Constrained Fast-time STAP Techniques for Interference Suppression in Multichannel SAR

Luke Rosenberg
University of Adelaide, Australia and
Defence Science and Technology Organisation
Email: luke.rosenberg@dsto.defence.gov.au

Doug Gray
University of Adelaide, Australia
Email: dgray@eleceng.adelaide.edu.au

Abstract—Forming a Synthetic Aperture Radar (SAR) image while suppressing an airborne broadband jammer can potentially destroy large regions of the image. In addition to this, multipath reflections from the ground, known as hot-clutter (HC) or terrain scattered interference will add a non-stationary interference component to the image. Using multiple antennas on a SAR provides spatial degrees of freedom and allows for adaptive beamforming to suppress the jammer signals. This paper presents a summary of constrained sub-optimal fast-time Space Time Adaptive Processing (STAP) techniques which reduce the interference level with minimal distortion to the SAR image.

I. INTRODUCTION

The goal of interference suppression for SAR is to successfully suppress the undesired signals while not significantly effecting the image quality by blurring, reducing the resolution or raising the sidelobe level. This can be hard to achieve in practice, especially if the interference is non-stationary and the training statistics change from pulse to pulse, causing traditional slow-time STAP techniques to be ineffective, [1]. Therefore adapting within each pulse is required by exploiting fast-time STAP. This offers the advantage of exploiting the coherency between the direct-path jammer and other HC signals to provide improved interference rejection. It will however cause secondary modulations during image formation, similar to that shown by [2]. In previous work, the use of derivative constraints to reduce potential signal suppression has shown to be an effective compromise to reduce the interference without compromising the target’s range profile, [3].

The first algorithm presented in this paper is a sub-optimal constrained Generalised Sidelobe Canceller (GSC) suitable for reducing the secondary modulations without damaging the desired image. Fully adaptive processing however can be very computationally intensive and not suitable for real time operation. Moreover, if the interference is non-stationary, the eigenvalues of the covariance matrix will spread, increasing the interference rank and therefore the degrees of freedom required to effectively cancel it. This problem is also analogous to the Moving Target Indication application where the ground clutter returns may not be stationary due to real world effects, such as aircraft crabbing, non-linear array geometry, intrinsic clutter motion, and scattering from near-field obstacles. Two modified GSC algorithms are therefore presented using techniques to reduce the rank associated with the interference plus noise covariance matrix. These approaches provide equivalent performance to the full rank version but with reduced sample support and computation.

II. SYSTEM MODELS AND GEOMETRY

A. SAR Signal Model

Consider a SAR platform travelling along the y-axis at \( v_p \) m/s, imaging a point in the slant-plane \( x \in [X_c - X_0, X_c + X_0] \), \( y \in [-Y_0, Y_0] \). The radar transmits a broadband chirp and the received signal \( x_n(t, u) \), is base-hand and sampled for each of the \( N \) channels of a linear antenna array with equi-spaced receivers along the azimuth direction. The variables \((t, u)\) represent (fast-time) samples within a pulse and the SAR platform position (slow-time) respectively. As the SAR bandwidth, \( B \) (Hz) is much smaller than the carrier frequency, \( \omega_c \) (rad/s), the SAR signal model can be split into temporal and spatial components.

The spatial reference signal is given by the time difference within a given area. The interference plus noise signal, \( z_n(t, u) \), or equivalently, an angular offset \( \theta(u) \), i.e.,

\[
s_n(u) = \exp\left[ j \frac{\omega_c}{c} d_n \sin(\theta(u)) \right]
\]

where \( c \) is the speed of light, \( \theta(u) = \arctan(-u/X_c) \) is the steering angle and \( d_n = n\delta \) is the antenna offset from the array phase centre with antenna spacing \( \delta \) and \( n \in \{- (N-1)/2, (N-1)/2\} \) for \( N \) (odd) antenna elements.

