k,r} \exp\Big[j2\pi f_{k,r}\big(t_0 + N_p T_p\big)\Big]\mathbf{s}_N(\phi_k, \theta_r) \\ &\quad \times \exp\Big[j2\pi f_{k,r}(m-1)T_p\Big] \\ &= \alpha_{k,r}\hat{\mathbf{s}}_N(\phi_k, \theta_r, f, t_0) \exp\Big[j2\pi f_{k,r}(m-1)T_p\Big], \end{aligned} \tag{15}$$

where $\hat{\mathbf{s}}_N(\phi, \theta, f, t_0) = \mathbf{s}_N(\phi, \theta)\rho(t_0, f)$, with $\rho(t_0, f) = \exp[j2\pi f(t_0 + N_p T_p)]$. The vector $\hat{\mathbf{s}}(\phi, \theta, f, t_0)$ can be considered as the secondary receivers spatial component of the steering vector of size $N \times 1$ which is synchronised to the same coherent clock as the first transmission. This is equivalent to the original spatial steering vector, but, it is a function of the angle of arrival, the Doppler frequency of interest, the switching delay, and the pulse repetition interval, related to the target or clutter patch of interest.

Before proceeding any further, one has to notice that, apart from the familiar ambiguities of the usual spatial steering manifold defined in (1), we have a new ambiguity that is present in the secondary steering manifold due to the switching delay t_0 given by the following formula:

$$\hat{\mathbf{s}}_N(\phi, \theta, f, t_0) = \hat{\mathbf{s}}_N\left(\phi, \theta, f, t_0 \pm \frac{n}{f}\right), \quad n = 1, 2, 3, \ldots. \tag{16}$$

Just as we avoid the spatial ambiguity by restricting our array spacing to half-wavelength, we can avoid this ambiguity by restricting the switching delay t_0 to less than one PRI ($= T_p$), because, in order to avoid Doppler ambiguities, we already have the restriction of possible Doppler frequencies to $(-1/(2T_p), +1/(2T_p))$. In any case, if one ever needs to resolve this ambiguity, the next possible value of the switching time is $t_0 \pm T_0$ $(T_0 > 2T_p)$, for some T_0. A procedure is developed later to estimate the switching time delay t_0 very accurately subject to the above ambiguity.

3.2. Properties of the Two Data Streams. Suppose that the first data stream uses the complex phase shifter weights ($N \times 1$ vector) $\mathbf{w}_A(\phi_t, \theta_r)$, with the property, $\mathbf{w}_A^H \mathbf{s}_N(\phi_t, \theta_r) = 1$ and the auxiliary (2nd receiver) data stream uses the $N \times 1$ weights vector $\mathbf{w}_B(\phi_t, \theta_r)$, with the property: $\mathbf{w}_B^H \mathbf{s}_N(\phi_t, \theta_r) = 1$ where (ϕ_t, θ_r) is the Tx/Rx direction (Figure 2). Here the target direction or look direction is ϕ_t, but the presence of the range cell of interest (its elevation) is maintained throughout the analysis as by θ_r, since all range cells are interrogated generally and $r = r_t$ contains a target for illustration when needed. From (1), for the first data stream we have

$$
\begin{aligned}
y_1^{(m)}(r) &= \mathbf{w}_A^H \mathbf{x}^{(m)}(r) \\
&= \sum_{k=1}^{N_s} \alpha_{k,r} \left[\mathbf{w}_A^H \mathbf{s}_N(\phi_k, \theta_r) \right] \exp\left(j2\pi f_{k,r}(m-1)T_p \right) \\
&\quad + \delta(r - r_t)\alpha_t \exp\left(j2\pi f_t T_p(m-1) \right) + \mathbf{w}_A^H \mathbf{e}^{m,r}, \\
&\quad\quad m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2,
\end{aligned}
\tag{17}
$$

and for the second (received) data stream we have

$$
\begin{aligned}
y_2^{(m)}(r) &= \sum_{k=1}^{N_s} \alpha_{k,r} \left[\mathbf{w}_B^H \mathbf{s}_N(\phi_k, \theta_r)\rho(t_0, f_{k,r}) \right] \\
&\quad \times \exp\left(j2\pi f_{k,r}(m-1)T_p \right) \\
&\quad + \delta(r - r_t)\alpha_t \exp\left(j2\pi f_t T_p(m-1) \right) \\
&\quad \times \left[\rho(t_0, f_t) \right] + \mathbf{w}_B^H \mathbf{e}^{m,r}, \\
&\quad\quad m = 1, 2, \ldots, N_p, \quad r = N_1, \ldots, N_2.
\end{aligned}
\tag{18}
$$

It should be noted that the first nonadaptive stage of this spatial filtering may eliminate some of the clutter points depending on the choice of $\mathbf{w}_A$, $\mathbf{w}_B$ since the patterns $\mathbf{w}_A^H \mathbf{s}(\phi, \theta)$ and $\mathbf{w}_B^H \mathbf{s}(\phi, \theta)$) generally contain a considerable number of nulls in the (ϕ, θ) domain. The spatially stacked 2×1 data vector corresponding to the mth pulse is expressed as

$$
\begin{aligned}
\begin{pmatrix} y_1^{(m)}(r) \\ y_2^{(m)}(r) \end{pmatrix} &= \sum_{k=1}^{N_s} \alpha_{k,r} \left[\mathbf{w}_A^H \mathbf{s}_N(\phi_k, \theta_r) \right] \begin{pmatrix} 1 \\ F(\phi_k, \theta_k)\rho(t_0, f_{k,r}) \end{pmatrix} \\
&\quad \times \exp\left(j2\pi f_{k,r}(m-1)T_p \right) + \delta(r - r_t) \\
&\quad \times \alpha_t \begin{pmatrix} 1 \\ \rho(t_0, f_t) \end{pmatrix} \exp\left(j2\pi f_t T_p(m-1) \right) \\
&\quad + \mathbf{w}_n^H \mathbf{e}^{m,r}, \quad m = 1, 2, \ldots, N_p, \\
&\quad\quad r = N_1, \ldots, N_2,
\end{aligned}
\tag{19}
$$

where $F(\phi, \theta) = \mathbf{w}_B^H \mathbf{s}_N(\phi, \theta) / \mathbf{w}_A^H \mathbf{s}_N(\phi, \theta)$ is the receive patterns ratio with the property $F(\phi_t, \theta_r) = 1$, $\mathbf{w}_n^H = \begin{pmatrix} \mathbf{w}_A^H & \mathbf{O}_{1 \times N} \\ \mathbf{O}_{1 \times N} & \mathbf{w}_B^H \end{pmatrix}$ is a combined weights matrix of size $2 \times 2N$, $\mathbf{e}^{m,r}$

represents a $2N \times 1$ independent random entries, and $\mathbf{O}_{1 \times N}$ is the $1 \times N$ matrix of zero entries.
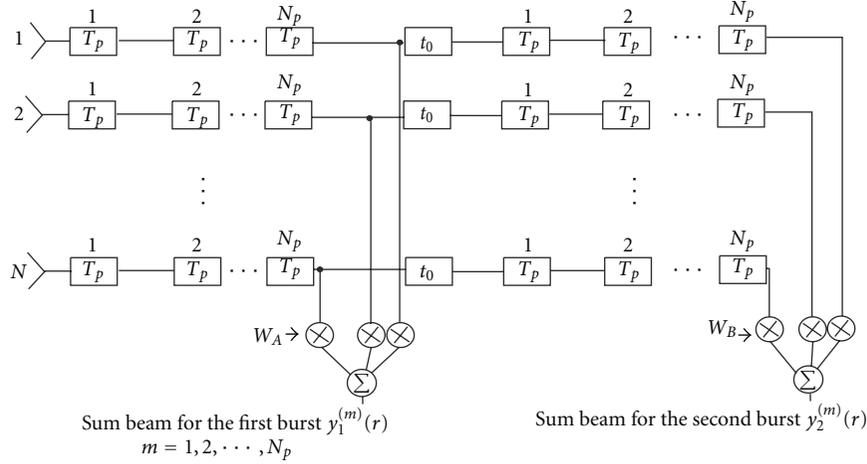
3.3. Space-Time Stacking. In this case the $(2N_p) \times 1$ space-time data snapshot is defined by stacking the data stream in (19) for all pulses as follows:

