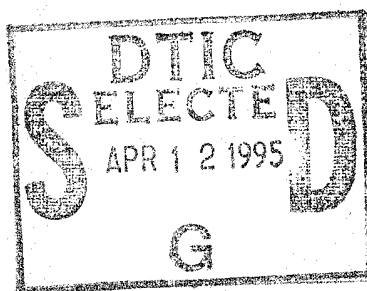


**Technical Report
1015**

Space-Time Adaptive Processing for Airborne Radar



J. Ward

13 December 1994

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS



Prepared for the Advanced Research Projects Agency under
Air Force Contract F19628-95-C-0002.

Approved for public release; distribution is unlimited.

19950410 076

DTIC QUALITY INSPECTED 6


This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Advanced Research Projects Agency under Air Force Contract F19628-95-C-0002.

This report may be reproduced to satisfy needs of U.S. Government agencies.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER


Gary Tutungian
Administrative Contracting Officer
Contracted Support Management

Non-Lincoln Recipients

PLEASE DO NOT RETURN

Permission is given to destroy this document
when it is no longer needed.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

**SPACE-TIME ADAPTIVE PROCESSING
FOR AIRBORNE RADAR**

*J. WARD
Group 102*

TECHNICAL REPORT 1015

13 DECEMBER 1994

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

Approved for public release; distribution is unlimited.

LEXINGTON

MASSACHUSETTS

ABSTRACT

Future airborne radars will be required to detect targets in an interference background comprised of clutter and jamming. Space-time adaptive processing (STAP) refers to multidimensional adaptive filtering algorithms that simultaneously combine the signals from the elements of an array antenna and the multiple pulses of a coherent radar waveform, to suppress interference and provide target detection. STAP can improve detection of low-velocity targets obscured by mainlobe clutter, detection of targets masked by sidelobe clutter, and detection in combined clutter and jamming environments. This report analyzes a variety of approaches to STAP problem. Optimum, or fully adaptive processing is reviewed. Computational complexity and the need to estimate the interference from a limited amount of available data make fully adaptive STAP impractical. As a result, partially adaptive space-time processors are required. A taxonomy of reduced-dimension STAP algorithms is presented where algorithms are classified based on the type of preprocessor employed. For example, beamspace algorithms use spatial preprocessing, while post-doppler approaches perform temporal (Doppler) filtering before adaptive processing. In some cases, the special structure of the clutter can be exploited to design preprocessors yielding minimum clutter rank. For each class, either sample-matrix-inversion (SMI) or subspace-based weight computation may be employed. Simulation results are presented to illustrate various performance metrics, including SINR, adapted patterns, minimum detectable velocity, and required degrees of freedom.

ACKNOWLEDGMENTS

This report has been influenced by many people within Lincoln Laboratory. I would like to thank Ken Senne for the technical guidance and encouragement he has given me throughout this work. Special thanks go also to Dave Martinez for providing me with time, support, and motivation to complete this report.

Much of my education in the area of space-time adaptive processing is a direct result of many enlightening discussions with Allan Steinhardt, Steve Krich, Ed Baranoski, Steve Smith, Ken Teitelbaum, and Dan Marshall. Their contributions are gratefully acknowledged. Early encouragement from Allan Steinhardt and Steve Krich was particularly appreciated. My thanks go again to Allan and Ed for their careful reviews of the draft; their comments have greatly improved the presentation. Professor Richard Bucy of the University of Southern California also provided helpful comments on portions of this report.

Finally, I would like to acknowledge the continued sponsorship of CAPT. Ryan Henry of the Advanced Research Projects Agency. Also, I thank Jim Hall of the Office of Naval Research, Nancy MacMeekin of Naval Air Warfare Center, and CAPT. Dale Babin of Naval Air Systems Command for their continued interest in the area of space-time adaptive processing technology. Their support over the last two years is gratefully acknowledged.

TABLE OF CONTENTS

Abstract	iii
Acknowledgments	v
List of Illustrations	ix
List of Tables	xv
 1. INTRODUCTION	 1
 2. AIRBORNE ARRAY RADAR SIGNAL ENVIRONMENT	 7
2.1 Introduction	7
2.2 Radar System Description	7
2.3 Target	12
2.4 Noise	17
2.5 Jamming	18
2.6 Clutter	20
2.7 Range-Doppler View of the Interference	49
2.8 Summary	51
 3. SPACE-TIME PROCESSING FUNDAMENTALS	 53
3.1 Introduction	53
3.2 General Architecture and Assumptions	53
3.3 Fully Adaptive STAP	57
3.4 STAP Performance Metrics and Fully Adaptive Performance	59
3.5 Sample Support and SMI	74
3.6 Computational Complexity	77
 4. PARTIALLY ADAPTIVE STAP	 81
4.1 Introduction	81
4.2 A Generic Architecture	81
4.3 Exploiting Low-Rank Interference	83
4.4 Application to STAP	88
 5. ELEMENT-SPACE STAP	 95
5.1 Introduction	95
5.2 Element-Space Pre-Doppler STAP	95
5.3 Element-Space Post-Doppler	109

TABLE OF CONTENTS (Continued)

6. BEAMSPACE STAP	135
6.1 Introduction	135
6.2 Beam-space Pre-Doppler STAP	136
6.3 Beam-Space Post-Doppler STAP	153
7. ADDITIONAL PERFORMANCE RESULTS	169
8. SUMMARY AND FUTURE WORK	185
APPENDIX A – PROOF OF BRENNAN’S RULE	187
APPENDIX B – PROOF OF THEOREM 2	191
APPENDIX C – DISPLACED PHASE CENTER ANTENNA (DPCA) PROCESSING	195
REFERENCES	201

LIST OF ILLUSTRATIONS

Figure No.		Page
1	Space-time adaptive processing. (a) The power spectral density of the interference (clutter and jamming) seen by an airborne radar (b) Example adapted response.	3
2	(a) Platform geometry. (b) Top view.	8
3	Processing for each array element channel.	10
4	The radar CPI datacube.	10
5	A ring of ground clutter for a fixed range.	21
6	The $\beta = 1$ clutter ridge. The PRF is 300 Hz. (a) Doppler frequency vs. $\sin \phi$. (b) Normalized Doppler vs. spatial frequency.	26
7	Clutter loci for different velocities. (a) Stationary platform, $v_a = 0$. (b) Doppler unambiguous, $\beta = 0.5$. (c) Doppler ambiguous, $\beta = 1.5$. (d) Doppler ambiguous, $\beta = 2.67$.	27
8	Effective array position for successive pulses of a CPI. $N = 4$, $M = 3$, $\beta = 1$. Clutter observations are repeated by different elements on different pulses.	32
9	The element pattern. A -30 -dB backlobe level is assumed.	32
10	Example scenario: received CNR per column as a function of azimuth.	35
11	Illustrating Brennan's rule: clutter eigenspectra for the example radar system with different platform velocities.	35
12	Array geometry with velocity misalignment.	38
13	Example clutter ridges with velocity misalignment, for $\beta = 1$. (a) $\phi_a = 0^\circ$. (b) $\phi_a = 10^\circ$. (c) $\phi_a = 45^\circ$. (d) $\phi_a = 90^\circ$.	40
14	Example clutter ridges with velocity misalignment, for Doppler-ambiguous clutter. (a) $\beta = 2.6$, $\phi_a = 10^\circ$. (b) $\beta = 2.67$, $\phi_a = 60^\circ$.	41
15	Clutter eigenspectra for different misalignment angles.	43
16	Clutter eigenspectra with misalignment for different backlobe levels.	43
17	Clutter eigenspectra for different levels of ICM.	47
18	Range-Doppler view of interference, for the case of a low-PRF waveform. The clutter is ambiguous in Doppler and unambiguous in range.	50

LIST OF ILLUSTRATIONS

(Continued)

Figure No.		Page
19	Range-Doppler view of interference, for a high-PRF waveform. The clutter is unambiguous in Doppler but ambiguous in range.	50
20	A general block diagram for a space-time processor.	54
21	Data-domain view of space-time adaptive processing.	55
22	Fully adaptive space-time processing	60
23	Example scenario: optimum fully adaptive STAP. (a) Adapted pattern. (b) Principal cuts at target azimuth and Doppler.	63
24	Example scenario: tapered fully adaptive STAP. (a) Adapted pattern. (b) Principal plane cuts at target azimuth and Doppler.	65
25	SINR for the optimum and tapered fully adaptive STAPs.	68
26	SINR for the optimum and tapered fully adaptive STAPs, including Doppler straddling losses.	69
27	SINR loss for the fully adaptive STAP.	70
28	SINR improvement factor for the optimum and tapered fully adaptive STAP.	72
29	Expected SINR loss for SMI with matched steering vector.	79
30	A generic partially adaptive STAP architecture.	82
31	(a) Reduced dimension adaptive processing with low-rank interference. (b) Whitening filter viewpoint.	86
32	Another block diagram for reduced-dimension space-time adaptive processing.	92
33	A taxonomy of reduced-dimension STAP algorithms, classified by the type of nonadaptive transformation applied to the CPI data.	93
34	Element-space pre-Doppler STAP. (a) Sub-CPI processing. (b) Full CPI processing.	97
35	$K = 2$ element-space pre-Doppler STAP. (a) Sub-CPI adapted pattern. (b) Principal cuts in angle and Doppler.	103
36	Sub-CPI weight vector decomposed into spatial beamformer components, illustrating the DPCA effect.	106

LIST OF ILLUSTRATIONS

(Continued)

Figure No.		Page
37	Composite pattern for Doppler bin 5. (a) Full pattern. (b) Principal cuts in angle and Doppler.	107
38	SINR loss for element space pre-Doppler. Zero intrinsic clutter motion.	109
39	Block diagram of post-Doppler adaptive beamforming (factored post-Doppler).	112
40	Example performance for Doppler bin 6 (100 Hz) with 40-dB Chebyshev Doppler filters. (a) Clutter power spectral density. (b) Adapted pattern.	115
41	Example performance for Doppler bin 6 (100 Hz) with 80-dB Chebyshev Doppler filters. (a) Clutter power spectral density. (b) Adapted pattern.	115
42	SINR loss as a function of Doppler filter sidelobe level.	117
43	Block diagram for multiwindow post-Doppler STAP	120
44	Multiwindow post-Doppler form of Brennan's rule. Example for $N = 4$, $M = 3$, $K = 2$, and $\beta = 1$. The grouping of pulses for the two filters satisfies Theorem 3.	124
45	Block diagram for PRI-staggered post-Doppler STAP.	126
46	Block diagram for adjacent-bin post-Doppler STAP.	127
47	Clutter eigenspectra for multiwindow post-Doppler approaches with $K = 2$. (a) PRI-staggered post-Doppler, covariance matrix for Doppler bin 0. (b) PRI-staggered post-Doppler, Doppler bin 3. (c) PRI-staggered post-Doppler, Doppler bin 9. (d) Adjacent-bin post-Doppler, Doppler bin 0. (e) Adjacent-bin, Doppler bin 3. (f) Adjacent-bin, Doppler bin 9.	129
48	SINR loss performance for PRI-staggered and adjacent-bin post-Doppler STAP. (a) Untapered (uniform) Doppler filters. (b) 30-dB Chebyshev Doppler filters. (c) 60-dB Chebyshev Doppler filters. (d) 90- dB Chebyshev Doppler filters.	133
49	Beamspace pre-Doppler STAP. (a) Sub-CPI processing. (b) Full CPI processing.	137
50	Beamspace version of Brennan's rule for the case $N = 4$, $K_t = 3$, $K_s = 2$, and $\beta = 1$.	142

LIST OF ILLUSTRATIONS

(Continued)

Figure No.		Page
51	SINR loss performance for beamspace pre-Doppler STAP in a clutter-only scenario. (a) Displaced beams, untapered. (b) Adjacent beams, untapered. (c) Displaced beams, 30-dB Chebyshev taper. (d) Adjacent beams, 30-dB Chebyshev taper.	145
52	Displaced-beam pre-Doppler STAP performance with ambiguous clutter and no jamming. A 30-dB Chebyshev taper is assumed. (a) $\beta = 2$. (b) $\beta = 3$.	149
53	A two-step nulling approach to beamspace pre-Doppler STAP.	151
54	Generic block diagram for beamspace post-Doppler. Single Doppler bin processing is shown.	153
55	Beamspace post-Doppler variation on Brennan's Rule. $N = 4$, $M = 3$, $K_s = K_t = 2$, $\beta = 1$. If the conditions are satisfied, clutter observations are repeated by different outputs.	157
56	Block diagram for beamspace post-Doppler utilizing a single 2D-DFT, Cai, Wang, and Wicks[10,39].	160
57	Filter selection for adjacent-filter beamspace post-Doppler. Key: 'o'-available filters, '⊗'-target filter. The box encircles the selected filters for $K_{sm} = 3$, $K_{tm} = 3$.	161
58	Beamspace post-Doppler in a clutter-only environment. $K = 4$, $K_{sm} = K_{tm} = 2$.	163
59	Beamspace post-Doppler in a clutter-only environment. $K = 9$, $K_{sm} = K_{tm} = 3$.	163
60	Beamspace post-Doppler in a clutter-only environment. $K = 16$, $K_{sm} = K_{tm} = 4$.	165
61	Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 4$, $K_{sm} = K_{tm} = 2$.	165
62	Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 9$, $K_{sm} = K_{tm} = 3$.	167
63	Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 16$, $K_{sm} = K_{tm} = 4$.	167

LIST OF ILLUSTRATIONS (Continued)

Figure No.		Page
64	Performance for example system, $\beta = 1$, no misalignment, no intrinsic clutter motion. (a) SINR loss. (b) MDV and UDSF performance.	171
65	Performance for example system, $\beta = 2.6$, no misalignment, no intrinsic clutter motion. (a) SINR loss. (b) MDV and UDSF.	175
66	Performance for example system, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, no velocity misalignment. (a) SINR Loss (b) MDV and UDSF.	177
67	Performance for example system, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 10° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.	179
68	Performance for example system with -40 -dB backlobe, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 10° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.	181
69	Performance for example system with -40 -dB backlobe, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 90° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.	183
C-1	Block diagram for DPCA processing.	196

LIST OF TABLES

Table No.		Page
1	Symbols for Radar System Parameters	11
2	Radar System and Antenna Array Parameters for Example Scenario	31
3	Assumptions Made for Radar System and Signal Model	52
4	Platform and Interference Scenario for Baseline Scenario.	61
5	MDV and UDSF for fully adaptive STAP	74

1. INTRODUCTION

Future airborne radars will be required to provide long-range detection of increasingly smaller targets. This function must be performed in overland or littoral environments where the ground clutter can be quite severe, and in the presence of hostile electronic countermeasures, or jamming. Future radars must therefore possess, in addition to sufficient power and aperture, the capability to suppress both clutter and jamming to near or below the noise level. Only then will the radar's sensitivity be fully utilized in signal environments containing unwanted interference.

The problems of clutter and jamming suppression have dominated the minds of radar engineers since the beginnings of radar. The ground clutter seen by an airborne radar is extended in both range and angle; it also is spread over a region in Doppler due to the platform motion. A potential target may be obscured by not only mainlobe clutter that originates from the same angle as the target but also by sidelobe clutter that comes from different angles but has the same Doppler frequency. The effect of sidelobe clutter may be eliminated with low- enough antenna sidelobes on transmit and receive. Achieving very low-sidelobes in practice can be quite difficult, especially in the complex electromagnetic environment on an aircraft, and very expensive. In many cases the need to detect slow-moving targets prevents the radar from utilizing high-pulse repetition frequency (PRF) waveforms to avoid clutter completely. Displaced-phase-center-antenna (DPCA) processing [2, 1] was developed to solve the clutter problem in airborne surveillance radar. With DPCA the receive aperture is shifted from pulse to pulse to compensate for the platform motion. Outputs from different apertures with the same effective phase center are subtracted on a pulse-to-pulse basis to cancel the clutter. DPCA requires velocity-based control of the radar PRF. DPCA's sensitivity to element mismatch errors limits the clutter cancellation achievable in practice, and DPCA alone has no inherent provisions for suppressing jamming signals.