The received SAR signal comprises the total ground return, \( \gamma_n(\cdot) \), interference from the direct-path and ground reflected path, \( q_n(\cdot) \) and receiver noise \( \nu_n(t, u) \).

\[
x_n(t, u) = \gamma_n(t, u) + q_n(t, u) + \nu_n(t, u)
\]

Fig. 1 shows the processing chain from transmission of the chirp signal, formation of the received SAR signal, range processing, adaption and image formation.

B. Jammer and Noise Models

The bistatic jammer model assumes there are \( K \) HC patches within a given area. The interference plus noise signal, \( z_n(\cdot) \) is formed by the superposition of the direct path, HC scatterers and receiver noise,

\[
z_n(t, u) = \sum_{k=0}^K b_k J(t - \bar{r}_{n,k}(t, u)) \exp[-j \omega_c \bar{r}_{n,k}(t, u)] \exp[-j \omega d_{n,k} t] + \nu_n(t, u)
\]
where \( \nu_0(\cdot) \) is modelled independently for each channel as white Gaussian noise with zero mean and variance \( \sigma^2_\nu \). \( J(\cdot) \) is the jamming signal waveform, \( \bar{\tau}_{n,k}(\cdot) \) is the bistatic delay for the \( k^{th} \) scatterer, \( \omega_{d,k} \) is the fast-time doppler frequency and \( b_k \) is defined as the relative magnitude between the direct-path and HC signal. The zero index refers to the direct-path with \( b_0 = 1 \).

The power spectral density of the jammer signal has an apparent bandwidth \( B \), centred at baseband with power level, \( \sigma^2_J \). Realisations of the jammer signal \( J(\cdot) \) can then be generated by an eigen-decomposition of the jammer autocovariance,

\[
   r_J(\tau) = \sigma^2_J \text{sinc}(B\tau). \tag{4}
\]

A physically based model for the multipath scattering is presented by Beckman, [4] and uses a surface roughness parameter to define the scattering distribution between the SAR and an airborne jammer at heights \( h_P \) and \( h_J \) respectively, separated by a distance \( \bar{x}_J \) in the ground plane. The coefficients, \( b_k = \rho B_k \) for \( k > 1 \) are formed with a HC scaling factor \( \rho \), relative to the direct-path and a random amplitude \( B_k \), determined from the scattering model.

### III. Fast-time STAP

For effective fast-time filtering, the fast-time sample rate, \( \Delta t \) is oversampled by a factor of two to provide increased correlation, [5]. Spatial beamforming for the \( l^{th} \) fast-time range bin requires stacking of both the received data and the signal model to form \( N \times 1 \) spatial vectors,

\[
   \mathbf{x}(t_l, u) = \frac{1}{\sqrt{N}} [x((-N-1)/2(t_l, u)), \ldots, x((N-1)/2(t_l, u))]^T, \\
   \mathbf{s}(u) = \frac{1}{\sqrt{N}} [s((-N-1)/2(u)), \ldots, s((N-1)/2(u))]^T
\]

where \( t_l = l \Delta t \). Beamforming is then performed by matching the received data vector with the spatial steering vector,

\[
   y(t_l, u) = \mathbf{H}(u)\mathbf{x}(t_l, u). \tag{5}
\]

To extend the processing to use fast-time taps, the spatial data vector is stacked over the future \( \hat{L} \ll L \) taps,

\[
   \mathbf{X}(t_l, u) = \frac{1}{\sqrt{L}} \left[ \mathbf{x}^T(t_l, u), \mathbf{x}^T(t_{l-1}, u), \ldots, \mathbf{x}^T(t_{l-L+1}, u) \right]^T
\]

with data components for \( l < \hat{L} \) set to zero. The fast-time component of the steering vector post range processing is given by,

\[
   g_k = \text{sinc}[B(k-1)\Delta t], \quad k = 1 \ldots \hat{L} \tag{6}
\]

and can be stacked to give the fast-time steering vector,

\[
   \mathbf{g} = \frac{1}{\sqrt{L}} [g_1, \ldots, g_{\hat{L}}]^T \in \mathbb{C}^{L \times 1}. \tag{7}
\]