$$
\begin{aligned}
\mathbf{Y}_1(r) &= \left(y_1^{(1)}(r), y_2^{(1)}(r), y_1^{(2)}(r), \right. \\
&\quad\quad \left. y_2^{(2)}(r), \ldots, y_1^{(N_p)}(r), y_2^{(N_p)}(r) \right)^T \\
&= \sum_{k=1}^{N_s} \hat{\alpha}_{k,r} \mathbf{t}_s(f_{k,r}) \otimes \begin{pmatrix} 1 \\ F(\phi_k, \theta_r)\rho(t_0, f_{k,r}) \end{pmatrix} + \delta(r - r_t) \\
&\quad \times \alpha_t \mathbf{t}_s(f_t) \otimes \begin{pmatrix} 1 \\ \rho(t_0, f_t) \end{pmatrix} + \left(\mathbf{I}_{N_p} \otimes \mathbf{w}_n^H \right) \hat{\mathbf{e}}, \\
&\quad\quad r = N_1, \ldots, N_2,
\end{aligned}
\tag{20}
$$

where $\hat{\alpha}_{k,r} = \alpha_{k,r} \mathbf{w}_A^H \mathbf{s}_N(\phi_k, \theta_r)$ is the tapered clutter amplitude at the receiver level due to primary receiver and $\hat{\mathbf{e}}$ refers to the $2NN_p \times 1$ random component corresponding to all the pulses and channels. We may now define the space time steering manifold for dual Tx/Rx case as

$$
\hat{\mathbf{v}}(\phi, \theta, f, t_0) = \mathbf{t}_s(f) \otimes \begin{pmatrix} 1 \\ F(\phi, \theta)\rho(t_0, f) \end{pmatrix}.
\tag{21}
$$

3.4. Choice of Receiver Patterns. It can be shown that, if $\mathbf{w}_A$ and $\mathbf{w}_B$ are not carefully selected, several clutter arrivals may share the same spatial steering vector. In other words, $\mathbf{S}(\phi, \theta, f) = (1, F(\phi, \theta)\rho(t_0, f))^T$ has the property $\mathbf{S}(\phi_t, \theta_{r_t}, f_t) = \mathbf{S}(\phi_k, \theta_{r_t}, f_{k,r_t})$ for multiple k values, for most of the choices of the $\mathbf{w}_A$ and $\mathbf{w}_B$. This means that the look-direction constraint is satisfied by a number of sidelobe arrivals as well. The search Doppler bin is associated with the spatial steering vector $\mathbf{S}(\phi_t, \theta_{r_t}, f_t) = (1, \rho(t_0, f_t))^T$, where $f_t = (n-1)/(N_p T_p)$ $(n = 1, 2, \ldots, N_p)$ is the natural choice of Doppler bin values in the look direction. Suppose that the pattern ratio has the property $|F(\phi, \theta)| = 1$ for all angles; this allows us to represent $F(\phi, \theta)$ in the following form: $F(\phi, \theta) = e^{j2\pi\psi(\phi, \theta)}$, where $\psi(\phi, \theta)$ is the phase. If any of the clutter discretes has the same 2×1 spatial steering vector as the current search Doppler bin related spatial steering vector, then we have $(1, \rho(t_0, f_t))^T = (1, \rho(t_0, f_{k,r_t})e^{j2\pi\psi(\phi, \theta_{r_t})})^T$ for some value of k, the solution for which is given by solving $e^{j2\pi f_t(t_0 + N_p T_p)} = e^{j2\pi f_{k,r_t}(t_0 + N_p T_p)} \cdot e^{j2\pi\psi(\phi_k, \theta_{r_t})}$. This leads to the equation $2\pi f_t(t_0 + N_p T_p) = 2\pi f_{k,r_t}(t_0 + N_p T_p) + 2\pi\psi(\phi_k, \theta_{r_t}) \pm 2\pi m_0$, where m_0 is any arbitrary integer value. This is equivalent to solving $f_{k,r_t} = f_t - \psi_{k,r_t}\Delta_1 \pm m_0\Delta_1$, where $\Delta_1 = 1/(t_0 + N_p T_p) \approx 1/(N_p T_p)$ is the Doppler resolution and $\psi_{k,r_t} = \psi(\phi_k, \theta_{r_t})$. But in order to be a valid clutter discrete, we have the requirement $f_{k,r_t} = f_0 \sin\phi_k \cos\theta_{r_t}$ with

Sum beam for the first burst $y_1^{(m)}(r)$
$m = 1, 2, \cdots, N_p$

Sum beam for the second burst $y_2^{(m)}(r)$

FIGURE 2: N-Channel receiver configuration with two pulse bursts.

$f_0 = 2V_A/\lambda$. As a result we have

$$\sin \phi_k = \frac{f_t}{f_0 \cos \theta_{r_t}} - \left(\frac{\Delta_1}{f_0 \cos \theta_{r_t}} \right) \psi_{k,r_t} \pm \left(\frac{\Delta_1}{f_0 \cos \theta_{r_t}} \right) m_0, \tag{22}$$

and $|\psi_{k,r_t}| = |\psi(\phi_k, \theta_{r_t})| \leq 1$. The solution will provide multiple results for k making it impossible to satisfy the desired qualities to beamform. As an example, for equispaced linear array with half wavelength spacing, for the first data stream, we choose

$\mathbf{w}_A$

$$= \frac{\left(1, z(\phi_t, \theta_{r_t}), z(\phi_t, \theta_{r_t})^2, z(\phi_t, \theta_{r_t})^3, \ldots, z(\phi_t, \theta_{r_t})^{N-2}, 0\right)^T}{(N-1)}, \tag{23}$$

where the last element is switched off, and

$$\mathbf{w}_B = \frac{\left(0, 1, z(\phi_t, \theta_{r_t}), z(\phi_t, \theta_{r_t})^2, \ldots, z(\phi_t, \theta_{r_t})^{N-2}\right)^T}{(N-1)}, \tag{24}$$

for the second data stream with the first element switched off. Now we have $\mathbf{w}_B^H \mathbf{s}_N(\phi, \theta) = z(\phi, \theta) \mathbf{w}_A^H \mathbf{s}_N(\phi, \theta)$, where $z(\phi, \theta) = \exp(j\pi \sin \phi \cos \theta)$. We have the pattern ratio $F(\phi, \theta) = z(\phi, \theta)$, and substituting $\psi(\phi_k, \theta_{r_t}) = (\sin \phi_k \cos \theta_{r_t})/2$ leads to the following result:

$$\sin \phi_k = \left(\frac{f_t}{f_0 + \Delta_1/2} \pm \frac{m_0 \Delta}{f_0 + \Delta_1/2} \right) \frac{1}{\cos \theta_{r_t}}. \tag{25}$$

This will provide us a number of clutter discretes in general that satisfy the undesired properties mentioned above making it impossible to beamform in a spatial sense. The solution to resolve this situation is not to have a unit value for the absolute value of the pattern ratio for all angles except for the look-direction. A choice of a function $|F(\phi, \theta)|$

with the property $F(\phi_t, \theta_{r_t}) = 1$ and then smooth varying $|F(\phi, \theta)|$ across all other angles with property that no other angle provides the same output value for $|F(\phi, \theta)|$ as for the look direction that is generally $|F(\phi, \theta)| < 1$, with $F(\phi_t, \theta_{r_t}) = 1$, is an excellent choice as seen later. Since $|F(\phi, \theta)| = 1$ occurs only for the look direction, this will make 2×1 spatial vectors, $(1, \rho(t_0, f_t)F(\phi_t, \theta_{r_t}))^T$ and $(1, \rho(t_0, f_{k,r_t})F(\phi_k, \theta_{r_t}))^T$, linearly independent ($f_t \neq f_{k,r_t}$, $\phi_t \neq \phi_k$) for any k regardless of the phase component of the term $\rho(t_0, f_{k,r_t})F(\phi_k, \theta_{r_t})$. Furthermore, the search Doppler bin is associated with $(1, \rho(t_0, f_t))^T$ and if the phase component is ignored in the second entry of this vector we have $(1, 1)^T$, and this cannot be linearly dependent with any of the clutter discretes since all of them can be made to associate with the form $(1, |F|)^T$ with $|F| < 1$ except for the look-direction clutter that is, traditionally known as mainlobe clutter discrete which cannot be avoided in general in beamforming. The above property in 2×1 spatial manifold gives us sufficient conditions to carry out space-time beamforming. In order to further support that this argument, for a general MTR case, let us suppose we do three transmissions in the same direction, using 3 different receiver beam patterns $\mathbf{w}_A, \mathbf{w}_B$, and $\mathbf{w}_C$ pointed at the same look-direction, where each pulse train is N_p pulses long, and apply a common switching delay. Then we would have the 3×1 spatial component $(1, F(\phi, \theta)\rho(f, t_0), F_1(\phi, \theta)\rho(f, t_0)^2)^T$. In this case we will be enforcing the second pattern ratio to satisfy

$$F_1(\phi, \theta) = \frac{\mathbf{w}_C^H \mathbf{s}_N(\phi, \theta)}{\mathbf{w}_A^H \mathbf{s}_N(\phi, \theta)} = F(\phi, \theta)^2. \tag{26}$$