Adaptive array processing [22, 23] techniques developed over the last 25–30 years have been successfully used to mitigate the effects of jamming on communication and radar systems. These techniques, by virtue of their adaptive nature, can provide nulling far below the sidelobe level limitation due to random errors. In an adaptive beamforming radar, however, special measures must be taken to avoid the inclusion of mainlobe clutter during the adaptation process. For example, the use of a special listening interval to sample the jamming [33] avoids this problem, but this solution takes away from the radar timeline and creates a vulnerability to a nonstationary jamming environment. Similarly, adaptive beamforming changes the receive antenna patterns and therefore affects the clutter, which must subsequently be suppressed.

Space-time adaptive processing (STAP) refers to the extension of adaptive antenna techniques to processors that *simultaneously* combine the signals received on multiple elements of an antenna array (the spatial domain) and from multiple pulse repetition periods (the temporal domain) of a coherent processing interval (CPI). STAP is therefore a generalization of DPCA processing that utilizes many elements and pulses as well as data-adaptive combination of the various signals. STAP offers the potential to improve airborne radar performance in several areas. First, it can

improve low-velocity target detection through better mainlobe clutter suppression. Second, STAP can permit detection of small targets that might otherwise be obscured by sidelobe clutter. Third, STAP provides detection in combined clutter and jamming environments. Finally, STAP adds robustness to system errors and a capability to handle nonstationary interference.

A basic illustration of space-time adaptive processing is given in Figure 1. A pictorial view of the interference environment seen by an airborne radar is shown in Figure 1(a). The signal-to-noise-ratio (SNR) resulting from clutter and a single jamming signal is shown as a function of angle and Doppler. Barrage noise jamming is localized in angle and distributed over all Doppler frequencies. The clutter echo from a single ground patch has a Doppler frequency that depends on its aspect with respect to the platform; clutter from all angles lies on the "clutter ridge" shown in Figure 1(a). Note that a mainlobe target competes with both mainlobe and sidelobe clutter as well as jamming. A space-time adaptive processor may be thought of as a two-dimensional filter that represents combined receive beamforming and target Doppler filtering. An example adapted response is shown in Figure 1(b). Note the high gain at the target angle and Doppler, and the deep nulls along both the jamming and clutter lines. Applying this filter to the data will suppress the interference and enable target detection. A bank of adaptive filters is then formed to cover all potential target angles and velocities.

The first published work on space-time adaptive processing for radar was Brennan and Reed [3] from 1973, in which optimum space-time filtering was described. Other work has been done by Klemm [5, 6, 7, 8], who first tried to understand the fundamental degrees of freedom occupied by airborne clutter. The significant advances over the last two decades in digital signal processing technology have made the real-time implementation of STAP feasible with current or near-term technology. As a result, STAP is becoming a more active research area. Barile et al. [9] have considered some practical limitations to STAP performance. Richardson [43] has studied the relationship between STAP and DPCA processing. DiPietro [12] has examined a STAP algorithm that utilizes several adjacent Doppler bins of data for adaptation. Cai and Wang [10] and Wicks [39] have described a novel approach where a subset of space-time filter outputs is adaptively weighted. Their architecture is coupled with a Generalized-Likelihood-Ratio Test (GLRT) that provides an embedded constant-false-alarm-rate (CFAR) property. Much of the recent work has been focused on a particular issue or algorithm. Many additional papers on adaptive detection and sensor array signal processing also have direct application to STAP.

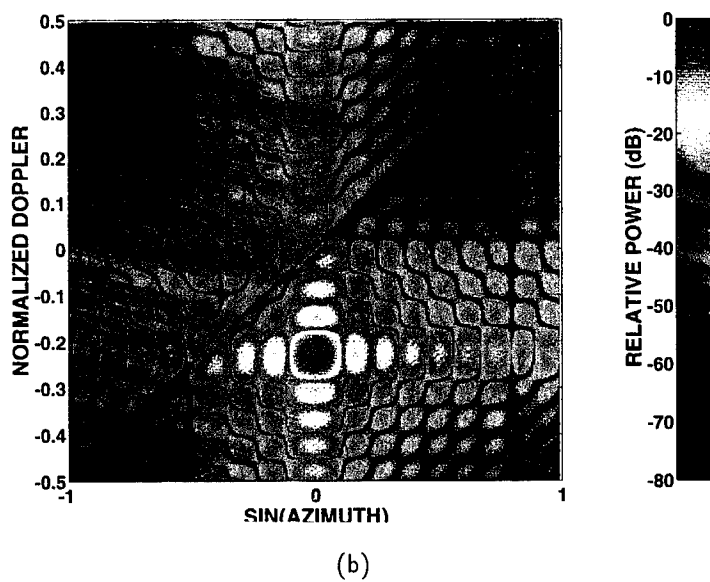
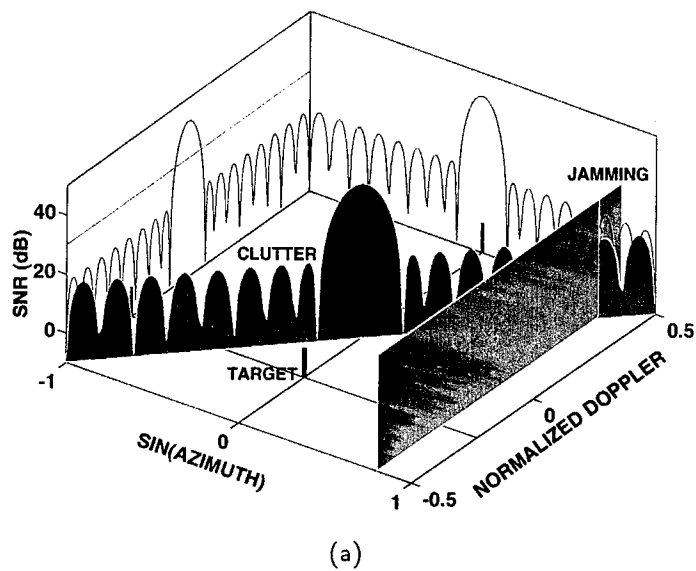


Figure 1. Space-time adaptive processing. (a) The power spectral density of the interference (clutter and jamming) seen by an airborne radar (b) Example adapted response.

The primary aim of this report is to analyze a variety of approaches to the space-time processing, starting from a common set of assumptions and system parameters. To this end, some algorithms are examined that have been described in the references above; variations to these approaches as well as some new approaches are also considered. Fully adaptive STAP, where a separate adaptive weight is applied to all elements and pulses, is addressed first. Fully adaptive STAP requires the solution to a system of linear equations of size MN , where N is the number of array elements and M is the number of pulses per CPI. For typical radar systems, the product MN may vary from several hundreds to tens of thousands. Fully adaptive STAP is impractical for two major reasons. First is the sheer computational power needed to solve large systems of equations in the time necessary for real-time radar operation. Second, the interference is typically unknown a priori and must be estimated from the limited data available from a radar dwell. As the dimension of the adaptive weight vector becomes larger, so does the amount of data required for a good estimate of the environment. The inherent nonstationarity of radar clutter further limits the amount of data over which the clutter can be assumed stationary, which in turn makes its estimation even more difficult. These concerns lead naturally to investigation of reduced-dimension, or partially adaptive, STAP architectures. A taxonomy of reduced-dimension STAP algorithms is presented in which algorithms are classified according to the domain in which the weight application occurs. The clutter seen by an airborne platform has a special structure induced by the platform motion; this structure is exploited in the design of some partially adaptive algorithms.

The topic of this report is the use of adaptive processing for improved airborne radar system performance. Although it is not discussed herein, an adaptive capability may in fact impact basic aspects of radar system design. In the past, the need for ultralow sidelobes or high-PRF waveforms was driven largely by the need to avoid clutter. A capability to suppress clutter adaptively may allow the rethinking of conventional wisdom in some areas. For example, medium- or low-PRF waveforms (with better range information) may be permissible in systems that had been primarily high PRF. The lessening of requirements on overall sidelobe level may have beneficial effects on the resulting component tolerances required and therefore even the cost of future radar systems.

The organization of this report is as follows. Chapter 2 defines the model for the signals received by an airborne radar utilizing an array antenna. Chapter 3 considers some fundamentals, including fully adaptive processing and a number of performance measures. Chapter 4 presents a theory for reduced-dimension adaptive processing and a framework for classifying STAP algorithms. Element-space approaches are discussed in detail in Chapter 5, and beamspace approaches are presented in Chapter 6. Chapter 7 provides additional performance comparisons. Finally, Chapter 8 gives a summary and some areas for future work. For reference, a brief review of DPCA processing is provided in Appendix 3.

2. AIRBORNE ARRAY RADAR SIGNAL ENVIRONMENT

2.1 Introduction

In this chapter, a model is developed for the signals received by an airborne pulsed-Doppler radar. The radar utilizes an array antenna with an independent receiver channel behind each element. The received signals will always contain a component due to receiver noise and may contain components due to both desired targets and undesired interference. Here, interference means either jamming, clutter, or both. The parameters for a generic radar system are used to derive expressions for each of these potential signal components. Since the clutter received by the airborne radar is the most complicated signal component, its development is the largest portion of the chapter. The results developed here form the basis for the analysis of the various space-time processing approaches that are discussed in the subsequent chapters of this report.

Space-time processing is a multidimensional filtering problem. To avoid a proliferation of subscripts, matrix notation is used where possible. Although the concepts discussed in this chapter are not new, the formulation of multidimensional radar data and covariance matrices in terms of Kronecker and Hadamard matrix products may be new to some readers.

2.2 Radar System Description

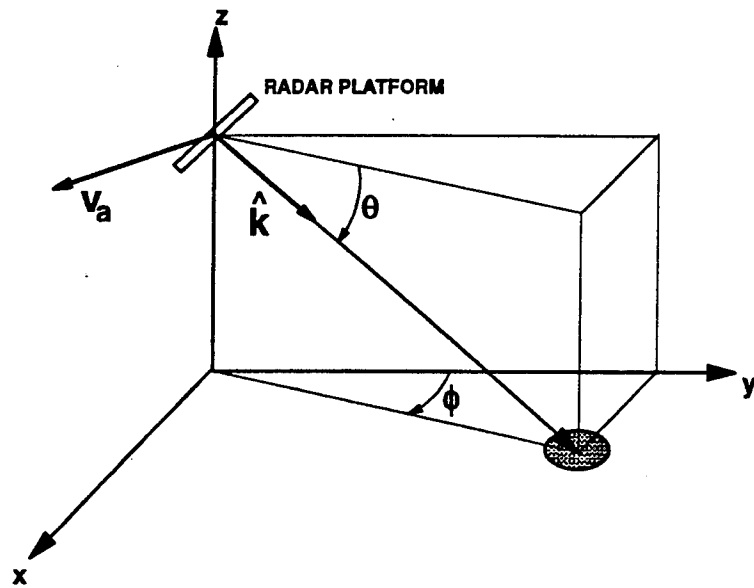
The system under consideration is a pulsed Doppler radar residing on an airborne platform. The radar antenna is a uniformly spaced linear array antenna (ULA) consisting of N elements. These elements may be the beamformed columns of a rectangular planar array. The platform is at an altitude h_a and moving with constant velocity vector \mathbf{v}_a . The chosen coordinate system is shown in Figure 2. The angle variables ϕ and θ refer to true azimuth and elevation, and not the standard spherical coordinate system angles. A unit vector $\hat{\mathbf{k}}$ pointing in the (ϕ, θ) direction is given by

$$\hat{\mathbf{k}}(\phi, \theta) = \cos \theta \sin \phi \hat{\mathbf{x}} + \cos \theta \cos \phi \hat{\mathbf{y}} + \sin \theta \hat{\mathbf{z}}, \quad (1)$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are the unit vectors of a Cartesian coordinate system. Let the array orientation be specified by the interelement position vector \mathbf{d} , so that the n th element, $n = 0, \dots, N - 1$, has position (at some reference time)

$$\mathbf{r}_n = n\mathbf{d}. \quad (2)$$

In this report, only horizontally oriented antennas are considered, and without loss of generality, $\mathbf{d} = d\hat{\mathbf{x}}$ is chosen, where d is the interelement distance. With this definition the chosen angle coordinates are referenced to the array normal. It is also assumed that the array elements are identical, with radiation (voltage) pattern denoted by $f(\phi, \theta)$ and power pattern $g(\phi, \theta) = |f(\phi, \theta)|^2$.



(a)

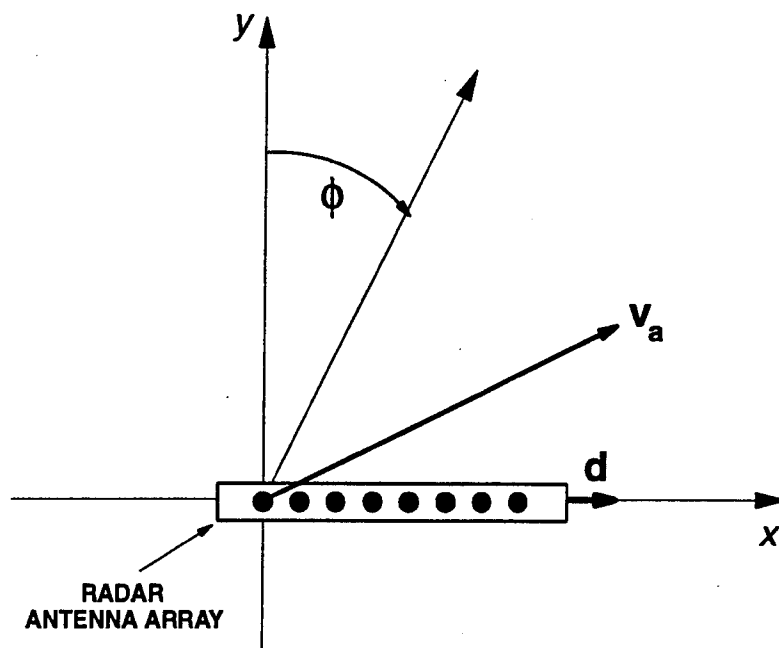


Figure 2. (a) Platform geometry. (b) Top view.

A fixed transmit pattern $G_t(\phi, \theta)$ for the array is assumed, which may or may not be tapered to produce low sidelobes. The radar transmits a coherent burst of M pulses at a constant PRF $f_r = 1/T_r$, where T_r is the pulse repetition interval (PRI). The transmitter carrier frequency is $f_o = c/\lambda_o$, where c is the propagation velocity. The time interval over which the waveform returns are collected is commonly referred to as the coherent-processing interval (CPI). The CPI length is equal to MT_r . A pulse waveform of duration T_p and bandwidth B is assumed. On receive, each element of the array has its own down-converter, matched filter receiver, and A/D converter, as shown in Figure 3.

Since the receiver is a matched filter, the receiver bandwidth B is taken to be equal to that of the transmitted pulse. Matched filtering is done separately on the returns from each pulse, after which the signals are sampled by the A/D converter and sent to a digital processor. The digital processor performs all subsequent radar signal and data processing. For each PRI, L time (range) samples are collected to cover the range interval. With M pulses and N receiver channels, the received data for one CPI comprises LMN complex baseband samples. This multidimensional data set is often visualized as the $L \times M \times N$ cube of complex samples shown in Figure 4. This assembly will be referred to as the radar or CPI datacube.

To facilitate this discussion, some notation must be developed for referring to portions of the radar data cube. First the mathematical conventions must be defined. Scalar quantities are denoted with italic typeface. Lowercase boldface quantities denote vectors and uppercase boldface quantities denote matrices. The operations of transposition, complex conjugation, and conjugate transposition are denoted by superscripts T , $*$, and H , respectively. MATLAB notation for separating the elements of a vector or matrix is employed, e.g., given scalars x and y , the vector $[x; y]$ is a 2×1 column vector, while the vector $[x, y]$ is a 1×2 row vector. Similarly, for two $N \times 1$ vectors \mathbf{a}_1 and \mathbf{a}_2 , $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2]$ is an $N \times 2$ matrix while $\mathbf{a} = [\mathbf{a}_1; \mathbf{a}_2]$ is a $2N \times 1$ vector. The symbols \odot and \otimes refer to the Hadamard and Kronecker¹ matrix products [13], respectively. The notation $\mathbf{T} = \text{diag}(\mathbf{t})$ denotes a diagonal matrix whose main diagonal is the vector \mathbf{t} . The notation $\mathbf{x} = \text{vec}(\mathbf{X})$ will define a vector \mathbf{x} formed by stacking the columns of the matrix \mathbf{X} . Also, the notation $\mathbf{R} = \text{Toeplitz}(\mathbf{c}, \mathbf{r})$ denotes a Toeplitz matrix whose first column is the vector \mathbf{c} and whose first row is the vector \mathbf{r} . Finally, the $E\{\}$ symbol denotes the expected value of a random quantity.