If no oversampling is used, the fast-time model matches the delta function commonly used in literature, [2] where it is assumed that the target occupies a single range bin. The fast-time filter is then represented as a convolution,

\[
   x_l(t_l, u) = \mathbf{S}(u) \mathbf{X}(t_l, u) \tag{8}
\]

with the space/fast-time steering vector formed by the Kronecker product of the spatial and temporal steering vectors,

\[
   \mathbf{S}(u) = \mathbf{g} \otimes \mathbf{s}(u) \in \mathbb{C}^{LN \times 1}. \tag{9}
\]

If the focussing vector in (8) is replaced with a weighted vector, than adaptive processing can be performed,

\[
   x_{wa}(t_l, u) = \mathbf{W}(u) \mathbf{X}(t_l, u). \tag{10}
\]

As shown in the following sections, different criteria can be used to form the weight vector, \( \mathbf{W}(\cdot) \). Note, that for convenience, fast-time samples will be referred to as \( t \).

### IV. Generalised Sidelobe Canceller

The GSC is a sub-optimal beamspace STAP implementation and is shown in Fig. 2. It forms a set of ‘beams’ with the main beam in the ‘desired’ target direction and the other ‘reference’ beams going through a blocking matrix \( \mathbf{B}(\cdot) \) to remove the desired signal from the data. This signal then goes through an adaptive filter to minimise the output power, before being subtracted from the main beam. The GSC is preferred over the traditional element space formulation due to its superior performance with steering vector mismatch. Also due to the loss of \( N_{\text{con}} \) degrees of freedom in the adaption, less training data is required for the same adaptive performance.

For a spatial only filter, the canceller’s output is given by

\[
   x_a(t_l, u) = \mathbf{w}_a^H(u)\mathbf{x}(t_l, u) - \mathbf{w}_a^H(u)\mathbf{B}(u)\mathbf{x}(t_l, u)
   = [\mathbf{w}_a(u) - \mathbf{B}(u)\mathbf{w}_a(u)]^H \mathbf{x}(t_l, u) \tag{11}
\]

where the desired weight \( \mathbf{w}_a(\cdot) \), is given by

\[
   \mathbf{w}_a(u) = \mathbf{c}(u) \left[ \mathbf{c}^H(u)\mathbf{c}(u) \right]^{-1} \mathbf{d} \in \mathbb{C}^{N \times 1} \tag{12}
\]

Fig. 1. SAR processing diagram
with $\mathbf{c}(\cdot)$ containing the $N_{\text{con}}$ adaptive constraints with desired response, $\mathbf{d}$. The adaptive weight vector $\mathbf{w}_a(\cdot)$ of size $(N - N_{\text{con}}) \times 1$, is designed to minimise the output power by solving the unconstrained optimisation [6],

$$\min_{\mathbf{w}_a} \left\{ E \{ |e_0 - 0|^2 \} \right\} \Rightarrow \mathbf{w}_a(u) = [\mathbf{B}^H(u)\mathbf{R}_x(u)\mathbf{B}(u)]^{-1}\mathbf{B}^H(u)\mathbf{R}_x(u)\mathbf{w}_d(u).$$

As the reference beam is orthogonal to the mainbeam and providing there is no mismatch between the input signal and reference beam, the following holds for both the total received covariance, $\mathbf{R}_x(\cdot)$ and the interference plus noise only covariance $\mathbf{R}_z(\cdot)$, [6],

$$\mathbf{B}^H(u)\mathbf{R}_x(u) = \mathbf{B}^H(u)\mathbf{R}_z(u).$$

(13)

The adaptive weight then reduces to,

$$\mathbf{w}_a(u) = [\mathbf{B}^H(u)\mathbf{R}_x(u)\mathbf{B}(u)]^{-1}\mathbf{B}^H(u)\mathbf{R}_x(u)\mathbf{w}_d(u) = \mathbf{R}^{-1}_{x_1, c_0}(u)\mathbf{r}_{x_1, c_0}(u) \in C^{(N - N_{\text{con}}) \times 1}$$

(14)

where $\mathbf{R}_{x_1}(\cdot)$ is the ‘reference’ covariance matrix at $x_1(\cdot)$ and $\mathbf{r}_{x_1, c_0}(\cdot)$ is the cross covariance between $x_1(\cdot)$ and $c_0$. To remove the desired signal, the blocking matrix must be orthogonal to the constraint matrix, $\mathbf{B}^H(u)\mathbf{C}(u) = 0$. A general method for the blocking matrix design has been presented in [3].