This will lead to the spatial component $(1, F(\phi, \theta)\rho(f, t_0), F(\phi, \theta)^2 \rho(f, t_0)^2)^T$ which follows a Vandermonde structure. When a "sinc" pattern is chosen for the first ratio $F(\phi, \theta)$ ($|F(\phi, \theta)| < 1$), with $F(\phi_t, \theta_{r_t}) = 1$, we are not able to express the look-direction-related spatial steering vector (i.e., $(1, 1, 1)^T$), as a sum of any two other spatial steering vectors which correspond to any two sidelobe-related clutter arrivals. Now, in space-time domain, we will satisfy the

requirement that the look-direction and Doppler-related $2N_p \times 1$ steering vector $\hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t)$ cannot be expressed as a linear combination of the clutter-related $df(= 2N_p - 1)$ steering vectors. The expected upper limit df would be the degrees of freedom. The most basic example of a pattern ratio is to choose what is known as "sinc" pattern. In general we can consider the case where we choose the Mth$(< N)$ order sinc function given by

$$F(\phi, \theta) = \left(\frac{1}{M}\right)\left(1 + zz_t^* + (zz_t^*)^2 + (zz_t^*)^3, \ldots, (zz_t^*)^{M-1}\right), \tag{27}$$

as the pattern ratio, where $z_t = z(\phi_t, \theta_{r_t}) = \exp(j\pi \sin\phi_t \cos\theta_{r_t})$ for a linear array with half wavelength spacing and $*$ denotes the complex conjugate. In order to achieve this result, we may choose the first receiver weights by

$$\mathbf{w}_A = \frac{(1, z_t, z_t^2, z_t^3, \ldots, z_t^{N_A-1}, 0, \ldots, 0)^T}{N_A}, \tag{28}$$

where $N_A = N - M + 1$. We can now estimate the desired weights for the second receiver by resolving the inverse problem

$$\mathbf{w}_B^H \mathbf{s}_N(\phi, \theta) = F(\phi, \theta) \cdot \mathbf{w}_A^H \mathbf{s}_N(\phi, \theta) = \left(\frac{1}{MN_A}\right)$$
$$\times \left(1 + (zz_t^*) + (zz_t^*)^2 + \cdots + (zz_t^*)^{M-1}\right)$$
$$\times \left(1 + (zz_t^*) + (zz_t^*)^2 + \cdots + (zz_t^*)^{N_A-1}\right)$$
$$= \left(c_0 + c_1 z + c_2 z^2 + \cdots, c_N z^{N-1}\right), \tag{29}$$

where c_p $(p = 0, 1, 2, \ldots, N)$ are (weights) easily obtainable by equating the coefficients of the above product which is of order N polynomial in z. These are the weights for the second receiver. Large value of M for the pattern ratio forces us to switch off too many elements at the first receiver.

4. Theoretical Performance Prediction

4.1. Comparison of Performances. For MTR-STAP, the interference only covariance matrix is expressed as a function of switching time using (20) by

$$\mathbf{R}_I(t_0) = \sum_{k=1}^{N_S} |\hat{\alpha}_{k,r_t}|^2 \hat{\mathbf{v}}(\phi_k, \theta_{r_t}, f_{k,r_t}, t_0) \hat{\mathbf{v}}(\phi_k, \theta_{r_t}, f_{k,r_t}, t_0)^H$$
$$+ \sigma_n^2 \left(\mathbf{I}_{N_p} \otimes \mathbf{w}_n^H\right)\left(\mathbf{I}_{N_p} \otimes \mathbf{w}_n^H\right)^H. \tag{30}$$

The optimal array weights are given by

$$\hat{\mathbf{w}} = \frac{\mathbf{R}_I^{-1}(t_0)\hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0)}{\left[\hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0)^H \mathbf{R}_I^{-1}(t_0)\hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0)\right]} \tag{31}$$

The output signal to clutter plus noise ratio is given by

$$SCN_o(\phi_t, \theta_{r_t}, f_t, t_0) = |\alpha_t|^2 \hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0)^H$$
$$\times \mathbf{R}_I^{-1}(t_0)\hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0). \tag{32}$$

In order to predict the performance of the MTR-STAP algorithm with the nonadaptive single A/D-based-FFT solution, as well as potential multichannel upgrades, we would like to establish a theoretical space-time clutter covariance matrix for each case using the parameters similar to MCARM system. Consider a 22-channel half wavelength equispaced airborne array with PRF = 1984 Hz, $\lambda = .24$ cm, $\nu = 100$ m/sec, $N = 22$, and $N_p = 64$ The estimation of the clutter covariance matrix was carried out using two methods. The continuous model described in [7] and another straightforward discrete method is to first determine a value for $N_s(\approx N_p)$ as the desired clutter degree of freedom. The discrete method considers a series of angles of arrivals to represent each Doppler bin of interest by using the equation (the ridge) $f_{k,r_t} = f_0 \sin\phi_k \cos\theta_{r_t} = (k - 1)/N_s T_p$. This equation provides us with a series of clutter angles for $\phi_k = \sin^{-1}((k - 1)/(f_0 N_s T_p \cos\theta_{r_t}))$ generally close to the figure N_s. This procedure creates nonuniform patches on the ground, and hence a series of power levels are associated with each patch, say σ_k^2 $(k = 1, 2, \ldots, N_s)$ which follow values proportionate to the patch size $(\phi_k - \phi_{k+1})r_t$. Finally the covariance matrix is estimated by summing $\sigma_k^2 \hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^H$ terms, where $\hat{\mathbf{v}}_k$ represents the appropriate manifold. In both approaches, we compute the rank of the covariance matrix to confirm the degrees of freedom.