Let x_{nml} be the complex sample from the n th element, m th pulse, at the l th sample time (range gate). Let $\mathbf{x}_{m,l}$ be the $N \times 1$ vector of antenna element outputs, or a spatial snapshot, at the time of the l th range gate and m th pulse. Now let the $N \times M$ matrix \mathbf{X}_l consist of the spatial snapshots for all pulses at the range of interest,

$$\mathbf{X}_l = \begin{bmatrix} \mathbf{x}_{0,l} & \mathbf{x}_{1,l} & \dots & \mathbf{x}_{M-1,l} \end{bmatrix}. \quad (3)$$

¹Also called the tensor product.

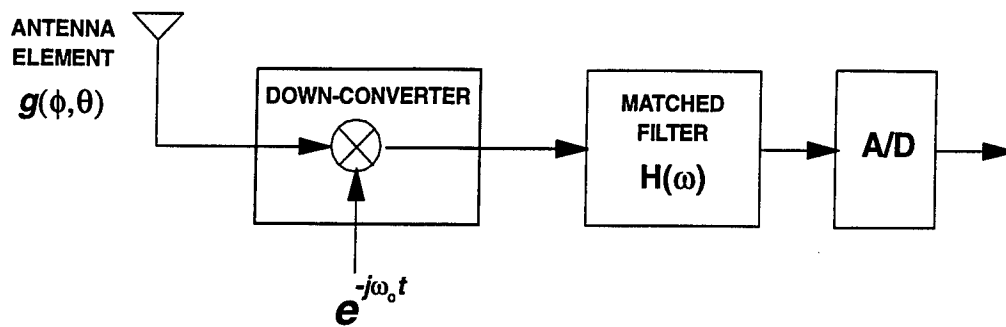


Figure 3. Processing for each array element channel.

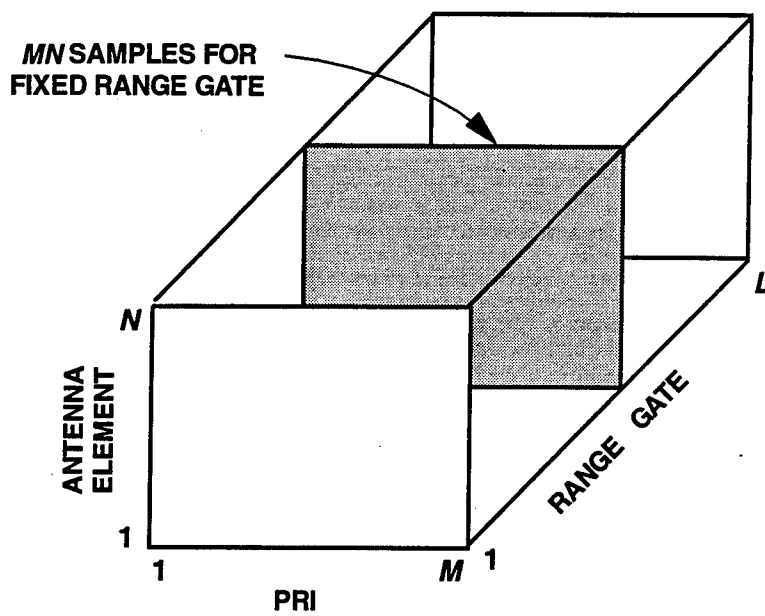


Figure 4. The radar CPI datacube.

TABLE 1
Symbols for Radar System Parameters

N	number of elements in array
M	number of pulses per CPI
L	number of range samples per PRI
$\omega_o = 2\pi f_o$	radar operating frequency
λ_o	radar operating wavelength
d	interelement spacing
T_r	pulse repetition interval (PRI)
f_r	pulse repetition frequency (PRF)
P_t	peak transmit power
T_p	transmit pulse width (uncompressed)
B	instantaneous bandwidth
$G_t(\theta, \phi)$	full array transmit power gain
$f(\theta, \phi)$	element pattern (voltage)
$g(\theta, \phi)$	element pattern (power)
L_t	system losses on transmit
L_r	system losses on receive
N_o	receiver noise power spectral density
h_a	platform height
v_a	platform velocity

For reference, Table 1 provides a list of the various radar system parameters. This matrix is represented by the shaded slice of the datacube in Figure 4. The rows of \mathbf{X}_l represent the temporal (pulse-by-pulse) samples for each antenna element. Beamforming is an operation that combines the rows of \mathbf{X}_l , while combining the columns is a temporal, or Doppler filtering operation. It is convenient to write the data for a single range gate as a the l th $MN \times 1$ vector χ_l , termed a *space-time snapshot*, by stacking the columns of \mathbf{X}_l :

$$\chi_l = \text{vec}(\mathbf{X}_l) = \begin{bmatrix} \mathbf{x}_{0,l}; & \mathbf{x}_{1,l}; & \dots; & \mathbf{x}_{M-1,l} \end{bmatrix} . \quad (4)$$

This report will work primarily with data from a single range gate, where it is understood that the components of the space-time snapshot may have a range dependency. For notational expediency, the l subscript will be omitted in much of what follows. The symbol χ will refer to a space-time snapshot at the range of interest, and \mathbf{x}_m will be the spatial snapshot for the m th PRI at this range. The function of a surveillance radar is to ascertain whether targets are present in the data. Thus, given a space-time snapshot, the signal processor must make a decision as to which of the two hypotheses is true:

$$\begin{aligned} \chi &= \chi_u & H_0: \text{Target absent} \\ \chi &= \alpha_t \mathbf{v}_t + \chi_u & H_1: \text{Target present} \end{aligned} \quad (5)$$

The vector \mathbf{v}_t is the (known) response of the system to a unit amplitude target and α_t is the (unknown) target amplitude. The component χ_u encompasses any interference or noise component of the data. Three undesired components will be considered: clutter, jamming, and thermal noise. These three components are assumed to be mutually uncorrelated. The target component is viewed as a shift in the mean of the data, so that under either hypothesis the data has covariance matrix

$$\mathbf{R}_u = E\{\chi_u \chi_u^H\} = \mathbf{R}_c + \mathbf{R}_j + \mathbf{R}_n, \quad (6)$$

where the subscript identifies the particular component. In the next sections, a model is developed for the target response and for each interference and noise component, starting from the set of radar system parameters.

2.3 Target

A target is defined as a moving point scatterer that is to be detected. The component of the space-time snapshot at the range gate corresponding to the target range R_t will be derived. The target is also described by its azimuth ϕ_t , elevation θ_t , relative velocity with respect to the radar v_t , and radar cross-section (RCS) σ_t . The derivation will begin by defining the transmitted waveform and stepping through the filtering that leads to the samples forming the space-time snapshot. The full array transmits a coherent burst of pulses

$$\tilde{s}(t) = a_t u(t) e^{j(\omega_0 t + \psi)}, \quad (7)$$

where

$$u(t) = \sum_{m=0}^{M-1} u_p(t - mT_r) \quad (8)$$

is the signal's complex envelope and $u_p(t)$ is the complex envelope of a single pulse. The transmit signal amplitude is a_t , and a random phase ψ , uniformly distributed on $[0, 2\pi)$, is also included. The pulse waveform $u_p(t)$ is of duration T_p and has unit energy,

$$\int_0^{T_p} |u_p(t)|^2 dt = 1 . \quad (9)$$

The energy in the transmitted signal is then

$$E_t = \int_0^{MT_r} |\tilde{s}(t)|^2 dt = M E_p , \quad (10)$$

where $E_p = a_t^2$ is the energy transmitted in a single pulse.

The target echo is received by each of the N elements. Ignoring relativistic effects, the target signal at the n th antenna element, $\tilde{s}_n(t)$, is given by [14]

$$\tilde{s}_n(t) = a_r u(t - \tau_n) e^{j2\pi(f_o + f_t)(t - \tau_n)} e^{j\psi} , \quad (11)$$

where a_r is the echo amplitude and

$$f_t = \frac{2v_t}{\lambda_o} . \quad (12)$$

is the target Doppler frequency. It will be convenient at times to use the *normalized Doppler frequency* defined by

$$\varpi_t = f_t T_r = \frac{f_t}{f_r} . \quad (13)$$

The target delay to the n th element, τ_n , consists of two components:

$$\tau_n = \tau_t + \tau'_n , \quad (14)$$

where $\tau_t = 2R_t/c$ is the round trip delay and

$$\tau'_n = -\frac{\hat{\mathbf{k}}(\phi_t, \theta_t) \cdot \mathbf{r}_n}{c} = -\frac{n\hat{\mathbf{k}}(\phi_t, \theta_t) \cdot \mathbf{d}}{c} = -n\frac{d}{c} \cos \theta_t \sin \phi_t \quad (15)$$

is the relative delay measured from the phase reference to the n th element. It will also be convenient to define the target **spatial frequency**

$$\vartheta_t = \frac{\hat{\mathbf{k}}(\phi_t, \theta_t) \cdot \mathbf{d}}{\lambda_o} = \frac{d}{\lambda_o} \cos \theta_t \sin \phi_t \quad (16)$$

so that the phase delay to the n th element may be expressed as

$$-\omega_o \tau'_n = n2\pi\vartheta_t \quad (17)$$

It is assumed that the transmitted waveform is narrowband; the relative delay term is insignificant within the complex envelope term of Equation (11). With this assumption,

$$\tilde{s}_n(t) = a_r e^{j\psi} e^{jn2\pi\vartheta_t} u(t - \tau_t) e^{j2\pi f_t t} e^{j2\pi f_o t} \quad (18)$$

where several of the fixed phase terms have been absorbed into the random phase ψ . This signal is now down-converted to baseband, matched filtered, and sampled with an A/D converter.

After down-conversion the n th element signal is

$$s_n(t) = \tilde{s}_n(t) e^{-j2\pi f_o t} = a_r e^{j\psi} e^{jn2\pi\vartheta_t} u(t - \tau_t) e^{j2\pi f_t t} \quad (19)$$

Each pulse of the baseband signal is matched filtered separately with receiver filter

$$h(t) = u_p^*(-t) \quad (20)$$

The matched filter output for the n th channel is the signal

$$x_n(t) = a_r e^{j\psi} e^{jn2\pi\vartheta_t} \sum_{m=0}^{M-1} e^{jm2\pi\varpi_t} \chi(t - \tau_t - mT_r, \varpi_t) \quad (21)$$

where $\chi(\tau, f)$ is the waveform ambiguity function [15]

$$\chi(\tau, f) = \int_{-\infty}^{\infty} u_p(\beta) u_p^*(\beta - \tau) e^{j2\pi f\beta} d\beta. \quad (22)$$

Because of the pulse waveform normalization,

$$\chi(0, 0) = 1. \quad (23)$$

Consider only the target range gate and let $t_m = \tau_t + mT_r$, $m = 0, \dots, M - 1$ be the sample times from each PRI at this range gate. The target samples are thus given by²

$$x_{nm} = x_n(t_m) = a_r e^{j\psi} \chi(0, f_t) e^{j2\pi\vartheta_t} e^{jm2\pi\omega_t}. \quad (24)$$

Furthermore, assume that the pulse waveform time-bandwidth product and the expected range of target Doppler frequencies are such that the waveform is insensitive to target Doppler shift, i.e., that

$$\chi(0, f) \approx 1. \quad (25)$$

It is also convenient to group the received amplitude and the random phase term of Equation (24) into a single complex random amplitude $\alpha_t = a_r e^{j\psi}$. This leaves the simple expression

$$x_{nm} = \alpha_t e^{jn2\pi\vartheta_t} e^{jm2\pi\omega_t}, \quad \begin{matrix} n = 0, \dots, N - 1 \\ m = 0, \dots, M - 1 \end{matrix}. \quad (26)$$

The target amplitude is obtained directly from the radar equation. Let ξ_t be the *single-pulse signal-to-noise ratio (SNR) for a single antenna element on receive*. This is given by

$$\xi_t = \frac{P_t T_p G_t(\theta, \phi) g(\theta, \phi) \lambda_o^2 \sigma_t}{(4\pi)^3 N_o L_s R_t^4}, \quad (27)$$

where the radar system parameters were given in Table 1. The *target power* is then given by

$$E \{ |\alpha_t|^2 \} = \sigma^2 \xi_t, \quad (28)$$

²Neglect range straddling losses and also assume that there are no target range ambiguities.

where σ^2 is the thermal noise power per element. The target amplitude is then $a_r = \sqrt{\xi_t \sigma^2}$. This target model is easily generalized to random amplitudes as well. For example, if it is assumed that the samples are zero mean complex Gaussian random variables, the target SNR would be exponentially distributed with mean ξ_t and the target amplitude would be correspondingly Rayleigh distributed.

Examination of Equation (26) shows that one exponential term depends on the spatial index n and the other depends on the temporal index m . The spatial snapshot for the m th pulse, $\mathbf{x}_m = [x_{0,m}; x_{1,m}; \dots; x_{N-1,m}]$, can be written as

$$\mathbf{x}_m = [x_{0,m}; x_{1,m}; \dots; x_{N-1,m}] = \alpha e^{jm2\pi\varpi_t} \mathbf{a}(\vartheta_t), \quad (29)$$

where the $N \times 1$ spatial steering vector $\mathbf{a}(\vartheta)$ is defined to be

$$\mathbf{a}(\phi, \theta) = \left[1; e^{j\frac{2\pi d}{\lambda_0} \cos \theta \sin \phi}; \dots; e^{j(N-1)\frac{2\pi d}{\lambda_0} \cos \theta \sin \phi} \right] \quad (30)$$

or

$$\mathbf{a}(\vartheta) = \left[1; e^{j2\pi\vartheta}; \dots; e^{j(N-1)2\pi\vartheta} \right]. \quad (31)$$

The spatial steering vector assumes a Vandermonde form because of the uniform linear array geometry and the assumption of identical element patterns. If the element patterns are not identical, they must be included in the steering vector definition.

The target data is assembled in the form of a space-time snapshot

$$\begin{aligned} \chi_t &= \alpha_t \left[\mathbf{a}(\vartheta_t); e^{j2\pi\varpi_t} \mathbf{a}(\vartheta_t); \dots; e^{j(M-1)2\pi\varpi_t} \mathbf{a}(\vartheta_t) \right] \\ &= \alpha_t \mathbf{b}(\varpi_t) \otimes \mathbf{a}(\vartheta_t), \end{aligned} \quad (32)$$

where

$$\mathbf{b}(\varpi) = \left[1; e^{j2\pi\varpi}; \dots; e^{j(M-1)2\pi\varpi} \right]. \quad (33)$$

is an $M \times 1$ temporal steering vector. It is Vandermonde also, because the waveform is a uniform PRF and the target velocity is constant.

The $MN \times 1$ space-time steering vector,

$$\mathbf{v}(\vartheta, \varpi) = \mathbf{b}(\varpi) \otimes \mathbf{a}(\vartheta) . \quad (34)$$

is defined to be the response of a target at spatial frequency ϑ and normalized Doppler ϖ . If a target is in the data, it contributes a term

$$\chi_t = \alpha_t \mathbf{v}_t , \quad (35)$$

where $\mathbf{v}_t = \mathbf{v}(\vartheta_t, \varpi_t)$ may also be called the *target steering vector*.