To extend this algorithm to use $\tilde{L}$ fast-time taps, the desired fast-time weights are given as,

$$\mathbf{W}_d(u) = \mathbf{g} \otimes \mathbf{w}_d(u) \in C^{	ilde{L}N \times 1}$$

(15)

with the fast-time blocking matrix expanded by,

$$\mathbf{B}_f(u) = \mathbf{I}_{\tilde{L}} \otimes \mathbf{B}(u) \in C^{	ilde{L}N \times \tilde{L}(N - N_{\text{con}})}$$

(16)

and the fast-time adaptive weight of size $\tilde{L}N \times 1$,

$$\mathbf{w}_a(u) = [\mathbf{B}^H(u)\mathbf{R}_{Z,DL}(u)\mathbf{B}_f(u)]^{-1}\mathbf{B}^H(u)\mathbf{R}_{Z,DL}(u)\mathbf{W}_d(u)$$

(17)

where $\mathbf{R}_{Z,DL}(\cdot)$ is the diagonally loaded space/fast-time interference plus noise covariance matrix. The loading is included to improve the robustness by smoothing the adaption via compression of the eigenvalues. The modified covariance matrix with $\eta$ dB of diagonal loading is defined by,

$$\mathbf{R}_{Z,DL}(u) = \mathbf{R}_Z(u) + \eta \mathbf{I}_{\tilde{L}N}. $$

The normalised interference plus noise covariance matrix $\mathbf{R}_Z(u) = \alpha^{-1} \mathbf{R}_Z^2(u)$ is determined by averaging over $L_t$ range bins,

$$\mathbf{R}_Z(u) = \frac{1}{L_t} \sum_{l=1}^{L_t} \mathbf{Z}(t_l, u)\mathbf{Z}^H(t_l, u) \in C^{	ilde{L}N \times \tilde{L}N}$$

(18)

with the normalising value, $\alpha = \text{Tr} (\mathbf{R}^2(u)) / (\tilde{L}N)$ providing a relative measure of the effect of diagonal loading. It is assumed techniques described in [7] can be used to get different realisations of the interference plus noise signal without any targets present. The interference plus noise vector, $\mathbf{Z}(\cdot)$ is formed similarly to the data vector $\mathbf{X}(\cdot)$.

A. GSC Results

A multichannel X-band SAR simulation has been implemented with parameters summarised in Table I. A moderately diffuse scattering scenario is used to demonstrate the worst case scenario with both direct and HC paths incident in the SAR mainbeam. Fig. 3 shows the sample image used with and without interference and after adaptive filtering. Image formation has been performed with a spatial matched filter interpolation algorithm.

### TABLE I

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier frequency ($f_c$) / bandwidth ($B$)</td>
<td>10 / 0.3 GHz</td>
</tr>
<tr>
<td>Number of elements ($N$) / spacing ($\delta$)</td>
<td>5 / $\frac{1}{2} \text{m}$</td>
</tr>
<tr>
<td>Number of pulses ($M$) / range bins ($\tilde{L}$)</td>
<td>100 / 250</td>
</tr>
<tr>
<td>Range / azimuth resolution</td>
<td>0.5 / 2.5 m</td>
</tr>
<tr>
<td>Fast-time sampling ($\Delta_f$) / training size ($L_t$)</td>
<td>$\frac{1}{2\pi} / 3\tilde{L}N$</td>
</tr>
<tr>
<td>SAR height ($h_P$) / jammer height ($h_J$)</td>
<td>3 / 3 km</td>
</tr>
<tr>
<td>Jam. offset ($\bar{x}_J$) / jam. power ($\sigma^2_J$)</td>
<td>50 km / 80 dB</td>
</tr>
<tr>
<td>No. HC scats. ($K$) / relative HC scaling ($\rho$)</td>
<td>200 / 0.6</td>
</tr>
<tr>
<td>Noise power ($\sigma^2_N$) / clutter noise ratio</td>
<td>0 dB / 20 dB</td>
</tr>
</tbody>
</table>