Figure 3(a) illustrates the example of the two patterns selected for receiving with a predetermined pattern ratio of order $M = 4$ suitable for array broadside look. Figure 3(b) illustrates the case $(M = 12)$ where the scan angle is -40 degrees. The MTR system uses 22-channel $(= N)$, 64 pulse $(= N_p)$ system for each transmission. For comparison we consider the ideal scenario (Full STAP) where 64-pulses are transmitted and received via 22 digitised channels and full adaptive degrees of freedom (22×64) are applied to the processor by creating a 1408×1408 covariance matrix inversion. In practice this is not possible due to the lack of training data, so for greater realism we use reduced STAP with 128 adaptive degrees of freedom (i.e., $N_B = 2; N_p = 64$) for comparison of performances. The most important performance measure is described in (32). This curve, assuming a target of unit power, for the reduced STAP is represented by the symbol $\bullet-$ in Figure 4, and this is possibly the best curve achievable via any multiple A/D system. The results for the MTR are shown in the plots of Figure 4 for $M = 4$ with the symbol $--$ and in the same plot with the symbol $\circ$ for the case with $M = 12$ which corresponds to rapidly changing $|F(\phi, \theta)|$. In our examples total clutter-to-noise power ratio is 72 dB, and we have ignored the transmit tapering in order to increase the received clutter power levels to test the algorithms under severe clutter. These results simulate switching time $t_0 = T$. The performance of the conventional solution is denoted in Figure 4 as $*-$, which is the nonadaptive FFT-based solution currently available for single A/D phase array radar with the
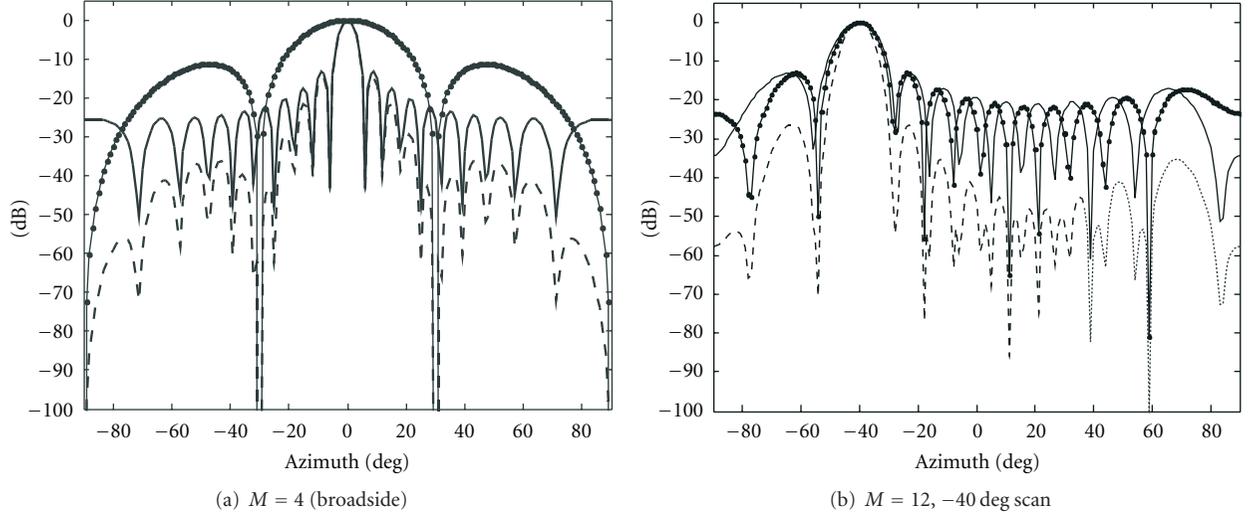
FIGURE 3: The two receiver patterns (azimuth only) and the associated pattern ratios. $-\mathbf{w}_A^H \mathbf{s}(\phi, \theta)$ pattern for the first transmission, - - $\mathbf{w}_B^H \mathbf{s}(\phi, \theta)$ pattern for the second transmission, and $\bullet - F(\phi, \theta))$ for the pattern ratio.
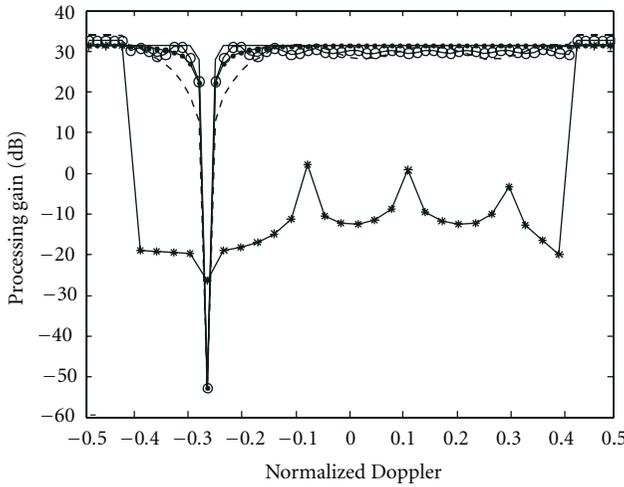

FIGURE 4: Performance comparison of processing options. - - Dual Tx/Rx single receiver system with 64 pulses in each Tx/Rx ($t_0 = T$; $M = 4$). $\circ-$ dual Tx/Rx single receiver system with 64 pulses in each Tx/Rx ($t_0 = T$, $M = 12$). $\bullet-$Reduced STAP with 64 pulses, 128 adaptive degrees of freedom ($N_B = 2$). $-$Ideal curve (standard STAP with 1408 degrees of freedom) with 64 pulses and 22 receiver channels. $\ast-$ FFT solution (available via single A/D).
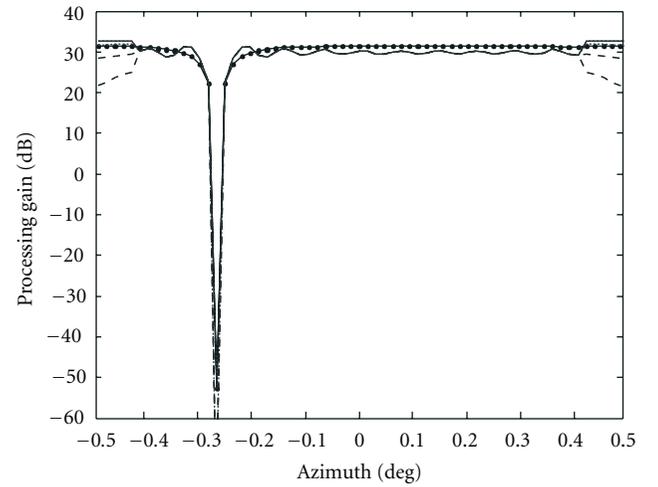

FIGURE 5: Sensitivity to switching time error ($t_0 = T_p/2$). $\bullet-$Reduced STAP with 128 adaptive degrees of freedom ($N_p = 64; N_B = 2$). $\bar{}\bar{}$MTR-STAP single receiver system with 64 pulses ($M = 12$) using $\tilde{t}_0 = 0, 0.2T_p, 0.8T_p$ (incorrect values), $-\tilde{t}_0 = 0.5T_p$ (true value).

same parameters. This FFT solution performs equally well only in 2 or 3 Doppler bins which are clutter free, that is, the far end of the spectrum. An important observation is that reduced STAP with 64 pulses and MTR with 64 pulses per transmission invert a matrix of size 128×128, but MTR can only handle no more than 64-degrees of freedom, beyond which it begins to fail. For clutter free Doppler bins, we can theoretically prove that MTR-STAP maintains a processing gain of NN_p.

4.2. Sensitivity to Switching Time Errors. A large number of simulations have confirmed that the filter performance is almost invariant to the selected value of t_0 for $0 < t_0 < T_p$. Next step is to estimate how accurately one has to know the value of t_0 to construct the arrays MTR steering vector. While representing $\hat{\mathbf{v}} = \hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, t_0)$ as the correct dual space-time steering manifold, we may now represent $\tilde{\mathbf{v}} = \hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, \tilde{t}_0)$ as the incorrect manifold, where $\tilde{t}_0$ represents the incorrectly chosen value of the switching time. Now the usual procedure is to find $\mathbf{w}$, for a given $\tilde{t}_0$ which is a guessed value for the switching time, which optimises the objective function $P = \mathbf{w}^H \mathbf{R}_I(t_0)\mathbf{w}$ subject to the constraint $\mathbf{w}^H \hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, \tilde{t}_0) = 1$ (written as $\mathbf{w}^H \tilde{\mathbf{v}} = 1$). It should be noted that the covariance matrix contains the correct value of the switching time (i.e., $\tilde{t}_0 = t_0$). The optimal solution ($\mathbf{w}$)

for any guess value of $\widetilde{t}_0$ is given by

$$\mathbf{w} = \frac{\mathbf{R}_I^{-1}(t_0)\widetilde{\mathbf{v}}}{[\widetilde{\mathbf{v}}^H \mathbf{R}_I^{-1}(t_0)\widetilde{\mathbf{v}}]}. \tag{33}$$

This leads to the output signal-to-interference ratio given by

$$SC\widetilde{N}_0\left(\phi_t, \theta_t, f_t, t_0, \widetilde{t}_0\right) = \frac{\left|\widetilde{\mathbf{v}}^H \mathbf{R}_I^{-1}(t_0)\widehat{\mathbf{v}}\right|^2}{\widetilde{\mathbf{v}}^H \mathbf{R}_I^{-1}(t_0)\widetilde{\mathbf{v}}}. \tag{34}$$

The Figure 5 shows the plots of the filter for various incorrect values of $\widetilde{t}_0$ by the symbol $--$ a, where $\widetilde{t}_0 = (0, 0.2T_p, 0.5T_p, 0.8T_p)$. The plot for the actual value is set at $t_0 = T/2$ shown with the symbol — ($M = 12, -40^0$ scan). When the correct value is assumed in setting up the steering manifold, we achieve the best performance where the curve is horizontal, and achieve the value $10\log_{10}(NN_p) = 10\log_{10}(22 \times 64) \approx 34.5$ for the detection in noise which corresponds to several Doppler bins at the two ends. For the bin with severe clutter or look-direction clutter, the performance is downgraded severely; that is, the depth of the clutter notch at the mainlobe clutter Doppler value is very deep.