2.4 Noise

The first undesired signal that a potential target must contend with is noise. Assume that the only noise source is internally generated receiver noise, which is always present on each channel. Because each element has its own receiver, assume that the noise processes on each element are mutually uncorrelated. Furthermore, assume that the instantaneous bandwidth is large compared with the PRF.³ Therefore, noise samples on a single element taken at time instants separated by a nonzero multiple of the PRI are temporally uncorrelated. Let x_{nm} be the noise sample on the n th element for the m th PRI. The first assumption above is a statement of spatial noise correlation

$$E\{x_{n_1, m} x_{n_2, m}^*\} = \sigma^2 \delta_{n_1 - n_2} , \quad (36)$$

where

$$\delta_m = \begin{cases} 1 & , m = 0 \\ 0 & , m \neq 0 \end{cases} \quad (37)$$

is the Kronecker delta and σ^2 is the noise power per element. The second assumption above leads to the temporal noise correlation

$$E\{x_{n, m_1} x_{n, m_2}^*\} = \sigma^2 \delta_{m_1 - m_2} . \quad (38)$$

³This is not inconsistent with the narrowband assumption made previously, merely a statement that for the radar systems under consideration, $f_r \ll B \ll f_o$.

Equations (36) and (38) lead to the noise component of the space-time covariance matrix being the scaled identity matrix

$$\mathbf{R}_n = E \left\{ \mathbf{x}_n \mathbf{x}_n^H \right\} = \sigma^2 \mathbf{I}_M \otimes \mathbf{I}_N = \sigma^2 \mathbf{I}_{MN} . \quad (39)$$

In terms of the radar system parameters, the noise power $\sigma^2 = N_o B$. For purposes of this analysis, a convenient normalization is to set $\sigma^2 = 1$ so that all signal levels can be referenced by their SNR per element and pulse.

The noise model above is realistic only when the dominant source of noise is internally generated receiver amplifier noise. If sky noise is a major contributor, a spatial correlation may need to be introduced into the model above. It will be shown later that preprocessing prior to adaptation will also introduce noise correlation between the various signals to be adaptively combined.

2.5 Jamming

In this section expressions are derived for the jamming contribution to a space-time snapshot vector and its covariance matrix. Only barrage noise jamming that originates from land-based or airborne platforms at long range from the radar will be considered. The jamming energy is assumed to fill the radar's instantaneous bandwidth. The narrowband assumption that a signal's propagation time across the array is small relative to $1/B$, i.e. there is no signal decorrelation across the array, is again invoked. Conversely, a radar PRF is assumed that is significantly less than the instantaneous bandwidth so that the jamming decorrelates from one pulse to the next. In other words, the jamming is spatially correlated from element to element and temporally uncorrelated from pulse to pulse. Thus, jamming looks like thermal noise temporally, but like a point target or a discrete clutter source in the spatial domain.

Consider first a single jammer located at elevation θ_j , azimuth ϕ_j , and range R_j . Let S_j denote the jammer's effective radiated power density (in W/Hz) in the direction of the radar. The jammer power spectral density received by one array element, J_o , is then obtained from the range equation [17] as

$$J_o = \frac{S_j g(\theta_j, \phi_j) \lambda_o^2}{(4\pi)^2 R_j^2 L_r} . \quad (40)$$

The received jammer-to-noise ratio (JNR) at an element is then given by

$$\xi_j = \frac{J_o}{N_o}, \quad (41)$$

where N_o is the receiver noise power spectral density. The jamming component of the spatial snapshot for the m th PRI is then

$$\mathbf{x}_m = \alpha_m \mathbf{a}_j, \quad (42)$$

where α_m is the jammer amplitude for the m th PRI and $\mathbf{a}_j = \mathbf{a}(\phi_j, \theta_j)$ is the jammer steering vector. The jammer space-time snapshot may be written as

$$\boldsymbol{\chi}_j = \boldsymbol{\alpha}_j \otimes \mathbf{a}_j, \quad (43)$$

where $\boldsymbol{\alpha}_j = [\alpha_0; \alpha_1; \dots; \alpha_{M-1}]$ is a random vector containing the jammer amplitudes. The jammer samples from different pulses are uncorrelated. Further assume for simplicity that the jamming signal (aspect and power spectral density) is stationary over a CPI. Thus,

$$E\{\alpha_{m_1} \alpha_{m_2}^*\} = \sigma^2 \xi_j \delta_{m_1 - m_2} \quad \text{or} \quad E\{\boldsymbol{\alpha}_j \boldsymbol{\alpha}_j^H\} = \sigma^2 \xi_j \mathbf{I}_M. \quad (44)$$

The jammer space-time covariance matrix is then

$$\begin{aligned} \mathbf{R}_j &= E\{\boldsymbol{\chi}_j \boldsymbol{\chi}_j^H\} = \sigma^2 \xi_j \mathbf{I}_M \otimes \mathbf{a}_j \mathbf{a}_j^H = \mathbf{I}_M \otimes \sigma^2 \xi_j \mathbf{a}_j \mathbf{a}_j^H \\ &= \mathbf{I}_M \otimes \boldsymbol{\Phi}_j, \end{aligned} \quad (45)$$

where $\boldsymbol{\Phi}_j$ is the jammer spatial covariance matrix

$$\boldsymbol{\Phi}_j = E\{\mathbf{x}_m \mathbf{x}_m^H\} = \sigma^2 \xi_j \mathbf{a}_j \mathbf{a}_j^H. \quad (46)$$

The extension to multiple jamming signals is straightforward. Let J be the number of jamming sources, and let θ_k, ϕ_k and ξ_k be the elevation, azimuth, and JNR of the k th jammer, respectively, for $k = 1, \dots, J$. The result is again given by Equation (45), but the spatial covariance matrix is now given by

$$\boldsymbol{\Phi}_j = \mathbf{A}_j \boldsymbol{\Xi}_j \mathbf{A}_j^H, \quad (47)$$

where

$$\mathbf{A}_j = \begin{bmatrix} \mathbf{a}(\theta_1, \phi_1), & \mathbf{a}(\theta_2, \phi_2), & \dots, & \mathbf{a}(\theta_J, \phi_J) \end{bmatrix} \quad (48)$$

is an $N \times J$ matrix of the jammer spatial steering vectors and $\mathbf{\Xi}_j$ is the $J \times J$ jammer source covariance matrix. The jammer space-time covariance matrix from Equation (45) is block diagonal. Off-diagonal $N \times N$ blocks are zero because jamming samples from different PRIs are uncorrelated. The stationary assumption results in the blocks along the diagonal all being equal to a single spatial covariance matrix. Let the rank of the jammer spatial covariance matrix be

$$\rho_j = \text{rank}(\mathbf{\Phi}_j). \quad (49)$$

Assume that the J jammer steering vectors are linearly independent. If no two jamming signals are perfectly coherent, then $\rho_j = J$, and the jammer space-time covariance matrix has rank

$$r_j = \text{rank}(\mathbf{R}_j) = M \text{rank}(\mathbf{\Phi}_j) = MJ. \quad (50)$$

The low rank nature of the jamming covariance matrix will be helpful in the design of efficient adaptive processing architectures.

2.6 Clutter

Radar clutter is generically defined as the echoes from any scatterers deemed to be not of tactical significance. For an airborne surveillance radar, the Earth's surface is the major source of clutter and is the only type of clutter to be considered in this report. Of the various sources of interference, clutter is the most complicated because it is distributed in both angle and range and is spread in Doppler frequency due to the platform motion. In this section, a model is developed for the ground clutter component of the space-time snapshot for a given range, and the properties of the clutter space-time covariance matrix are considered. The clutter in many cases is shown to be low rank. The effects of velocity misalignment (due to aircraft crab, for example) and intrinsic clutter motion are also modeled.

2.6.1 General Clutter Model

The return from a discrete ground clutter source has the same form as a target echo defined in Section 2.2. Since the ground is assumed to have zero inherent velocity, the relative velocity of a ground clutter source depends only on its aspect with respect to the radar and on the platform velocity. Unlike a target, ground clutter is distributed in range; it exists over a region extending from the platform altitude to the radar horizon. Ground clutter also exists over all azimuths and a region in elevation angle bounded by the horizon elevation.

Consider the clutter present at range R_c from the radar, as shown in Figure 5. Assume a spherical earth with a $4/3$ effective radius [16]. The elevation angle θ_c to this clutter ring is

$$\theta_c = \theta_c(R_c) = -\sin^{-1} \left(\frac{R_c^2 + h_a(h_a + 2a_e)}{2R_c(a_e + h_a)} \right) \approx -\sin^{-1} \left(\frac{R_c}{2a_e} + \frac{h_a}{2a_e} \right), \quad (51)$$

where $a_e = 4/3r_e$ is the effective earth radius. The elevation angle is measured with respect to the horizontal at the antenna. The grazing angle ψ_c is formed by the ray from antenna to clutter patch and the surface tangent at the clutter patch,⁴

$$\psi_c = \psi_c(R_c) = -\sin^{-1} \left(\frac{R_c^2 - h_a(h_a + 2a_e)}{2R_c a_e} \right) \approx -\sin^{-1} \left(\frac{R_c}{2a_e} - \frac{h_a}{2a_e} \right). \quad (52)$$

The radar horizon range

$$R_h = \sqrt{2a_e h_a + h_a^2} \approx \sqrt{2a_e h_a} \quad (53)$$

is the range for which the grazing angle is zero.

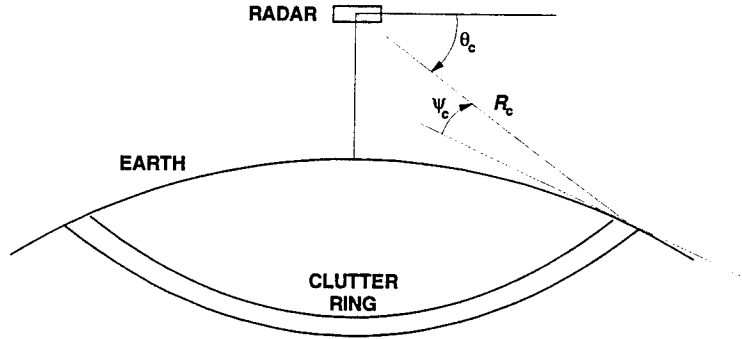


Figure 5. A ring of ground clutter for a fixed range.

⁴For a flat-earth model, the elevation and grazing angles are the same and equal to $-\sin^{-1}(h_a/R_c)$.

Let $R_u = c/2f_r$ be the radar's unambiguous range. Consider the clutter return for the l th range gate, which corresponds to true range R_c where $0 < R_c \leq R_u$. If the unambiguous range is greater than the horizon range, $R_u > R_h$, the clutter is said to be unambiguous in range. In this case the clutter component of the space-time snapshot consists of clutter from at most one range. If the radar horizon is larger than the unambiguous range, some or all range gates will have clutter contributions from multiple ranges. The clutter in this case is said to be range-ambiguous. Let $R_i = R_c + (i - 1)R_u$ be the i th ambiguous range corresponding to the range gate of interest. Each ambiguous range has corresponding elevation and grazing angles $\theta_i = \theta_c(R_i)$ and $\psi_i = \psi_c(R_i)$ found from Equation (51) and Equation (52). The clutter component consists of the superposition of returns from all ambiguous ranges within the radar horizon. Denote the number of range ambiguities by N_r .

As an approximation to a continuous field of clutter, the clutter return from each ambiguous range will be modeled as the superposition of a large number N_c of independent clutter sources that are evenly distributed in azimuth about the radar. The location of the ik th clutter patch is described by its azimuth ϕ_k and ambiguous range R_i (or elevation θ_i). The corresponding spatial frequency is

$$\vartheta_{ik} = \frac{\hat{\mathbf{k}}(\theta_i, \phi_k) \cdot \mathbf{d}}{\lambda_o} = \frac{d}{\lambda} \cos \theta_i \sin \phi_k . \quad (54)$$

The normalized Doppler frequency of the ik th patch will be denoted by ϖ_{ik} . The clutter component of the space-time snapshot is then given by

$$\chi_c = \sum_{i=1}^{N_r} \sum_{k=1}^{N_c} \alpha_{ik} \mathbf{v}(\vartheta_{ik}, \varpi_{ik}) , \quad (55)$$

where α_{ik} is the random amplitude from the ik th clutter patch and $\mathbf{v}_{ik} = \mathbf{v}(\vartheta_{ik}, \varpi_{ik})$.

The power of each clutter contribution is obtained from the radar equation for area clutter [17]. Each patch represents an effective area bounded in azimuth by the granularity of the angle sampling, $\Delta\phi = 2\pi/N_c$, and in range by the radar's range resolution $\Delta R = c/2B$. The effective RCS of the ik th clutter patch is therefore

$$\sigma_{ik} = \sigma_o(\phi_k, \theta_i) \times \text{Patch Area} = \sigma_o(\phi_k, \theta_i) R_i \Delta\phi \Delta R \sec \psi_i , \quad (56)$$

where $\sigma_o(\phi_k, \theta_i)$ is the area reflectivity of the ground at the location of the ik th patch. The dependence on position is necessary to account for clutter nonhomogeneities in range and angle. Various models for the clutter reflectivity have been proposed, depending on the terrain type, radar frequency, polarization, etc. For the simulations in this report, the *constant gamma* model [17],

$$\sigma_o = \gamma \sin \psi_c , \quad (57)$$

is used, where γ is the terrain-dependent parameter. With the definitions above, the contribution from the ik th clutter patch has a clutter-to-noise ratio (CNR) given by

$$\xi_{ik} = \frac{P_t T_p G_t(\phi_k, \theta_i) g(\phi_k, \theta_i) \lambda_o^2 \sigma_{ik}}{(4\pi)^3 N_o L_s R_i^4} . \quad (58)$$

The clutter amplitudes satisfy $E\{|\alpha_{ik}|^2\} = \sigma^2 \xi_{ik}$. Assume that returns from different clutter patches are uncorrelated:

$$E\{\alpha_{ik} \alpha_{jl}^*\} = \sigma^2 \xi_{ik} \delta_{i-j} \delta_{k-l} . \quad (59)$$

The clutter space-time covariance matrix follows directly from Equations (55) and (59):

$$\mathbf{R}_c = E\{\chi_c \chi_c^H\} = \sigma^2 \sum_{i=1}^{N_r} \sum_{k=1}^{N_c} \xi_{ik} \mathbf{v}_{ik} \mathbf{v}_{ik}^H . \quad (60)$$

Alternatively, Equation (60) can be expressed as

$$\mathbf{R}_c = \sigma^2 \sum_{i=1}^{N_r} \sum_{k=1}^{N_c} \xi_{ik} (\mathbf{b}_{ik} \mathbf{b}_{ik}^H) \otimes (\mathbf{a}_{ik} \mathbf{a}_{ik}^H) , \quad (61)$$

where $\mathbf{b}_{ik} = \mathbf{b}(\varpi_{ik})$ and $\mathbf{a}_{ik} = \mathbf{a}(\vartheta_{ik})$. Each scatterer contributes a term that is the Kronecker product of a temporal covariance matrix with a spatial covariance matrix. These two components are coupled because the clutter Doppler is a function of angle. The matrix \mathbf{R}_c is an $M \times M$ block matrix, where each block is an $N \times N$ cross-covariance of the spatial snapshots from two PRIs. The p th row of \mathbf{R}_c corresponds to element $n(p)$ and pulse $m(p)$, where

$$n(p) = \text{mod}(p, N) , \quad m(p) = \text{floor}\left(\frac{p}{N}\right) . \quad (62)$$

The pq th element of \mathbf{R}_c can be written, from Equation (61), as

$$[\mathbf{R}_c]_{pq} = \sum_{i=1}^{N_r} \sum_{k=1}^{N_c} \sigma^2 \xi_{ik} \exp \{j [(n(p) - n(q)) 2\pi \vartheta_{ik} + (m(p) - m(q)) 2\pi \varpi_{ik}]\} . \quad (63)$$

Since Equation (63) depends only on the difference between the pulse indices, the matrix \mathbf{R}_c is block-Toeplitz. Moreover, Equation (63) depends only on difference $n(p) - n(q)$ because the antenna is an ideal uniform linear array and the clutter patches were assumed mutually uncorrelated; each $N \times N$ block is itself Toeplitz. A matrix of this type is said to be Toeplitz-block-Toeplitz [19].