![SAR image](image1.png) ![SAR image with hot-clutter](image2.png) ![Filtered SAR image](image3.png)

**Fig. 3.** Simulated image comparison

Both the Minimum Variance Distortionless Response (MVDR) and first order derivative constraints are used for this study with the former having a unity response in the steering direction and the latter also setting the first derivative equal to zero. This has shown to be effective as it allows less potential signal suppression [3]. The combined spatial constraint matrix, $\mathbf{c}(\cdot)$ and desired vector, $\mathbf{d}$ are defined as,

$$\mathbf{c}(u) = \left[ \frac{\partial \phi(u)}{\partial (\mathbf{B}(u))} \right]; \quad \mathbf{d} = [1, 0]^T$$

(19)

They are also related to the space/fast-time versions by,

$$\mathbf{C}(u) = \mathbf{I}_{\tilde{L}} \otimes \mathbf{c}(u) \in C^{\tilde{L}N \times \tilde{L}N_{\text{con}}},$$

$$\mathbf{D} = \mathbf{g} \otimes \mathbf{d} \in C^{\tilde{L}N_{\text{con}} \times 1}.$$  

(20)
The adaptive performance is measured by the Signal Distortion Ratio (SDR) which is a measure of the signal power of the adapted image relative to an ideal image with no interference present. For comparison, the conventional SDR with no adaption is 3.8 dB. The results in Fig. 4 show the full rank GSC and the SDR improvement gained with using fast-time taps. The plot of the left is with the MVDR constraint and the one on the right uses derivative constraints. With -60 dB of diagonal loading and 15 fast-time taps, the MVDR SDR improves from 4 to 6.5 dB. The derivative constraint SDR however reaches a maximum 7.1 dB and maintains this level between -90 dB to -60 dB of diagonal loading.

The choice of transform, $U(\cdot)$ can be found by an eigen-decomposition of the reference interference plus noise covariance matrix,

$$
R_{X_1}(u) = QAQ^{H} \in \mathbb{C}^{L(N-N_{con}) \times L(N-N_{con})} \quad (22)
$$

and choosing $C$ eigenvalue/eigenvector pairs, $(\lambda_c, q_c)$ according to a ranking criteria. If the eigen-pairs are ranked according the eigenvalues, the transform is known as the Principle Component (PC) decomposition,

$$
U(u) = [q_1, \ldots, q_C] \in e^{L(N-N_{con}) \times C}. \quad (23)
$$

By using this approach, the reduction in computational complexity of the matrix inverse is reduced from $O(L(N-N_{con})^3)$ to $O(C)^3$.

A. Reduced Rank GSC Results

Fig. 6 shows how the SDR varies as a function of the rank and the level of diagonal loading. The MVDR results are on the left and are very sensitive to diagonal loading. The derivative constraint result however is not as greatly affected by the diagonal loading level. With a filter rank of only 10 and a small amount of diagonal loading, this filter can safely achieve the same SDR level as the full rank case.

### V. REDUCED RANK GSC

Using the space/fast-time formulation, the reduced rank GSC is shown in Fig. 5 where a reduced rank transform is now contained in the matrix $U(u)$. This acts to reduce the size and rank of the reference beam after going through the blocking matrix. The output of the reduced rank GSC is given by,

$$
x_{L_1}(t, u) = W^H_{a}(u)X(t, u) - W^H_{a}(u)U^H(u)B^H_t(u)X(t, u)
= [W^H_{a}(u) - B^H_t(u)U^H(u)W^H_{a}(u)]X(t, u) \quad (21)
$$

where the adaptive weight vector of size $C \times 1$ is again designed to minimise the mean square error between $\hat{e}_0$ and $e_0$. The solution is given by [8] with the covariance matrix substitution used from (13) and diagonal loading now included after the rank reduction,

$$
W_{a}(u) = [U^H(u)B^H_t(u)R_{Z}(u)B_t(u)U(u) + \eta I_C]^{-1}U^H(u)B^H_t(u)R_{Z}(u)W_{a}(u)
= [U^H(u)R_{X_1}(u)U(u) + \eta I_C]^{-1}U^H(u)r_{X_1,e_0}(u).
$$