4.3. Optimisation with respect to Switching Time.

As we have seen, the knowledge of switching time is important in clutter-free Doppler bins, and in other areas it does not degrade the performance considerably except at the mainlobe. However, it would be possible to optimise the desired output at the beamformer with respect to the space-time weights vector as well as switching time. The final expression for the signal-to-interference ratio in (34) contains the term

$$d\left(\widetilde{t}_0\right) = \widetilde{\mathbf{v}}^H \mathbf{R}_I^{-1}\widetilde{\mathbf{v}} = \widehat{\mathbf{v}}\left(\phi_t, \theta_{r_t}, f_t, \widetilde{t}_0\right)^H \mathbf{R}_I^{-1}(t_0)\widehat{\mathbf{v}}\left(\phi_t, \theta_{r_t}, f_t, \widetilde{t}_0\right), \tag{35}$$

below the line which we would like to optimise with respect to $\widetilde{t}_0$, in order to further improve the final processing gain. This leads to the following result:

$$\widetilde{t}_0 = \frac{1}{j2\pi \sin\phi_t \cos\theta_{r_t}}$$
$$\times \log_e\left(\frac{-(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b)^H \mathbf{R}_I^{-1}(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a)}{\left|(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b)^H \mathbf{R}_I^{-1}(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a)\right|}\right), \tag{36}$$

where

$$\mathbf{v}_a = \left(\mathbf{w}_A^H \mathbf{s}_N(\phi_t, \theta_{r_t}), 0\right)^T,$$
$$\mathbf{v}_b = \left(0, \mathbf{w}_B^H \mathbf{s}_N(\phi_t, \theta_{r_t}) \exp\left(j2\pi f_t N_p T_p\right)\right)^T, \tag{37}$$

and $||$ refers to the absolute value of a complex number (see the appendix for the proof). Simulation study has shown that the formula in (36) always produces a 99.9% accurate estimate of the switching time for all look directions which excludes broadside. This result is tested using MCARM data.

5. Analysis of MCARM Data

5.1. Selection of Pattern Ratios.

The US Air Force Research Laboratory, Rome Research Site collected a large amount of multichannel airborne radar measurement (MCARM) data [5]. The size of the MCARM array's calibrated matrix $\mathbf{s}_{i,j}$ ($i = 1, 2, \ldots, 22$, $j = 1, 2, \ldots, 129$) is 22×129, where 129 is the number of possible beamforming angles available in azimuth. Other important MCARM parameters are as follows: transmit frequency = 1240 MHz, the number of coherent pulses = 128, pulse repetition frequency = 1984 Hz ($T_p = 5.0403 \times 10^{-4}$ sec.), and number of cells = 680 (0.8 μsec pulses).

In order to generate data to suit the MTR scenario we combine the first 63 pulses with array weights vector $\mathbf{w}_A$, and the pulse numbers 65,..., 123 are combined with the weights vector $\mathbf{w}_B$. The 64th pulse is ignored allowing a switching time. This will simulate a delay (t_0) equal to one PRI. We use $M = 6$ to determine the weight vectors $\mathbf{w}_A$ and $\mathbf{w}_B$ as follows. As an example for the broadside look ($j = 65$) $\mathbf{w}_A(i) = (1/c_a)\mathbf{s}_{i,65}$, $i = 1, 2, \ldots, 17$, and $\mathbf{w}_A(i) = 0$, $i = 18, 19, \ldots, 22$, where c_a is the normalisation constant given by $c_a = (\mathbf{s}_{1,65}, \mathbf{s}_{2,65}, \ldots, \mathbf{s}_{17,65})^H(\mathbf{s}_{1,65}, \mathbf{s}_{2,65}, \ldots, \mathbf{s}_{17,65})$. This will make the last 5 elements inactive at the first receiver. Now we have to determine the pattern ratio, before estimating the second receiver weights. We define this by $\mathbf{w}_F(i) = 0, i = 1, 2, \ldots, 16$ and $\mathbf{w}_F(i) = (1/c_f)\mathbf{s}_{i,65}$, $i = 17, 18, 19, \ldots, 22$, where

$$c_f = (\mathbf{s}_{17,65}, \mathbf{s}_{18,65}, \ldots, \mathbf{s}_{22,65})^H(\mathbf{s}_{17,65}, \mathbf{s}_{18,65}, \ldots, \mathbf{s}_{22,65}). \tag{38}$$

In fact this gives us the pattern related to the last 6 elements of the array which would generally follow a uniform pattern to be the pattern ratio. Finally, we create the pattern $\mathbf{w}_B$ buy convolving the two patterns to obtain $\mathbf{w}_B$ which is a set of 22 complex numbers corresponding to the polynomial product. Finally $\mathbf{w}_B$ has to be normalised using the constant $c_b = (\mathbf{s}_{1,65}, \mathbf{s}_{2,65}, \ldots, \mathbf{s}_{22,65})^H(\mathbf{s}_{1,65}, \mathbf{s}_{2,65}, \ldots, \mathbf{s}_{22,65})$. This way we use all elements for the second receiver which receives pulse numbers from 65 to 123. This gives remarkable results as seen in Figure 6(a) for the angle index 65 (broadside) and in Figure 6(b) for the angle index 106.

As an example, the angle Doppler map of the dataset numbers rd50153 and rd50575 is shown in Figures 7(a) and 7(b), respectively. These plots are algorithm independent, and we simply apply the Fourier Transform on mutichannel data for all 129 beams to Doppler domain. However, one may use the new algorithm to produce the same plot with less resolution (3 dB) when multichannel data is not available. One important point to notice regarding all the MCARM datasets is the fact that the clutter center has shifted from the zero Doppler value. In other words, the Doppler value corresponding to the array broadside (with index = 65), we have nonzero Doppler value as clearly seen in Figures 7(a) and 7(b). Generally, this will not degrade the STAP performance. What this means for MCARM data sets is that we have the clutter ridge given by the format $f = f_s + f_0 \sin(\phi)$, where f_s is the clutter center shift. As long as we impose the above formula for the clutter ridge in optimising
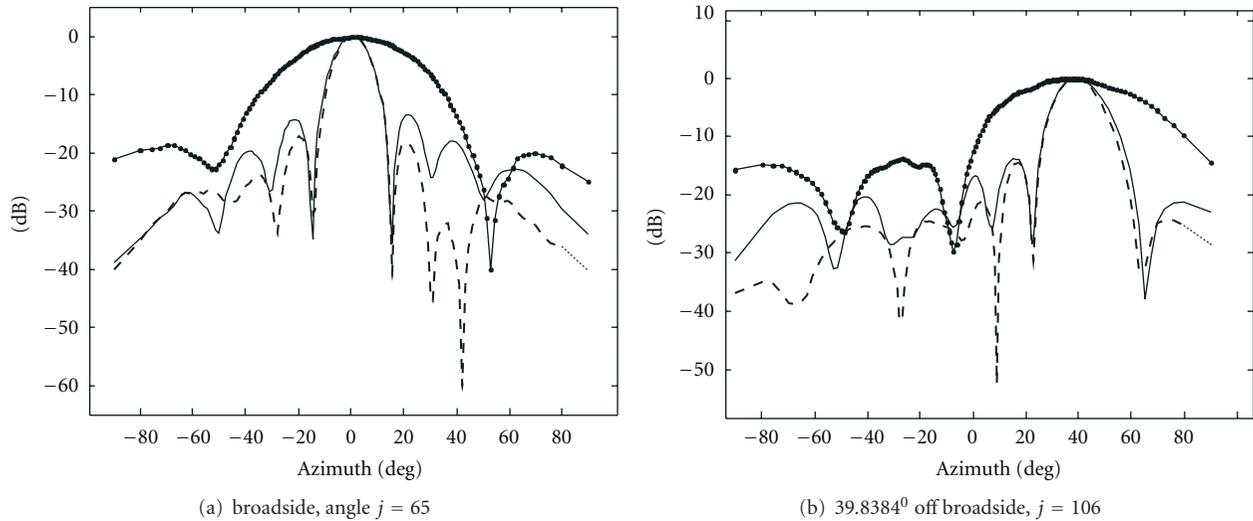
(a) broadside, angle $j = 65$

(b) 39.8384^0 off broadside, $j = 106$

FIGURE 6: The two receiver patterns and the pattern ratios $M = 6$ ($-\mathrm{w}_A^H \mathrm{s}(\phi)$, $--\mathrm{w}_B^H \mathrm{s}(\phi)$, and $\bullet - F(\phi)$).