The expressions above apply to the general case of range-ambiguous clutter. **The range-unambiguous case will be mainly considered in the remainder of this report, and the i subscript denoting ambiguous range will be dropped.** The clutter covariance matrix can also be expressed compactly as

$$\mathbf{R}_c = \mathbf{V}_c \mathbf{\Xi}_c \mathbf{V}_c^H , \quad (64)$$

where

$$\mathbf{V}_c = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_c}] \quad (65)$$

is an $MN \times N_c$ matrix of clutter space-time steering vectors and

$$\mathbf{\Xi}_c = \sigma^2 \text{diag}([\xi_1, \dots, \xi_{N_c}]) \quad (66)$$

contains the clutter power distribution.

2.6.2 Clutter Ridges

The aircraft platform motion induces a very special structure to the clutter due to the dependence of the Doppler frequency on angle. Consider again a single clutter patch located at azimuth ϕ_c and range R_c (or elevation θ) and let $\hat{\mathbf{k}}(\theta_c, \phi_c)$ be a unit vector pointing from the platform to this patch. The clutter spatial frequency is

$$\vartheta_c = \frac{\hat{\mathbf{k}}(\theta_c, \phi_c) \cdot \mathbf{d}}{\lambda_o} = \frac{d}{\lambda_o} \cos \theta_c \sin \phi_c . \quad (67)$$

The Doppler frequency of the echo from this patch is defined by

$$f_c(\theta_c, \phi_c) = \frac{2\hat{\mathbf{k}}(\theta_c, \phi_c) \cdot \mathbf{v}_a}{\lambda_o} . \quad (68)$$

For this section, assume that the velocity vector is aligned with the array axis, as it would be with a side-mounted antenna array and no crab. Therefore, $\mathbf{v}_a = v_a \hat{\mathbf{x}}$ and the clutter Doppler frequency becomes

$$f_c(\theta_c, \phi_c) = \frac{2v_a}{\lambda_o} \cos \theta_c \sin \phi_c . \quad (69)$$

In terms of normalized Doppler and spatial frequency,

$$\varpi_c = f_c T_r = \left(\frac{2v_a T_r}{d} \right) \vartheta_c . \quad (70)$$

The clutter Doppler frequency is a linear function of $\sin \phi$ and the normalized Doppler is linear in spatial frequency. With normalized coordinates the slope of the clutter line is

$$\beta = \frac{2v_a T_r}{d} , \quad (71)$$

which represents the number of half-interelement spacings traversed by the platform during one PRI. For half-wavelength interelement spacing, $\beta = 4v_a/\lambda_o f_r$ is equivalently the number of times the clutter Doppler spectrum aliases into the unambiguous Doppler space.

Equation (68) or (70) defines the locus in an angle-Doppler space where clutter is present. This locus is referred to as the “clutter ridge,” shown in Figure 1. Figure 6 shows two views of the clutter ridge for $\beta = 1$, which corresponds to the clutter exactly filling the Doppler space once.

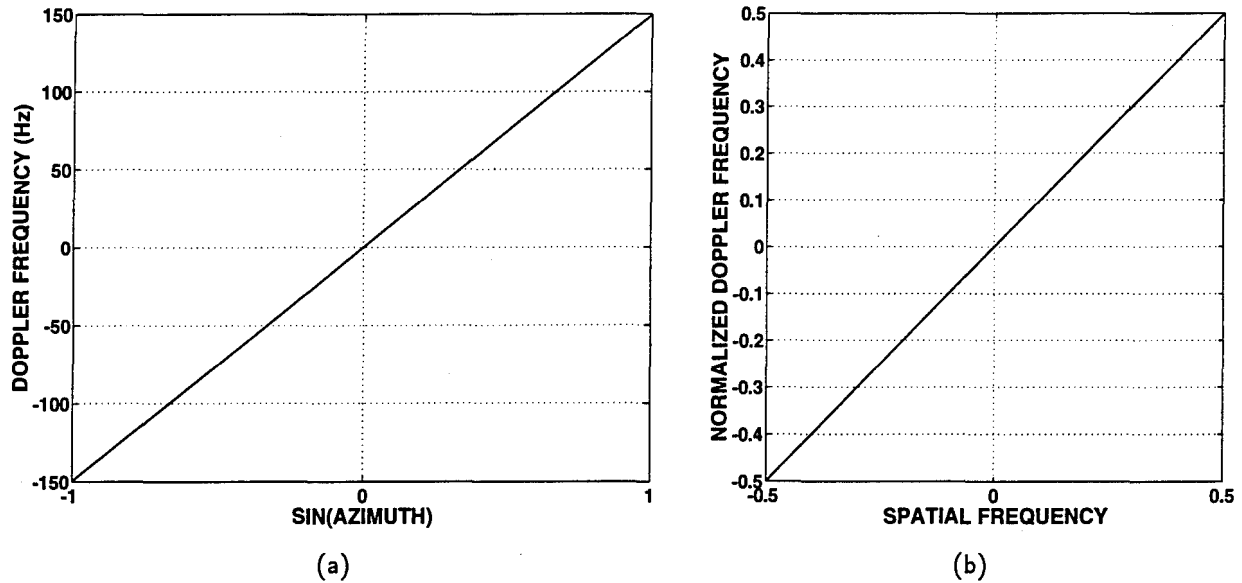
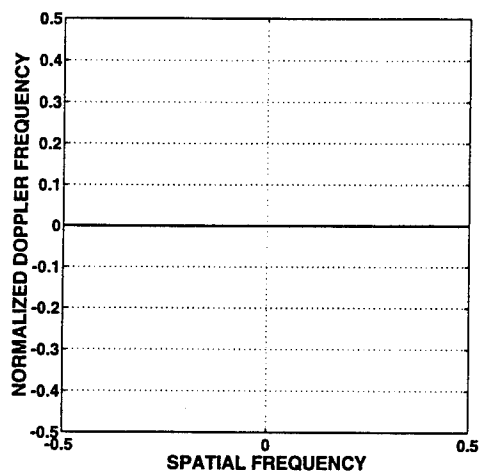
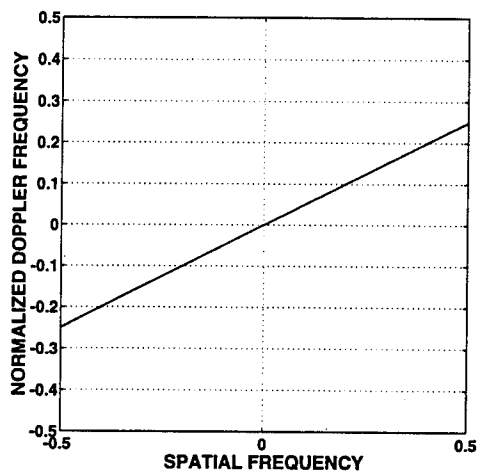


Figure 6. The $\beta = 1$ clutter ridge. The PRF is 300 Hz. (a) Doppler frequency vs. $\sin \phi$. (b) Normalized Doppler vs. spatial frequency.

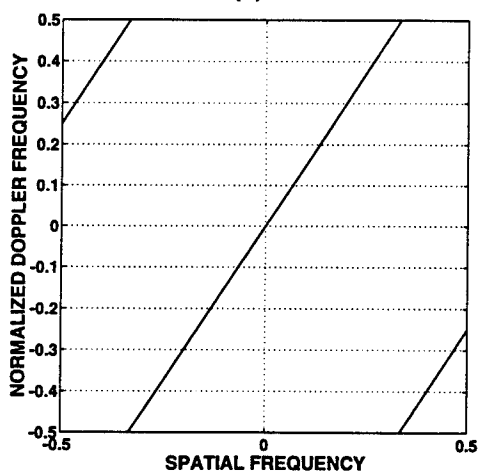
In general, the clutter ridge may span a portion of the Doppler space, or the whole Doppler space, depending on the platform velocity, the operating wavelength, and the radar PRF. Figure 7 shows four examples. In these examples the PRF is fixed and the platform velocity varies. The default case of a stationary platform and zero Doppler clutter is shown in Figure 7(a). If $f_r \geq 4v_a/\lambda_o$ (or $\beta \leq 1$), the clutter is said to be unambiguous in Doppler. The $\beta = 1$ case is shown in Figure 6; Figure 7(b) shows the $\beta = 0.5$ ridge. If the clutter is unambiguous in Doppler, there may be a clutter-free region in Doppler, and there is at most one angle where the clutter has the same Doppler as a target. When $f_r < 4v_a/\lambda_o$ or $\beta > 1$, the clutter is said to be Doppler-ambiguous. In this case the clutter spectrum extends over a region larger than the PRF and folds over (aliases) into the observable Doppler space. Figures 7(c) and (d) show Doppler-ambiguous clutter ridges. In this case there may be multiple angles at which sidelobe clutter has the same Doppler as a target. Furthermore, as β increases, the mainlobe clutter occupies a larger portion of the Doppler space. The more Doppler-ambiguous the clutter, the more difficult it will be to suppress. For low-PRF radars operating at UHF, the case of Figure 7(d) is typical.



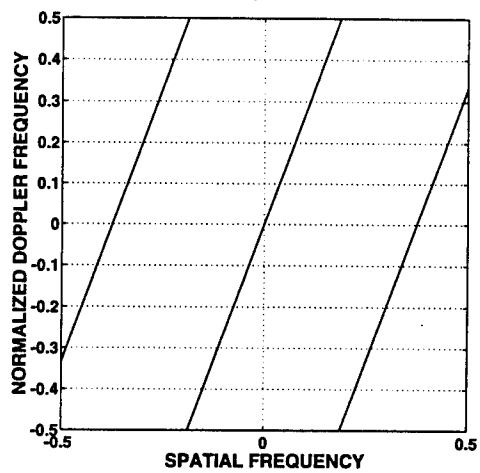
(a)



(b)



(c)



(d)

Figure 7. Clutter loci for different velocities. (a) Stationary platform, $v_a = 0$. (b) Doppler unambiguous, $\beta = 0.5$. (c) Doppler ambiguous, $\beta = 1.5$. (d) Doppler ambiguous, $\beta = 2.67$.

The space-time steering vector to a clutter patch can therefore be written as a function of angle (or spatial frequency) alone:

$$\mathbf{v}_c(\phi_c) = \mathbf{v}(\phi_c, f_c(\phi_c)) \text{ or } \mathbf{v}_c(\vartheta_c) = \mathbf{v}(\vartheta_c, \beta \vartheta_c). \quad (72)$$

By substituting the clutter ridge equation into Equation (61), the clutter covariance matrix elements may be expressed as

$$[\mathbf{R}_c]_{pq} = \sum_{k=1}^{N_c} \sigma^2 \xi_k \exp \{j2\pi \vartheta_k [(n(p) - n(q)) + \beta (m(p) - m(q))]\} . \quad (73)$$

It is sometimes useful to express the clutter covariance matrix, Equation (60), in integral form,

$$\mathbf{R}_c = \int_{-\pi}^{\pi} s_c(\phi) \mathbf{v}_c(\phi) \mathbf{v}_c^H(\phi) d\phi , \quad (74)$$

where $s_c(\phi)$ is the clutter power spectral density obtained from the radar equation. In the simple case of clutter power uniformly distributed in azimuth, $s_c(\phi) = 1/2\pi$, it can be shown that the elements of the clutter covariance matrix have closed form

$$[\mathbf{R}_c]_{pq} = J_0 \left(\frac{2\pi d \cos \theta}{\lambda} [n(p) - n(q) + \beta (m(p) - m(q))] \right) , \quad (75)$$

where $J_0(z)$ is the zeroth order Bessel function.

2.6.3 Rank of the Clutter Covariance Matrix

The rank of the clutter covariance matrix will now be considered, as it is an indicator of both the severity of the clutter scenario and the number of degrees of freedom required to produce effective cancellation. The special structure embodied in the clutter ridge suggests that the clutter is of low rank. Klemm [5] has hypothesized that the clutter rank is approximately equal to $N + M$. Based upon extensive computer simulation, Brennan and Staudaher [18] developed the following rule regarding the clutter rank.

Theorem 1 (Brennan's Rule): *The rank of the clutter covariance matrix is approximately given by*

$$r_c \approx \lfloor N + (M - 1)\beta \rfloor , \quad (76)$$

where β , given by Equation (71), is the number of one-half interelement spacings traversed by the platform during one PRI. The brackets $\lfloor \cdot \rfloor$ indicate rounding to the nearest integer.

The rounding operation in Equation (76) is required as β is not necessarily an integer. When β is an integer, Equation (76) can be replaced by an equality, and a rigorous proof for this case is provided in Appendix A.

To see why Brennan's rule holds, recall that each clutter patch contributes a term $\mathbf{v}(\vartheta_c, \beta\vartheta_c)$ to the space-time snapshot. Examination of the clutter space-time steering vector shows that the phase of the clutter signal on the n th element and m th pulse is given by

$$\varphi_{nm} = 2\pi(n + m\beta)\vartheta_c \quad (77)$$

$$= \frac{2\pi}{\lambda} [(n + m\beta)d] \cos \theta \sin \phi . \quad (78)$$

The clutter Doppler makes the spatial snapshot for the m th pulse appear as though it is received by an array whose position has moved by $m\beta d$. Therefore, the *effective* position of the n th element for the m th pulse is

$$d_{nm} = (n + m\beta)d . \quad (79)$$

Figure 8 illustrates this concept for the case $N = 4$, $M = 3$, and $\beta = 1$. Note that element #1 on pulse #0 is effectively at the same position as element #0 on pulse #1. Clutter observations are, in effect, repeated by different elements on different pulses as the platform moves during a CPI. Therefore, the number of independent clutter observations is less than the total snapshot dimension MN . Only independent observations contribute to the clutter rank. For the case in Figure 8, there are exactly $N + M - 1$ distinct observations, which is equal to the rank predicted by Brennan's rule. From Equation (76), the clutter rank increases linearly with β or equivalently with the amount of Doppler ambiguity. Moreover, because the above argument did not rely on a range-unambiguous assumption, Brennan's rule holds for range-ambiguous clutter as well.

To further illustrate Brennan's rule, consider an example radar system whose parameters are given in Table 2. The radar operates at UHF and utilizes an 18×4 planar array antenna. The columns of this antenna are beamformed to form the elements (or channels) of the adaptive array. The element pattern is a cosine in azimuth, where in the backlobe region the pattern is scaled by an assumed backlobe level b_e ,

$$f(\theta, \phi) = \begin{cases} \cos \phi, & -90^\circ \leq \phi < 90^\circ \\ b_e \cos \phi, & 90^\circ \leq \phi < 270^\circ \end{cases} \quad (80)$$

This pattern is shown in Figure 9 for a nominal backlobe level of -30 dB. The radar PRF is 300 Hz and 18 pulses are transmitted within a coherent processing interval. A uniform taper on the transmit pattern is assumed. Assume the platform altitude is 9 km and the range of interest is 130 km. A $\gamma = -3$ dB is chosen for the reflectivity to model heavy land clutter. The clutter is divided among $N_c = 361$ clutter patches equally distributed in azimuth about the platform. Figure 10 shows the resultant CNR from each patch. The total CNR per element per pulse is 47 dB.

For this scenario the parameter β is varied by changing the platform velocity while keeping the PRF fixed. For each case the spectral decomposition

$$\mathbf{R}_c = \mathbf{E}_c \mathbf{\Lambda}_c \mathbf{E}_c^H, \quad (81)$$

is computed where $\mathbf{\Lambda}_c = \text{diag}(\lambda_1, \dots, \lambda_{MN})$ is a diagonal matrix of the eigenvalues $\{\lambda_i\}$. Brennan's rule says that only the first r_c eigenvalues are nonzero. Figure 11 shows the results. The value obtained from Brennan's rule is indicated by the dashed lines. When β is an integer, the eigenspectrum exhibits a sharp cutoff, as the covariance matrix is singular. Note also that the rank becomes larger as the platform velocity is increased. The shape of the spectrum is determined by the transmit pattern and the clutter power distribution in angle.

For noninteger β , the eigenspectrum exhibits a gradual decrease as opposed to the sharp cutoff present in a singular matrix. In these cases the clutter covariance is no longer singular, although many of the eigenvalues are extremely small. Brennan's rule no longer provides the exact rank, but, from Figure 11 it does accurately predict the knee in the eigenspectrum. In practical situations only a few of the eigenvalues beyond the knee may be significant, as most of them will be buried significantly below the thermal noise (at 0 dB level in Figure 11).