The Multistage Wiener Filter (MWF) is the final algorithm and provides a faster rank reduction using a nested chain of traditional Wiener filter stages. This method does not need an eigenvector decomposition or large covariance matrix inversion which makes it more suitable for real world implementation. The space/fast-time constrained MWF of order $P$ is formed from $P$ filter stages as shown in Fig. 7, where $null[W_{a}(u)]$ represents the nullspace of $W_{a}(u)$.
The output from the $p^{th}$ MWF stage is given by,
\[ x_{d}(t, u) = [W_{d}(u) - L_{p}(u)W_{a,p}(u)]^{H}X(t, u) \]  
(24)
where the sequential vector, $L_{p}(\cdot)$ is defined by,
\[ L_{p}(u) = \{B_{1}(u)h_{1}(u), B_{1}(u)B_{2}(u)h_{2}(u), \ldots, B_{1}(u)B_{2}(u) \cdots B_{p}(u)h_{p}(u)\} \in \mathbb{C}^{LN \times p}. \]
(25)
The rank one space/fast-time basis vectors, $h_{p}(\cdot)$ is designed to maximise the cross-correlation energy between $e_{p}$ and $e_{p-1}$,
\[ h_{p}(u) = \frac{r_{x_{p},e_{p-1}}}{\sqrt{r_{x_{p},e_{p-1}}^{H}(u)r_{x_{p},e_{p-1}}(u)}} \in \mathbb{C}^{L(N-N_{con}) \times 1} \]  
(26)
with reference covariance and cross covariance,
\[ R_{x_{p}}(u) = B_{p}^{H}(u)R_{x_{p-1}}(u)B_{p}(u) \in \mathbb{C}^{L(N-N_{con}) \times L(N-N_{con})}, \]
\[ r_{x_{p},e_{p-1}}(u) = B_{p}^{H}(u)R_{x_{p-1}}(u)h_{p-1}(u) \in \mathbb{C}^{L(N-N_{con}) \times 1} \]
where $h_{0}(u) = W_{d}(u)$, $R_{x_{0}}(u) = R_{Z}(u)$ and the blocking matrices of size $L(N-N_{con}) \times L(N-N_{con})$ for $p > 1$,
\[ B_{p}(u) = I_{L(N-N_{con})} - h_{p-1}(u)h_{p-1}^{H}(u). \]  
(27)
The size of the $p^{th}$ order space/fast-time weight vector is then determined by the MWF order. Diagonal loading level is again included after the rank reduction,
\[ W_{a,p}(u) = \left[ L_{p}^{H}(u)R_{Z}(u)L_{p}(u) + \eta I_{p} \right]^{-1} \]
\[ L_{p}^{H}(u)R_{Z}(u)W_{d}(u) \in \mathbb{C}^{p \times 1}. \]  
(28)

**A. MWF Results**

Fig. 8 shows the simulated results with both the filter order and diagonal loading level varied with $L = 15$ fast-time taps. It takes an order of 14 before the MWF with MVDR constraints behaves like the full rank case. In contrast, the derivative constraint results show that a small order of 5 can meet the full rank case without diagonal loading! This is huge difference of 11 filter orders and demonstrates the superiority of using the derivative constraints with the MWF.

**VII. CONCLUSION**

This paper has shown how the constrained fast-time GSC can be formulated with reduced rank and computational complexity. Equivalent full rank results were achieved with derivative constraints for both the PC decomposition and the MWF. The latter filter however offers the most reduced computational complexity by using the adaptive degrees of freedom more effectively to remove the HC interference and minimise distortion to the final SAR image.

**REFERENCES**