(35), we can estimate the switching time as well as the clutter shift without having any knowledge of f_0 the value of which is in fact = 827.8619 Hz, and this knowledge is not needed to estimate f_s as seen below.

5.2. Switching Time Estimation. Let us assume that the switching time is unknown and we would like to estimate its value using the data set constructed for the MTR scenario. We can apply the result in (35) to estimate t_0 by using the formula $f = f_s + f_0 \sin(\phi)$. For the angle index 65 (broadside look, $\sin(\phi) = 0$), we have $f = f_s$ which is an unknown quantity. Therefore, it is only possible to estimate the value of $\tilde{t}_0$, for any guessed value of f_s and then evaluate the value of the objective function $d(\tilde{t}_0)$ in (35) which would optimise the processing gain. For some value of f_s, we may find that the objective function is absolutely optimum or the processing gain maximum, at which point we have the best pair of ($f_s, \tilde{t}_0$). Such a plot is illustrated in Figure 8 for several data sets. The data sets rd15015x ($x = 2, 3, 4, 5$) all have very similar curves. The data set rd150575 has a very different clutter center (-108 Hz), whereas only two data sets (rd150150 and rd150151) have almost zero as the clutter centre to within 1 Hz accuracy. For the data set rd50151 we encountered a singularity due to the fact that the clutter center is zero. In this case one should steer the beam to the next position (angle index = 66), which is 0.9 degrees off the array broadside. The estimated missing pulse length (T_p) is reasonably well estimated as illustrated in Table 1.

5.3. Signal Processing Gain. In order to compare (Figures 9(a) and 9(b)) with the multichannel (22 A/D solution), we use the reduced STAP using the channels 1, 2,..., 21 to form one channel (using uniform weights) and then channels 2, 3,..., 22 to form the second stream of data. The two data streams are combined to form the covariance matrix of size 128×128 using the first 64 pulses only (pulses 65 to 128 are discarded). This would make it the same size covariance

TABLE 1: Clutter center estimate for several MCARM data sets and the corresponding optimal switching time estimates.

Data set number	f_s (estimated clutter center (Hz))	$\tilde{t}_0/T_p$ (estimated switching time (seconds))
Rd50150	0	0.8350
Rd50151	0	0.9890
Rd50152	310	0.9337
Rd50153	341	1.0735
Rd50154	325	0.9627
Rd50155	372	1.1052
Rd50575	−108	0.8404

matrix (128×128) which we used in MTR demonstrations. This would apply 128 adaptive degrees of freedom, which seems to be twice the MTR solution is capable of. Applying any more adaptive degrees of freedom to multiple A/D-based solution will not give us a fair comparison.

6. Concluding Remarks

The most important observation is that the MTR inverts a matrix of size $2N_p \times 2N_p$, but it does not mean it's adaptive degrees of freedom is $2N_p$. The simulation has confirmed that it is limited to N_p. At this stage this can only be verified using extensive simulation. Another observation based on simulation data as well as MCARM data is that the order of pattern ratio is best to be around half the total number of sensors in the array. In our theoretical simulation, even though we use 128×128 matrix inversions for both MTR and beamspace solutions, we always validated this using covariance matrix of rank ≈ 60 via both continuous and discrete clutter models. As soon as the rank of the covariance

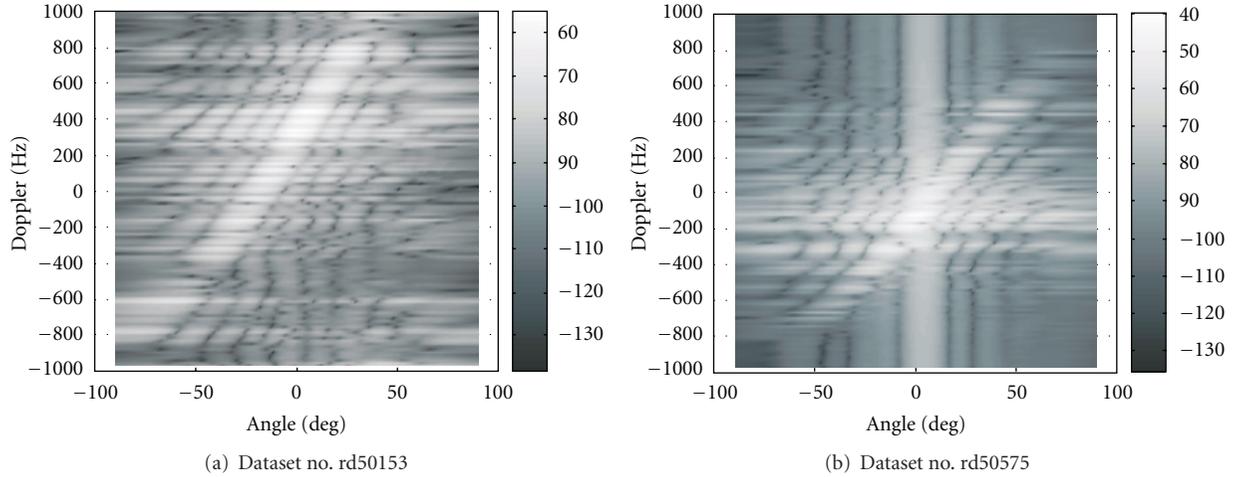(a) Dataset no. rd50153
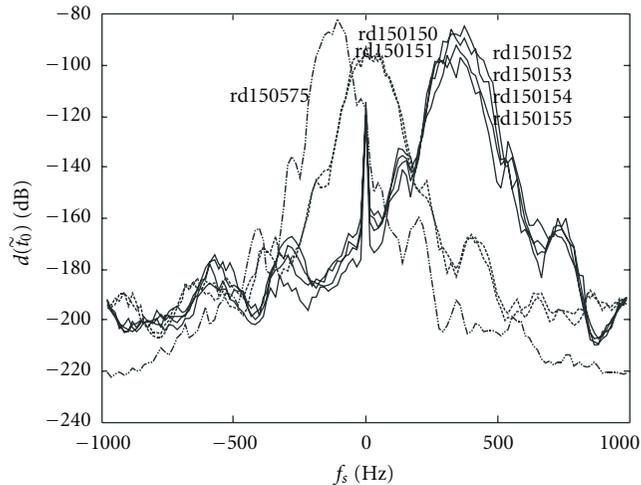

(b) Dataset no. rd50575

FIGURE 7: Angle Doppler Map of the MCARM dataset no. rd50575 (range cell number 200 is chosen).

FIGURE 8: Plot of the optimal value of $d(\widetilde{t}_0)$ for every guessed value of f_s.

matrix increases beyond 64, the MTR with 128×128 matrix solution begins to fail, and one has to increase the length of the pulse train accordingly. This also explains why MTR processing gain is marginally inferior when it comes to MCARM data. The reduced STAP solution is able to apply 128 adaptive degrees of freedom, while the MTR is able to apply up to 64, with the same size matrix inversion. It is also important to notice the nonzero clutter centers where the clutter notch occurs in Figures 9(a) and 9(b). The solution presented in this study is much more robust than the multichannel multiple A/D solution when no jammers are encountered. This procedure can avoid all the complications involved in synchronising a number of A/D converters to achieve good results. This is not really a new STAP algorithm; rather, it provides a way to apply many standard STAP algorithms by constructing multichannel data out of a single A/D converter.

Furthermore, it also makes it much easier to calibrate the array with only a single A/D. The simulation study has shown that the optimal configuration would be to make M equal to around half the number of sensors in the array. The major drawback in the software approach is that we need twice as many pulses to maintain the same performance or else a 3 dB loss occurs in the Doppler resolution. It is also possible to extend the algorithm to null sidelobe jammers as well. This analysis is beyond the scope of this paper.