TABLE 2
Radar System and Antenna Array Parameters for Example Scenario

Radar System Parameter	Value
Operating Frequency	450 MHz
Peak Power	200 kW
Duty Factor	6 %
Transmit Gain	22 dB
Column Receive Gain	10 dB
Instantaneous Bandwidth	4 MHz
Noise Figure	3 dB
System Losses	4 dB
PRF	300 Hz
Number of Pulses / CPI	18
Pulse Width	200 μ s

Antenna Array Parameter	Value
Number of Elements x	18
Number of Elements z	4
Element Pattern	Cosine
Element Gain	4 dB
Transmit Taper	Uniform

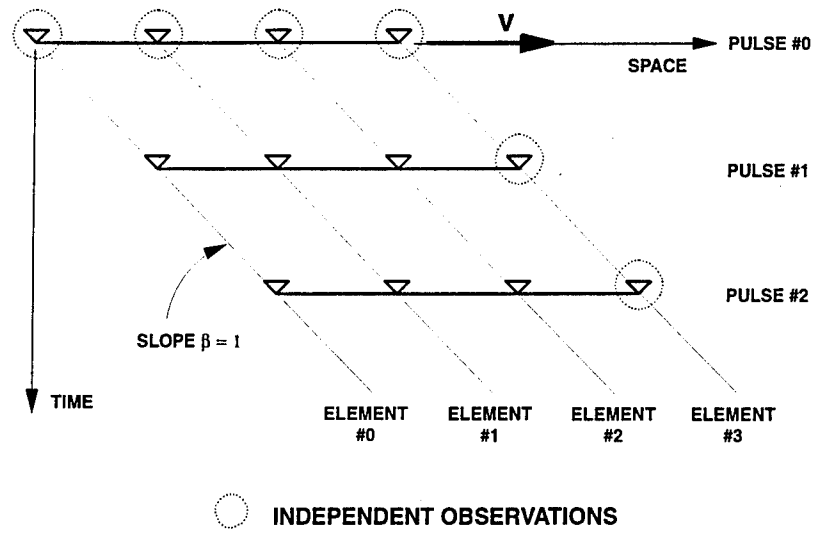


Figure 8. Effective array position for successive pulses of a CPI. $N = 4$, $M = 3$, $\beta = 1$. Clutter observations are repeated by different elements on different pulses.

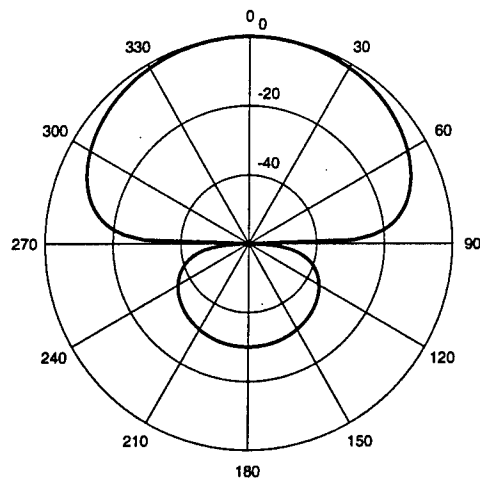


Figure 9. The element pattern. A -30 -dB backlobe level is assumed.

Thus, for noninteger β , Brennan's rule may slightly underestimate the clutter rank.

Brennan's rule is extremely useful and provides much insight into the characteristics of the clutter seen by an airborne platform. It will help motivate the processing algorithms in the next chapters. In practice, however, there are several other phenomena that tend to increase the clutter rank significantly beyond that predicted by Equation (76). Among these are the misalignment of the velocity vector and the array axis (e.g., due to aircraft crab), intrinsic clutter motion, and element pattern mismatches. The first two of these are considered in the next subsections.

2.6.4 Velocity Misalignment

The analysis leading to Brennan's rule assumed that the platform velocity vector \mathbf{v}_a was perfectly aligned with the antenna array axis vector \mathbf{d} . In practice, this may not be true for several reasons. With a side-mounted array antenna, the direction of motion may be slightly offset from the array axis due to aircraft crab. In a radar that utilizes a rotating antenna, misalignment is nearly always present; the array axis is only aligned with the velocity vector for the two instants during each scan when the array normal is perpendicular to the velocity vector. In forward-looking airborne interceptor radars, the antenna is typically mounted in the aircraft nose; in this case the array axis is approximately normal to the velocity vector.

Expressions for clutter Doppler and spatial frequency in terms of arbitrary \mathbf{v}_a and \mathbf{d} were given in Equations (68) and (67). In general, two angles are required to specify the direction of each. Here it is assumed that both the velocity vector and array axis lie in the horizontal plane but differ in direction by an angle ϕ_a called the *misalignment angle*,⁵ as shown in Figure 12. The presence of misalignment, regardless of its source, is potentially troublesome if it is combined with significant clutter from the antenna backlobes. To see this, consider the single clutter patch at azimuth ϕ in Figure 12. Due to the inherent symmetry of a linear array, this patch and its corresponding backlobe patch at $180^\circ - \phi$ produce the same interelement phase shift, i.e., they appear to be at the same azimuth. However, if $\phi_a \neq 0$, these two clutter patches have different Doppler frequencies and may require more adaptive degrees of freedom to cancel.

⁵This assumption is not restrictive for surveillance platforms; for other types of radars where diving scenarios are typical, the more general expressions will be necessary.

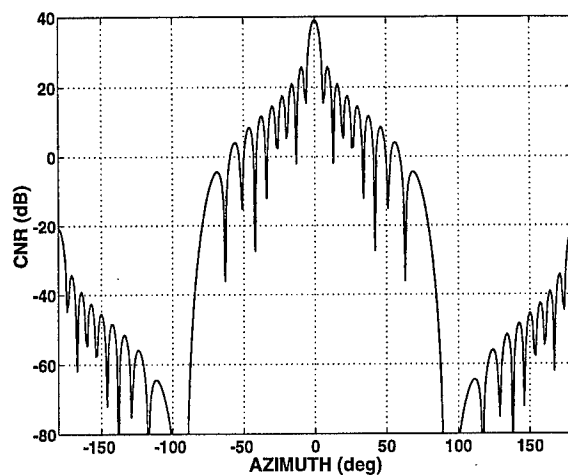


Figure 10. Example scenario: received CNR per column as a function of azimuth.

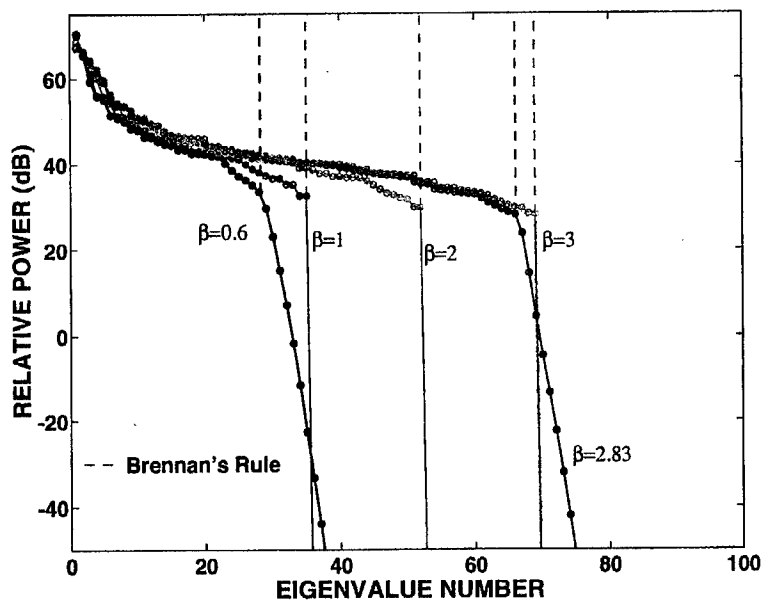


Figure 11. Illustrating Brennan's rule: clutter eigenspectra for the example radar system with different platform velocities.

Velocity misalignment produces a clutter ridge that, instead of being a line in the variables $\sin \phi$ and Doppler frequency, is an ellipse. Note from the figure or from Equation (68) that the clutter Doppler frequency with misalignment is

$$f_c(\phi) = \frac{2v_a}{\lambda_o} \cos \theta \sin(\phi + \phi_a) . \quad (82)$$

It is again convenient to work with normalized frequencies. The clutter spatial frequency and normalized Doppler are given by

$$\vartheta = \eta \sin \phi , \quad \varpi = \beta \eta \sin(\phi + \phi_a) , \quad (83)$$

where $\eta = d \cos \theta / \lambda_o$. Manipulation of Equation (83) leads to the standard elliptical form

$$\beta^2 \vartheta^2 - (2\beta \cos \phi_a) \vartheta \varpi + \varpi^2 - \beta^2 \eta^2 \sin^2 \phi_a = 0 . \quad (84)$$

The presence of the $\vartheta \varpi$ term above indicates that the major and minor axes of the clutter ellipse are rotated from the ϑ and ϖ axes by an angle ψ_a that satisfies [21]

$$\cot 2\psi_a = \frac{1 - \beta^2}{2\beta \cos \phi_a} , \quad (85)$$

where the rotation angle is measured clockwise from the positive ϖ -axis. Let $2a$ and $2b$ be the lengths of the major and minor axes. A coordinate rotation through angle ψ_a applied to Equation (84) yields the results:

$$a = \beta \eta \sin \phi_a \left(\frac{1}{2}(\beta^2 + 1) + |c| \right)^{-1/2} , \quad b = \beta \eta \sin \phi_a \left(\frac{1}{2}(\beta^2 + 1) - |c| \right)^{-1/2} , \quad (86)$$

where

$$c = \frac{1}{2}(\beta^2 - 1) \cos 2\psi_a - \beta \cos \phi_a \sin 2\psi_a . \quad (87)$$

Examples of clutter ridges with velocity misalignment are shown for different values of ϕ_a in Figure 13. For these plots $\beta = 1$. Equation (85) shows that the clutter ellipse will be rotated by 45° from the $+\omega$ -axis.

For each plot the portion of the ellipse corresponding to the half-space in front of the antenna is drawn with a solid line, while the portion corresponding to backlobe clutter is drawn with a dashed line. The aligned case is seen to be a degenerate ellipse where the frontlobe and backlobe portions lie on top of each other. As soon as the misalignment angle becomes nonzero, the backlobe portion of the clutter ridge becomes distinct; in essence there are two clutter ridges. One indicator of the separation between the backlobe and front lobe portions is the length of the minor axis. At $\phi_a = 90^\circ$, the clutter ridge is a circle. Note also that at the target angle of $\vartheta_t = 0$, there is clutter at two Doppler frequencies, one from the front and one from the back. Therefore, severe backlobe clutter and misalignment may reduce the Doppler space available for detecting targets. Figure 14 shows additional examples for ambiguous clutter.

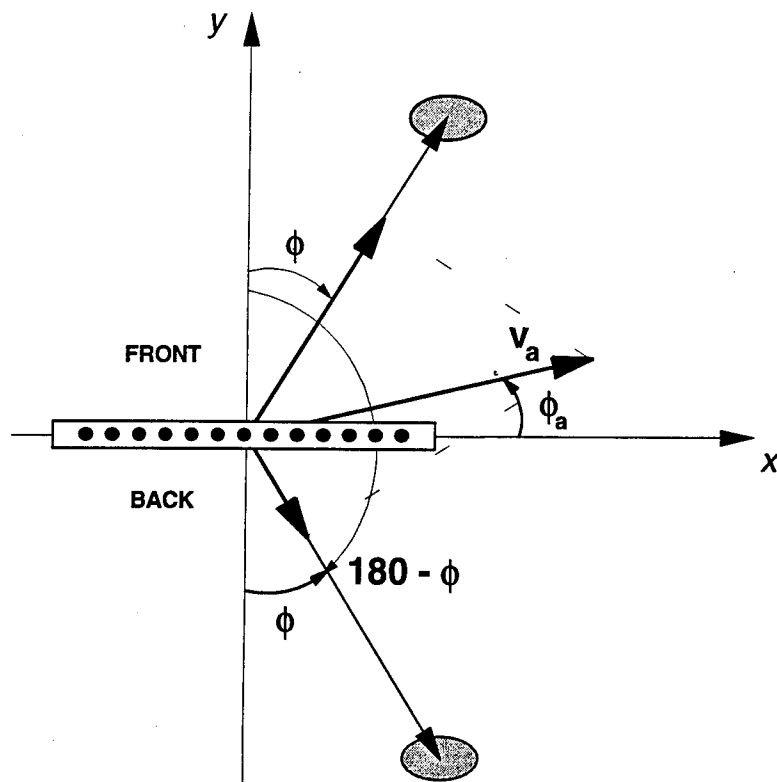


Figure 12. Array geometry with velocity misalignment.

The number of potentially interfering clutter patches is doubled with velocity misalignment.

All of the previous analysis leading to the clutter covariance matrix is still valid with misalignment, except that the Doppler frequencies of the individual clutter patches are given by Equation (82) instead of Equation (70). Brennan's rule, because it relies on a linear relationship between Doppler and spatial frequency, is no longer applicable when $\phi_a \neq 0$. Whether misalignment causes a significant increase in the clutter severity is a strong function of the antenna backlobe level. If the backlobe level is very low, the clutter power on the backlobe ridge will have little impact. On the other hand, if the backlobe clutter is high, the rank of the clutter will approximately double because of the backlobe clutter ridge.

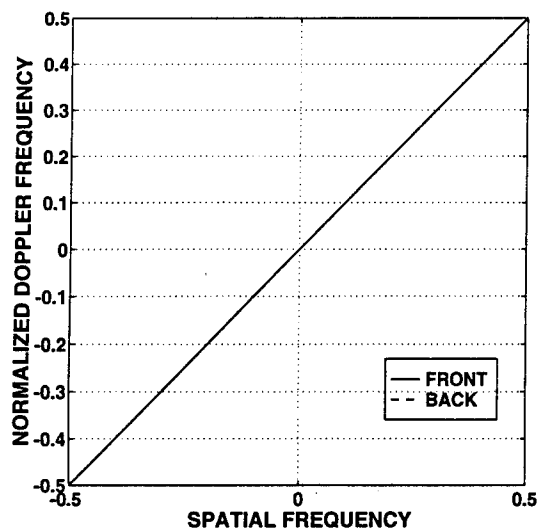
Figure 15 shows the clutter eigenspectra for different values of the misalignment angle. The element pattern is that of Figure 9 and has a -30 -dB backlobe level. The platform velocity corresponds to $\beta = 1$.

For all nonzero ϕ_a , the number of nonzero eigenvalues is approximately doubled as expected due to the backlobe clutter ridge. The power distribution of the highest eigenvalues is changed slightly as the Doppler distribution of clutter power is altered by misalignment. For ϕ_a beyond 1° there does not seem to be much difference in the eigenspectra tails. The effect of backlobe level is shown in Figure 16. Here the backlobe level b_e from Equation(80) is varied as $\phi_a = 10^\circ$ is fixed.

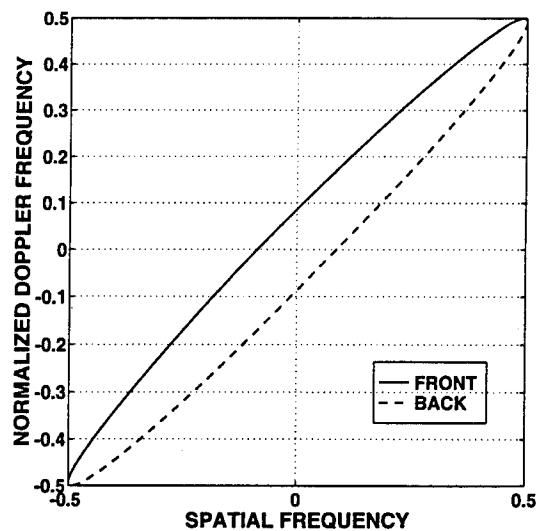
Note that the tails of the eigenspectra show a decrease in level that is essentially linear with the backlobe level. This portion of the eigenspectra is due solely to the backlobe portion of the clutter ridge. With very low backlobes, the numerical rank of the clutter is only slightly higher than that predicted by Brennan's rule. If the backlobes are large the clutter rank will approximately double. These results suggest that more degrees of freedom will be required for effective clutter suppression when misalignment and severe backlobe clutter are present.