Appendix

It should be noted that $\mathbf{w}_A^H \mathbf{s}_N(\phi_t, \theta_{r_t}) = 1$, $\mathbf{w}_B^H \mathbf{s}_N(\phi_t, \theta_{r_t}) = 1$ are physical requirements and not relevant to the analysis to follow. Even though we have ignored the numerator in (34), we can achieve the desired result (i.e., $\widetilde{t}_0 \approx t_0$) very accurately by only optimising the denominator term. Further evidence is given later to prove this result. In order to further optimise (i.e., to minimise)

$$d\left(\widetilde{t}_0\right) = \widetilde{\mathbf{v}}^H \mathbf{R}_I^{-1} \widetilde{\mathbf{v}} = \widehat{\mathbf{v}}\left(\phi_t, \theta_{r_t}, f_t, \widetilde{t}_0\right)^H \mathbf{R}_I^{-1} \widehat{\mathbf{v}}\left(\phi_t, \theta_{r_t}, f_t, \widetilde{t}_0\right)$$
(A.1)

with respect to $\widetilde{t}_0$, we define $\rho(\widetilde{t}_0, f_t) = \beta \exp(j2\pi f_t N_p T_p))$, where $\beta = \gamma \exp(j2\pi f_t \widetilde{t}_0)$ is a new parameter that contains $\widetilde{t}_0$ and the parameter γ (amplitude of β) which is introduced in order to formulate the optimisation problem to suit the Lagrange multiplier method, where we will be forcing the requirement that $\gamma = 1$ (or equivalently $\beta\beta^* = 1$) at the optimisation. We may now write the MTR space-time steering vector with guessed switching time as

$$\widehat{\mathbf{v}}\left(\phi_t, \theta_{r_t}, f_t, \widetilde{t}_0\right)$$
$$= \left(\mathbf{t}_s(f_t) \otimes \left[\mathbf{w}_A^H \mathbf{s}_N(\phi_t, \theta_{r_t}), \beta \exp\left(j2\pi f_t N_p T_p\right)\right.\right.$$
(A.2)
$$\left.\left. \times \mathbf{w}_B^H \mathbf{s}_N(\phi_t, \theta_{r_t})\right]^T\right),$$
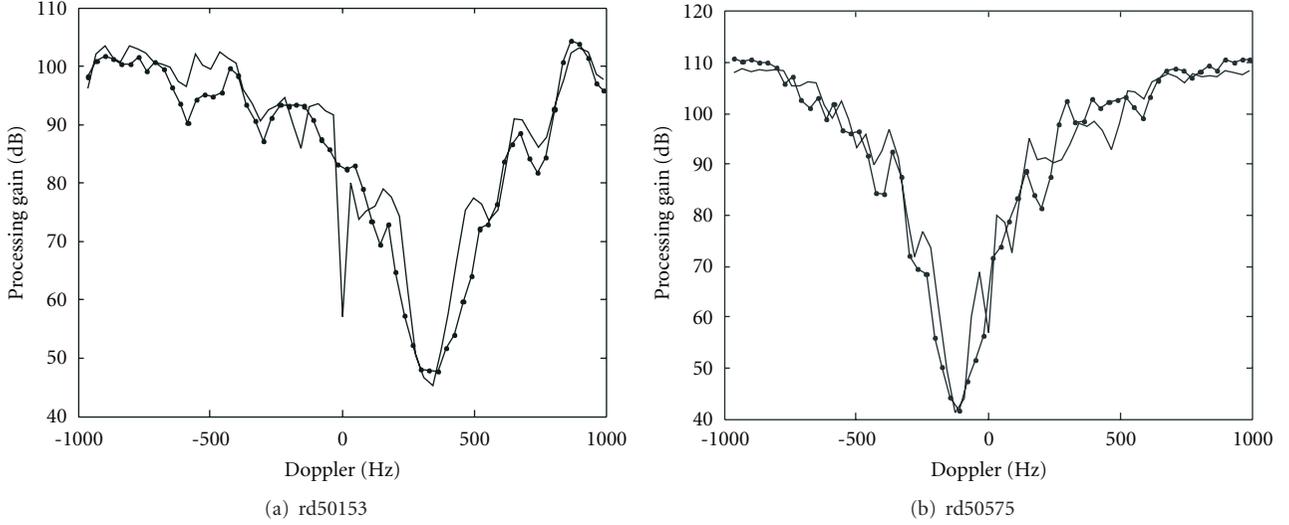
(a) rd50153

(b) rd50575

FIGURE 9: Processing gain across all Doppler bins when scanned to array broadside (angle index = 65). $\bullet-$ MTR solution, — Reduced STAP.

and express $d(\widetilde{t_0})$ (a function of β and β^*) as follows:

$$
\begin{aligned}
d(\beta,\beta^*) &= \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a + \beta \mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \mathbf{R}_I^{-1} \\
&\quad \times \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a + \beta \mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right) \\
&= \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right)^H \mathbf{R}_I^{-1} \mathbf{t}_s(f_t) \otimes \mathbf{v}_a \\
&\quad + \beta \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right)^H \mathbf{R}_I^{-1} \mathbf{t}_s(f_t) \otimes \mathbf{v}_b \\
&\quad + \beta^* \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \mathbf{R}_I^{-1} \mathbf{t}_s(f_t) \otimes \mathbf{v}_a \\
&\quad + \beta\beta^* \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \mathbf{R}_I^{-1} \mathbf{t}_s(f_t) \otimes \mathbf{v}_b.
\end{aligned}
\tag{A.3}
$$

Now we would like to optimise $d(\beta,\beta^*)$ subject to the constraint $\beta\beta^* = 1$. Using the Lagrange Multiplier method, we need to optimise the objective function $\psi(\beta,\beta^*) = d(\beta,\beta^*) + \eta(\beta\beta^* - 1)$ by solving $\partial\psi/\partial\beta^* = 0$, where η is an unknown parameter (or alternatively $\partial\psi/\partial\beta = 0$). As a result we have $\beta = -A/(\eta + C)$, where

$$
\begin{aligned}
A &= \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \mathbf{R}_I^{-1}\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right), \\
C &= \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \mathbf{R}_I^{-1}\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right).
\end{aligned}
\tag{A.4}
$$

The unknown parameter (real) is estimated by enforcing $\beta\beta^* = 1$. That is, $|A|^2 = |C + \eta|^2 = |C|^2 + \eta^2 + \eta(C + C^*)$ or $\eta^2 + \eta(C + C^*) + |C|^2 - |A|^2 = 0$.

The solution is given by

$$
\eta = -\frac{(C + C^*)}{2} \pm \sqrt{\frac{(C + C^*)^2}{4} - \left(|C|^2 - |A|^2\right)}.
\tag{A.5}
$$

Furthermore, we have $C^H = C(= C^*)$, and noting that covariance matrix is positive definite (i.e., $C \geq 0$), replacing $(C + C^*)/2 = C$, and taking the positive sign for a physically realisable result (i.e., $\eta \geq 0$), we can conclude that $\eta + C = \sqrt{|A|^2} = |A|$ and hence the desired result in (36). The above optimisation was carried out for a given value of f_t usually expressed by $f_t = (n - 1)/N_p T$, $n = 1, 2, \ldots, N_p$. Rather than estimating $\widetilde{t_0}$ for every Doppler bin, for the best results, one may choose the value corresponding to the bin given by $f_t = f_0 \sin\phi_t \cos\theta_{r_t}$ which is a point on the clutter ridge. As a result, for the secondary optimisation we may consider the Doppler bin which contains the most clutter by assigning the value $f_t = f_0 \sin\phi_t \cos\theta_{r_t}$. Now along the clutter ridge, the look direction or in other words, the processing gain related to mainlobe clutter power at the output needs to be maximised with respect to $\widetilde{t_0}$. However, in case of the broadside look where $\phi_t = 0, f_t = 0$, such optimisation is not relevant and cannot be implemented since $\widehat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_t, \widetilde{t_0})$ is independent of switching time. Further evidence can be justified as follows.

Since we are restricting our optimisation to the look direction's mainlobe clutter-related Doppler bin only where the clutter power is severe at the receiver, the processing gain which needs to be maximised can be expressed as (from (34))

$$
SC\widetilde{N}\left(\phi_t, \theta_{r_t}, f_0 \sin\phi_t \cos\theta_{r_t}, t_0, \widetilde{t_0}\right) = \frac{\left|\widetilde{\mathbf{v}}^H \widehat{\mathbf{R}}_I^{-1}(t_0)\widehat{\mathbf{v}}_c\right|^2}{\widetilde{\mathbf{v}}^H \widehat{\mathbf{R}}_I^{-1}(t_0)\widetilde{\mathbf{v}}},
\tag{A.6}
$$

where $\widehat{\mathbf{v}}_c = \widehat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_0 \sin\phi_t \cos\theta_{r_t}, t_0)$.

The objective here is to prove that the denominator of the above expression is minimised for $\widetilde{t_0} = t_0$, and at the same time the numerator is maximized; hence, the processing gain achieves a possible overall maximum value.