2.6.5 Intrinsic Clutter Motion

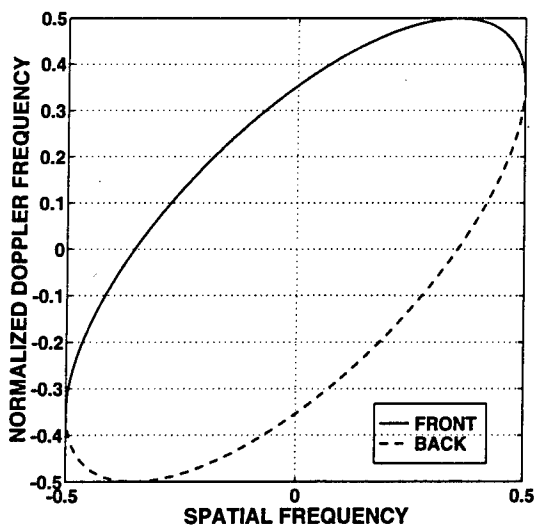
Another assumption made in the development thus far was that the echo from a single clutter patch does not fluctuate from pulse to pulse. This allowed the snapshot from a single scatterer to be expressed in terms of a space-time steering vector and eventually resulted in the convenient sum-of-outer-product form of the space-time covariance matrix given in Equation (60). Many factors, due to both the radar system design and the environment, may in practice cause small pulse-to-pulse fluctuations in the clutter return. The dynamics of a mechanically scanning antenna and any pulse-to-pulse instabilities of the radar system components will produce fluctuations. Natural variation in clutter reflectivity may occur with land clutter due to wind and with sea clutter due to the motion of ocean waves. Any source of fluctuation causes a broadening of the Doppler spectrum of a single clutter echo. The pulse-to-pulse fluctuations due to any of these sources will be referred to as intrinsic clutter motion (ICM). The presence of ICM requires a wider clutter notch or, equivalently, more adaptive degrees of freedom for effective cancellation.



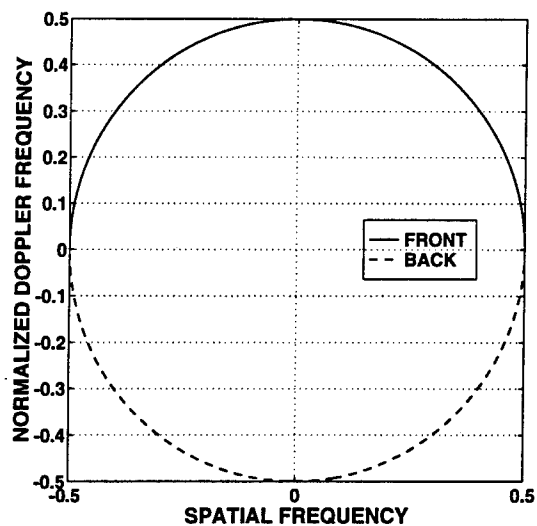
(a)



(b)

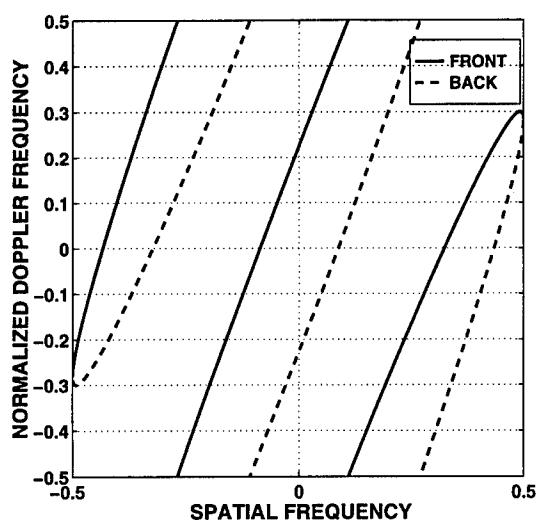


(c)

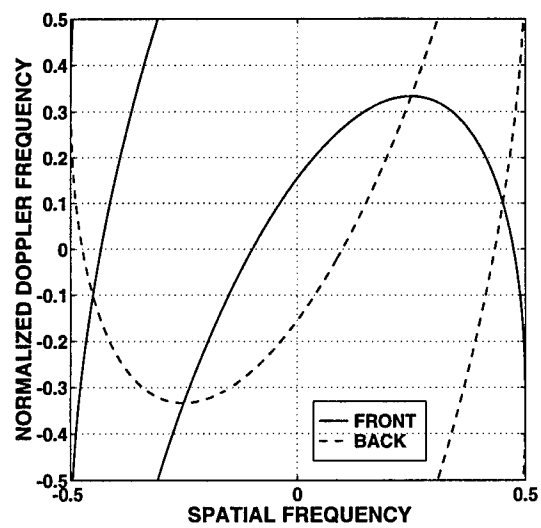


(d)

Figure 13. Example clutter ridges with velocity misalignment, for $\beta = 1$. (a) $\phi_a = 0^\circ$. (b) $\phi_a = 10^\circ$. (c) $\phi_a = 45^\circ$. (d) $\phi_a = 90^\circ$.



(a)



(b)

Figure 14. Example clutter ridges with velocity misalignment, for Doppler-ambiguous clutter. (a) $\beta = 2.6$, $\phi_a = 10^\circ$. (b) $\beta = 2.67$, $\phi_a = 60^\circ$.

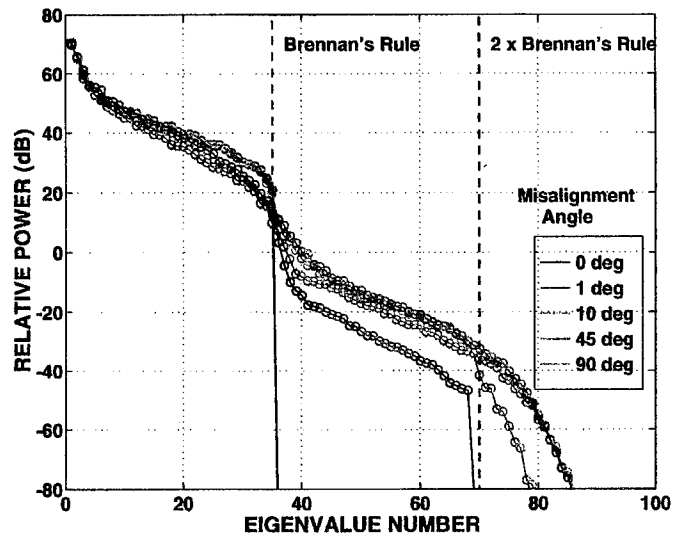


Figure 15. Clutter eigenspectra for different misalignment angles.

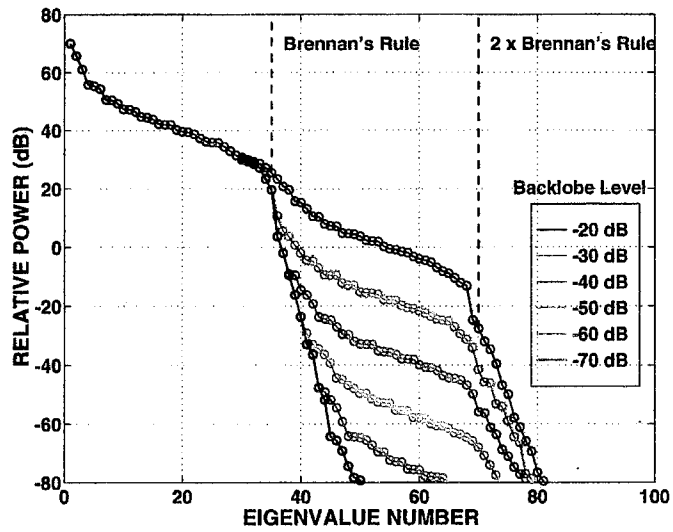


Figure 16. Clutter eigenspectra with misalignment for different backlobe levels.

Recall from Equation (55) the expression for the echo from the k th clutter patch,

$$\chi_c = \alpha_k \mathbf{v}_k = \alpha_k \mathbf{b}_k \otimes \mathbf{a}_k. \quad (88)$$

Temporal fluctuation in the return is represented by replacing the single amplitude α above with the $M \times 1$ vector

$$\boldsymbol{\alpha}_k = \begin{bmatrix} \alpha_{k,0} ; & \alpha_{k,1} ; & \dots ; & \alpha_{k,M-1} \end{bmatrix}, \quad (89)$$

where $\alpha_{k,m}$ is the random amplitude for the k th scatterer from the m th PRI. The result is a space-time snapshot including ICM,

$$\chi_c = (\boldsymbol{\alpha}_k \odot \mathbf{b}_k) \otimes \mathbf{a}_k. \quad (90)$$

The fluctuation will be modeled as a wide-sense stationary random process, and the common assumption of a Gaussian Doppler spectrum will be used [17]. The temporal autocorrelation of the fluctuations is then also Gaussian in shape:

$$\gamma(m) \triangleq E\{\alpha_{l+m}\alpha_l^*\} = \sigma^2 \xi_k \exp\left\{-\frac{\kappa_c^2 T_r^2}{2} m^2\right\}, \quad (91)$$

where ξ_k is the clutter CNR from Equation (58). The spectral standard deviation κ_c is normally expressed in terms of a velocity standard deviation σ_v , defined by

$$\sigma_v = \frac{\lambda_o \kappa_c}{4\pi}. \quad (92)$$

Measured values of σ_v for various types of clutter and conditions are common in most radar systems textbooks [17]. The space-time covariance matrix for a single clutter patch including ICM is then given by

$$\mathbf{R}_c = \xi_k (\boldsymbol{\Gamma}_k \odot \mathbf{b}_k \mathbf{b}_k^H) \otimes \mathbf{a}_k \mathbf{a}_k^H, \quad (93)$$

where

$$\mathbf{\Gamma}_k = E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^H\} = \text{Toeplitz}(\gamma_c(0); \dots; \gamma_c(M-1)) \quad . \quad (94)$$

is the $M \times M$ covariance matrix of the fluctuations for the k th patch.

Superposition is invoked to yield the result when many independent clutter sources are present:

$$\mathbf{R}_c = \sum_{k=1}^{N_c} \xi_k (\mathbf{\Gamma}_k \odot \mathbf{b}_k \mathbf{b}_k^H) \otimes \mathbf{a}_k \mathbf{a}_k^H \quad . \quad (95)$$

It is assumed here that although the clutter patches are independent, they have the same intrinsic velocity spectrum. Note that the clutter covariance is no longer equal to a sum of outer products of space-time steering vectors.

The effect of ICM on the clutter eigenspectrum is illustrated in Figure 17. The example system of Table 2 with zero crab angle and $\beta = 1$ is assumed. The velocity standard deviation is varied from 0 to 0.5 m/s.

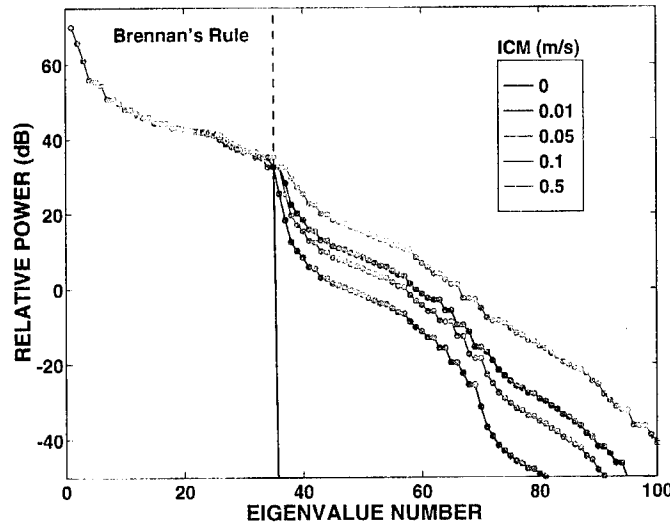


Figure 17. Clutter eigenspectra for different levels of ICM.

Nonzero intrinsic clutter motion has little effect on the portion corresponding to the largest eigenvalues. As σ_v increases, the tails of the eigenspectrum become larger as the rank of the covariance increases. ICM is another phenomenon that, if present, results in the actual clutter rank's being larger than what would be predicted by Brennan's rule. The increased rank will require broader nulls in Doppler frequency, which in turn means more adaptive degrees of freedom.

2.7 Range-Doppler View of the Interference

This analysis has concentrated on the components of the space-time snapshot corresponding to a single range gate. Because the snapshot comprises the samples from the array elements and from multiple pulses, the signals were characterized in terms of their angle and Doppler content. The full radar data cube consists of snapshots from all ranges. In practice, data from multiple ranges is averaged to derive estimates of the interference. Therefore, the range dependence of the various signal components must also be understood.

The assumption of continuous barrage noise jamming results in a jamming signal that is present at all range gates. Moreover, assume that the jamming is stationary over the time scale of a PRI so that its statistics are independent of radar range; the jamming components of all range gates are identically distributed. Therefore, the jamming component of the radar datacube, like thermal noise, is assumed to occupy all of the range-Doppler space.

Ground clutter, on the other hand, has a very definite range dependence. Not only does the clutter power depend on range through the radar equation, Equation (58), but its elevation angle is a function of range and its reflectivity in practice is a function of range, as real clutter is very nonhomogeneous. The extent of the clutter in angle and Doppler has been discussed; its extent in range is a function of the platform altitude and the radar PRF. For low-PRF systems, the clutter is typically ambiguous in Doppler, as shown in Figure 7, but unambiguous in range because the radar unambiguous range R_u is typically greater than the radar horizon R_h . This situation is shown in Figure 18. Note that beyond the horizon, range gates are corrupted only by jamming and noise. This is significant for two reasons. First, no clutter cancellation is required for these ranges. Second, these range gates can be utilized to obtain jamming estimates that are uncorrupted by clutter. Ways in which this information can result in efficient adaptive processing architectures will be described in later sections.

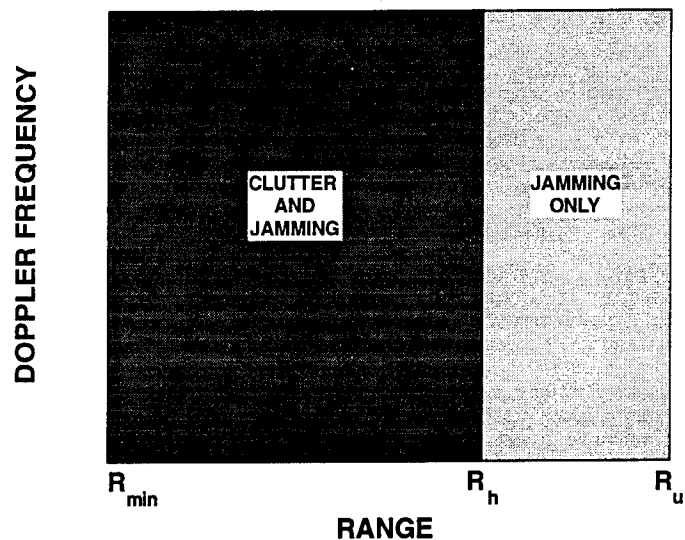


Figure 18. Range-Doppler view of interference, for the case of a low-PRF waveform. The clutter is ambiguous in Doppler and unambiguous in range.