Let us represent the $2N_p \times 2N_p$ space-time covariance matrix by the dominant clutter discrete (mainlobe clutter) and write

$$\mathbf{R}_I(t_0) = \left(\sigma_c^2 \hat{\mathbf{v}}_c \hat{\mathbf{v}}_c^H + \sigma_n^2 \mathbf{I}_{2N_p} \right), \qquad (A.7)$$

where σ_c represents the power level. Now we may use the following well-known matrix inversion lemma.

Lemma 1. *Suppose that the square matrix $\mathbf{Q}$ is added to an additional dyad term $\mathbf{u}\mathbf{u}^H$, where $\mathbf{u}$ is a column vector; then the inversion of the new matrix is given by (See in p1348 [8])*

$$\left(\mathbf{Q} + \mathbf{u}\mathbf{u}^H \right)^{-1} = \mathbf{Q}^{-1} - \frac{\mathbf{Q}^{-1}\mathbf{u}\mathbf{u}^H\mathbf{Q}^{-1}}{1 + \mathbf{u}^H\mathbf{Q}^{-1}\mathbf{u}}. \qquad (A.8)$$

Substituting $\mathbf{Q} = \sigma_n^2 \mathbf{I}_{2N_p}$ and $\mathbf{u} = \sigma_c \hat{\mathbf{v}}_c$, the inversion of the covariance matrix in (A.7) is given by

$$\mathbf{R}_I^{-1}(t_0) = \frac{1}{\sigma_n^2}\left(\mathbf{I}_{2N_p} - \frac{\sigma_c^2 \hat{\mathbf{v}}_c \hat{\mathbf{v}}_c^H}{\sigma_n^2 + 2N_p\sigma_c^2} \right), \qquad (A.9)$$

and we have

$$\tilde{d}(t_0) = \tilde{\mathbf{v}}^H \hat{\mathbf{R}}_I(t_0)^{-1}\tilde{\mathbf{v}} = \frac{1}{\sigma_n^2}\left(2N_p - \frac{\sigma_c^2 \left| \tilde{\mathbf{v}}^H \hat{\mathbf{v}}_c \right|^2}{\sigma_n^2 + 2N_p\sigma_c^2} \right), \qquad (A.10)$$

where $\tilde{\mathbf{v}} = \hat{\mathbf{v}}(\phi_t, \theta_{r_t}, f_0 \sin\phi_t \cos\theta_{r_t}, \tilde{t}_0)$ contains the incorrect switching time. Now by expressing $\hat{\mathbf{v}}_c$ and $\tilde{\mathbf{v}}$ as

$$\hat{\mathbf{v}}_c = \mathbf{t}_s(f_t) \otimes [\mathbf{v}_a + \beta_0 \mathbf{v}_b],$$
$$\tilde{\mathbf{v}} = \mathbf{t}_s(f_t) \otimes [\mathbf{v}_a + \beta \mathbf{v}_b], \qquad (A.11)$$

where $\beta_0 = \exp(j2\pi f_t t_0)$, $\beta = \exp(j2\pi f_t \tilde{t}_0)$, and $f_t = f_0 \sin\phi_t \cos\theta_{r_t}$.

$$\mathbf{v}_a = \left(\mathbf{w}_A^H \mathbf{s}_N(\phi_t, \theta_{r_t}), 0 \right)^T = (1, 0)^T,$$

$$\mathbf{v}_b = \left(0, \mathbf{w}_B^H \mathbf{s}_N(\phi_t, \theta_{r_t}) \exp\left(j2\pi f_t N_p T \right) \right)^T \qquad (A.12)$$

$$= \left(0, \exp\left(j2\pi f_t N_p T \right) \right)^T.$$

The value of A, as defined above in (A.4), is given by

$$A = \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b \right)^H \mathbf{R}_I^{-1}\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a \right)$$

$$= \frac{1}{\sigma_n^2}\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b \right)^H$$

$$\times \left(\mathbf{I}_{2N_p} - \frac{\sigma_c^2 \left[\mathbf{t}_s(f_t) \otimes (\mathbf{v}_a + \beta_0 \mathbf{v}_b) \right]\left[\mathbf{t}_s(f_t) \otimes (\mathbf{v}_a + \beta_0 \mathbf{v}_b) \right]^H}{\sigma_n^2 + 2N_p\sigma_c^2} \right)$$

$$\times \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a \right). \qquad (A.13)$$

Now expanding the above expression and noting that $\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right) = 0$ and $\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right)^H \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right) = 0$,

$\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right) = N_p$ and $\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right)^H \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right) = N_p$, we have the result

$$A = \frac{-\sigma_c^2 \beta_0 N_p^2}{\sigma_n^2 \left(\sigma_n^2 + 2N_p\sigma_c^2 \right)}. \qquad (A.14)$$

This will lead us to the value $A/|A| = \beta_0$, and hence we have $\tilde{t}_0 = t_0$ as the estimate.

Now the numerator in (A.6) can be expressed as (noting that $\hat{\mathbf{v}}_c^H \hat{\mathbf{v}}_c = 2N_p$)

$$\left| \tilde{\mathbf{v}}^H \hat{\mathbf{R}}_I^{-1} \hat{\mathbf{v}}_c \right|^2 = \left| \frac{1}{\sigma_n^2} \tilde{\mathbf{v}}^H \left(\mathbf{I}_{2N_p} - \frac{\sigma_c^2 \hat{\mathbf{v}}_c \hat{\mathbf{v}}_c^H}{\sigma_n^2 + 2N_p\sigma_c^2} \right) \hat{\mathbf{v}}_c \right|^2$$

$$= \frac{1}{\sigma_n^4} \left| \tilde{\mathbf{v}}^H \hat{\mathbf{v}}_c \right|^2 \left| \left(1 - \frac{\sigma_c^2 \hat{\mathbf{v}}_c^H \hat{\mathbf{v}}_c}{\sigma_n^2 + 2N_p\sigma_c^2} \right) \right|^2$$

$$= \frac{\left| \tilde{\mathbf{v}}^H \hat{\mathbf{v}}_c \right|^2}{\left(\sigma_n^2 + 2N_p\sigma_c^2 \right)^2}. \qquad (A.15)$$

Substituting (A.11) in the above expression using the properties of those vectors used in the previous case (i.e., $\left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_b\right)^H \left(\mathbf{t}_s(f_t) \otimes \mathbf{v}_a\right) = 0$, etc.), we arrive at

$$\left| \tilde{\mathbf{v}}^H \hat{\mathbf{R}}_I^{-1} \hat{\mathbf{v}}_c \right|^2 = \frac{N_p^2 \left| 1 + \exp\left(j2\pi f_t\left(\tilde{t}_0 - t_0 \right) \right) \right|^2}{\left(\sigma_n^2 + 2N_p\sigma_c^2 \right)^2}, \qquad (A.16)$$

which produces its maximum value at $\tilde{t}_0 = t_0$.

Acknowledgments

References

[1] R. Klem, *Space-Time Adaptive Processing*, The Institution of Electrical Engineers, London, UK, 1999.

[2] R. Klem, *Applications of Space-Time Adaptive Processing*, The Institution of Electrical Engineers, London, UK, 2004.

[3] D. Madurasinghe and P. E. Berry, "Pre-Doppler direct data domain approach to STAP," *Signal Processing*, vol. 85, no. 10, pp. 1907–1920, 2005.

[4] R. D. Brown, R. A. Schneible, M. C. Wicks, H. Wang, and Y. Zhang, "STAP for clutter suppression with sum and difference beams," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 634–646, 2000.

[5] B. N. S. Babu, J. A. Torres, and W. L. Melvin, "Processing and evaluation of multichannel airborne radar measurements (MCARM) measured data," in *Proceedings of IEEE International Symposium on Phased Array Systems and Technology*, pp. 395–399, October 1996.

[6] L. E. Brennan and L. S. Reed, "Theory of adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 9, no. 2, pp. 237–252, 1973.

[7] S. T. Smith, "Space-time clutter covariance matrix computation and interference subspace tracking," in *Proceedings of the 29th Asilomer Conference Signals, Systems and Computers*, pp. 1193–1197, 1995.

[8] H. L. Van Trees, *Optimum Array Processing*, Detection, Estimation and Modulation Theory, Part IV, John Wiley & Sons, New York, NY, USA, 2002.