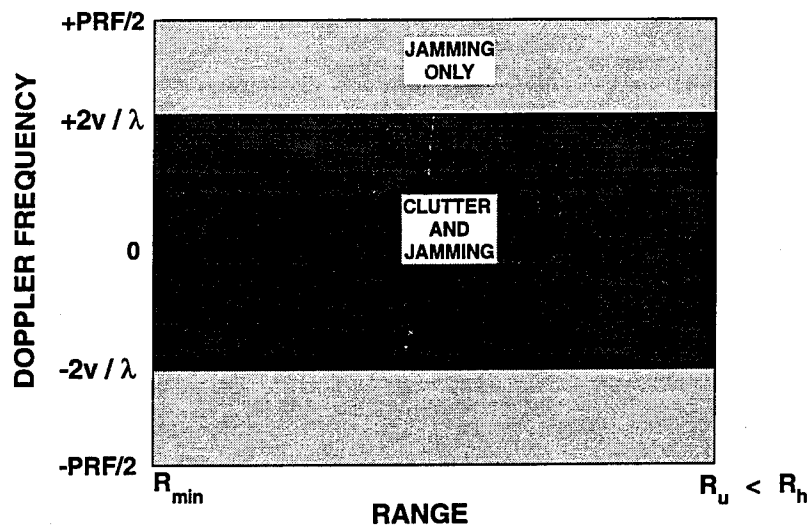


Figure 19. Range-Doppler view of interference, for a high-PRF waveform. The clutter is unambiguous in Doppler but ambiguous in range.

The other extreme applies to a high-PRF waveform. In this case, shown in Figure 19, the clutter is unambiguous in doppler but ambiguous in range. Again, jamming spans all of range and Doppler, but a clutter-free region in Doppler exists. Other variants are possible with different PRF, frequency, and platform altitude combinations.

2.8 Summary

Beginning with a generic radar system model, a model has been developed for the signals received by an airborne radar. A space-time snapshot was defined to be the slice of the datacube corresponding to a single range gate. This data may contain a target component as well as undesired components due to noise, jamming, and clutter. The target signal is modeled as a random amplitude times a space-time steering vector that has the target's angle and Doppler. The undesired signal components are modeled as random processes and expressions for their covariance matrices are derived. The results are summarized below.

A space-time snapshot may be decomposed as

$$\chi = \alpha_t \mathbf{v}_t + \chi_u, \quad (96)$$

where the undesired component

$$\chi_u = \chi_c + \chi_j + \chi_n \quad (97)$$

consists of clutter, jamming, and thermal noise. These components are mutually uncorrelated, so that the interference plus noise covariance matrix is given by

$$\mathbf{R}_u = E\{\chi_u \chi_u^H\} = \mathbf{R}_c + \mathbf{R}_j + \mathbf{R}_n, \quad (98)$$

where

$$\mathbf{R}_n = \sigma^2 \mathbf{I}_{MN} \quad (99)$$

$$\mathbf{R}_j = \mathbf{I}_M \otimes \mathbf{A}_j \Xi_j \mathbf{A}_j^H \quad (100)$$

$$\mathbf{R}_c = \sum_{k=1}^{N_c} \xi_k (\Gamma_k \odot \mathbf{b}_k \mathbf{b}_k^H) \otimes \mathbf{a}_k \mathbf{a}_k^H. \quad (101)$$

For reference, the assumptions made in this chapter are summarized in Table 3. The model developed in this chapter will serve as the input to various space-time adaptive processing algorithms. A general discussion of STAP and fully adaptive space-time processing will begin the next chapter.

TABLE 3
Assumptions Made for Radar System and Signal Model

ASSUMPTION	PAGE NO.
Radar System	
Non-relativistic velocities	13
Narrowband waveforms and receivers ($B \ll f_o$)	14
Transmitted pulse insensitive to expected Doppler shifts	15
Constant PRF waveform	9
Constant platform velocity	7
Array axis and platform velocity vector lie in the horizontal plane	7,33
Uniform linear antenna array (possibly columns of a planar array)	7
Identical element patterns	7
Target Model	
Point target	12
Constant velocity	12
Noise Model	
Internally generated receiver noise is dominant source	17
Noise signals on different elements are mutually uncorrelated	17
Noise decorrelates over a PRI ($f_r \ll B$)	17
Jamming Model	
Only continuous barrage noise jamming is considered	18
Jamming signal decorrelates over a PRI	18
Jamming may be assumed stationary over a CPI	18
Clutter Model	
Clutter from a single range approximated by large number of independent scatterers	22
Constant gamma reflectivity model	23
Gaussian intrinsic clutter motion spectrum	45
ICM spectrum is the same for each clutter patch	47

3. SPACE-TIME PROCESSING FUNDAMENTALS

3.1 Introduction

For the purposes of this report, a space-time processor is defined as a linear combiner that sums the spatial samples from the elements of an antenna array and the temporal samples from the multiple pulses of a coherent waveform. The major components of a STAP algorithm are described: the algorithm architecture, the weight training and application strategy, and the weight computation approach. Fully adaptive space-time processing is discussed, which refers to a brute force approach whereby the signals at each element and pulse are adaptively weighted. If the statistics of the interference are known, fully adaptive STAP is optimum. A fully adaptive suboptimum approach is also described. Various STAP algorithm performance metrics are defined and illustrated, with fully adaptive STAP as an example. Finally, the chapter concludes with a discussion of the limitations of fully adaptive STAP. These shortcomings lead in the next chapter to the consideration of reduced-dimension algorithms. Fully adaptive STAP (optimum), however, will serve as a useful performance bound on any suboptimum approach. The metrics developed here will prove useful in subsequent chapters as well.

3.2 General Architecture and Assumptions

The function of a surveillance radar is to search a specified volume of space for potential targets. Within a single coherent processing interval, the search is confined in angle to the sector covered by the transmit beam for that CPI, but otherwise it covers all ranges. Consider a fixed range gate which is to be tested for target presence. The data available to the radar signal processor consists of the M pulses on each of N elements. A space-time processor is defined to be a linear filter that combines all the samples from the range gate of interest to produce a scalar output. This process is depicted in Figure 20. The tapped delay line on each element represents the multiple pulses of a CPI, with the time delay between taps equal to the PRI. Thus, a space-time processor utilizes the spatial samples from the elements of an array antenna and the temporal samples provided by the successive pulses of a multiple-pulse waveform. The space-time processor can be represented by an MN -dimensional weight vector \mathbf{w} . Its output z can be represented as the inner product of the weight vector and the snapshot of interest,

$$z = \mathbf{w}^H \boldsymbol{\chi} . \quad (102)$$

One way to view a space-time weight vector is as a combined receive array beamformer and target Doppler filter. Ideally, the space-time processor provides coherent gain on target while forming angle and Doppler response nulls to suppress clutter and jamming. As the clutter and jamming scenario is not known in advance, the weight vector must be determined in a data-adaptive way from the radar returns. A single weight vector is optimized for a specific angle and Doppler.

Since the target angle and velocity are also unknown a priori, a space-time processor typically computes multiple weight vectors that form a filter bank to cover all potential target angles and Doppler frequencies.

A more complete picture of a space-time processor is shown in Figure 21. Here the full CPI datacube is shown, with the shaded slice of data, labeled "target data," representing the data at the range gate of interest. This shaded portion is exactly the data represented by the tapped delay line on each element in Figure 20. The space-time processor consists of three major components. First a set of rules called the *training strategy* is applied to the data. This block derives from the CPI data a set of training data that will be used to estimate the interference. The second step is *weight computation*. Based on the training data, the adaptive weight vector is computed. Typically, weight computation requires the solution of a linear system of equations. This block is therefore a very computation-intensive portion of the space-time processor. New weight computations are performed with each set of training data. Finally, given a weight vector, the process of *weight application* refers to the computing of the scalar output or test statistic. Weight application is an inner product, or digital beamforming, operation. The scalar output is then compared to a threshold to determine if a target is present at the specified angle and Doppler. The output of the processor is a separate scalar (or decision) for each range, angle, and velocity at which target presence is to be queried. Each of these blocks will be discussed in more detail.

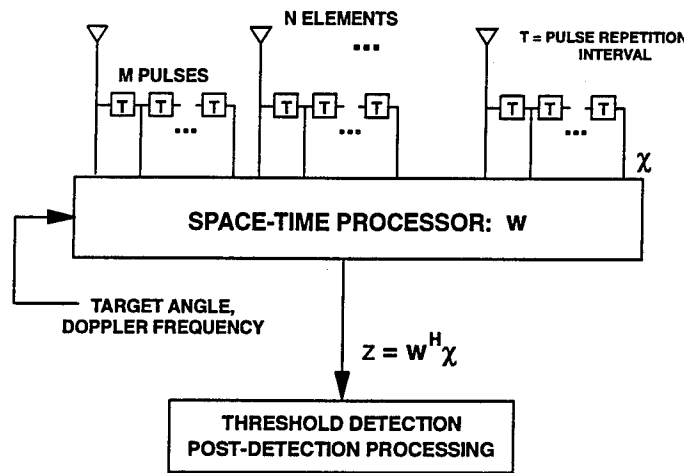


Figure 20. A general block diagram for a space-time processor.

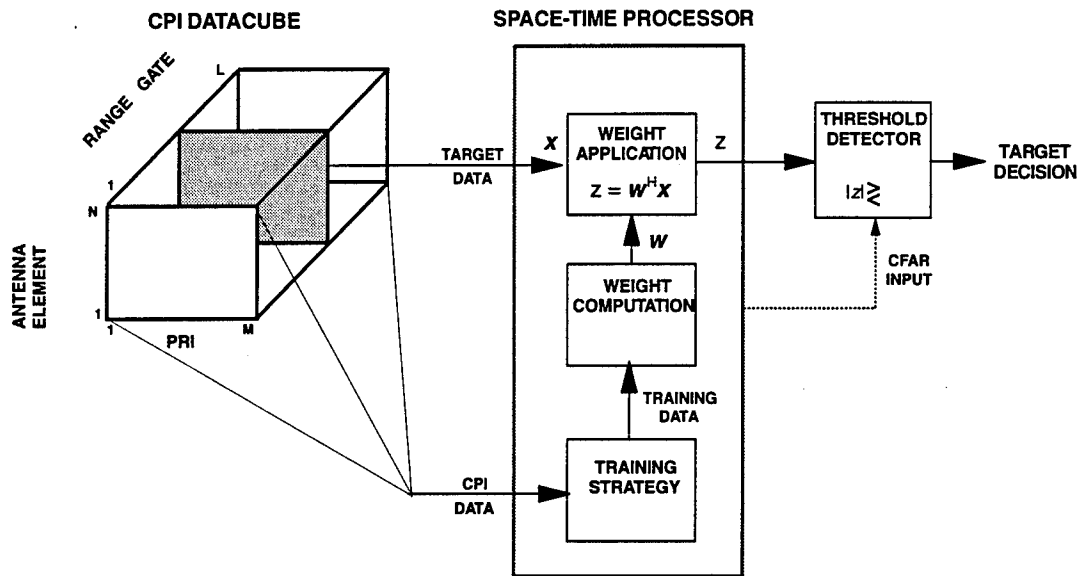


Figure 21. Data-domain view of space-time adaptive processing.

Because the interference is unknown a priori, it must be estimated data-adaptively from the finite amount of data comprising the CPI. The processor accepts the CPI data and implements a set of training rules to derive a secondary data set, called the training data, to be used for weight computation. The goal of the training strategy is to obtain the best estimate of the interference that exists at the range gate under test. Typically, the data from several range gates near the range gate of interest are used. A training strategy is defined with a number of factors in mind. First, the number of range gates in the training set must be sufficient to guarantee an interference estimate good enough for effective nulling; this issue is taken up further in Section 3.5. In a stationary environment, the required number of samples is well understood. Secondly, the training set must be updated or changed in accordance with the nonstationarity of the interference. The training strategy and weight computation requirements are coupled, because for each change in the training set a new weight vector must be computed. Training strategies in nonstationary environments are less understood and tend to be more heuristic. Since the number of range gates in the CPI data is dependent on the PRI and on the instantaneous bandwidth, training strategies are most strongly affected by these two radar system parameters.

Given a training set, weight computation strategies fall loosely into three categories. The first category will be called sample matrix inversion, or SMI [24]. The SMI weight computation refers to all approaches that effectively compute the weight vector from the inverse of the sample covariance matrix of the training data. SMI also refers to the more numerically stable algorithms whereby the weights are computed from a QR-decomposition of the matrix of training set data. The members of the SMI class differ by the way in which the QR-decomposition is performed. The second class will be called subspace projection. Algorithms in this class first estimate the subspace spanned by the interference by performing an eigenanalysis of the sample covariance matrix or a singular value decomposition (SVD) of the matrix of training data. The weight vector is then computed by projecting the desired response into the subspace orthogonal to the interference subspace. In this way the weight vector is forced to null the interference. The third class, subspace SMI, combines attributes from the first two classes. Here the data is first projected into a lower dimension space by means of some transformation, then a small SMI problem is solved with the projected data (and steering vector) in the lower dimension space. The transform may be either fixed or data-adaptive. In contrast with subspace projection, subspace SMI preserves SINR optimality in the ideal case. Because the size of the SMI problem is reduced, computational and training requirements are lessened. SMI and training is discussed further in Section 3.5.

Weight application is the formation of the processor output given the computed weight vectors. In practice, a single weight vector may be applied to data comprising many range gates. The design of the weight application regions is usually coupled with the training set design, with each application region corresponding to a single training set. Weight application is an inner product, or matrix-vector product, operation. The computational load of this portion of the space-time processor scales linearly with the weight vector dimension and the number of range gates. The number of range gates in turn depends on the radar PRF and its instantaneous bandwidth.

The processor output scalar is compared to a threshold to determine target presence at each range-Doppler-angle cell. Typically, a background noise estimate is provided to the detector so that it provides constant-false-alarm rate (CFAR). The selection of a training region resembles the choice of a CFAR stencil; the training set or application region data may be used to set the CFAR constant. It has been shown that with appropriate normalization of an SMI weight vector, a CFAR property can be embedded into the weight computation. Different normalizations lead to the adaptive-matched-filter (AMF) detector [28] or the Generalized-Likelihood-Ratio (GLR) detector [29]. The detection performance of these two approaches is similar at moderate to high SNR, but subtle differences in performance at low target SNR and with respect to sidelobe targets have been documented [25, 10, 35]. The approach in this report will not presume a detector structure. A number of different STAP algorithm architectures will be considered, each of which may be implemented with various detection methods, e.g., AMF, GLR, or another type of CFAR.

With this general breakdown of space-time processing in mind, consider the first STAP algorithm.

3.3 Fully Adaptive STAP

A space-time processor that computes and applies a separate adaptive weight to every element and pulse is said to be fully adaptive. The weight vector for a fully adaptive processor is of size MN . Fully adaptive space-time processing for airborne radar was first proposed in Brennan et al. [4], and is a natural extension of adaptive antenna processing [22, 23] to a two-dimensional space-time problem.

Assume that at the range of interest, a target signal is present in a background of interference. Let the target angle, Doppler, and amplitude be given by ϑ_t , ϖ_t , and α_t , respectively. The data snapshot at the range of interest may be written, from Equation (96), as

$$\chi = \alpha_t \mathbf{v}_t + \chi_u, \quad (103)$$

where $\mathbf{v}_t = \mathbf{b}(\varpi_t) \otimes \mathbf{a}(\vartheta_t)$ is the target steering vector and χ_u denotes the interference (clutter, jamming) plus noise components of the data. It is well known that the optimum space-time filter [3] is given to within a scale factor by

$$\mathbf{w} = \mathbf{R}_u^{-1} \mathbf{v}_t, \quad (104)$$

where

$$\mathbf{R}_u = E \{ \chi_u \chi_u^H \} \quad (105)$$

is the interference-plus-noise covariance matrix.

The weight vector in Equation (104) is optimum under several criteria [27]. It maximizes signal-to-interference-plus-noise-ratio (SINR), maximizes probability of detection for a given false alarm probability, and with the proper choice of scale factor minimizes output power subject to a unity gain constraint in the target direction. The optimum processor response has high sidelobes in both angle and Doppler, however, because of the implied windowing of the data. These high sidelobes make the optimum processor susceptible to the detection of sidelobe targets.

It is sometimes desirable to consider a suboptimum fully adaptive processor given by

$$\mathbf{w} = \mathbf{R}_u^{-1} \mathbf{g}_t. \quad (106)$$

Here \mathbf{g}_t is not the target steering vector but a desired weight vector that may include tapering to obtain low sidelobes. One way to form \mathbf{g}_t is to apply a low-sidelobe window to an assumed target steering vector. Let \mathbf{t}_a be an $N \times 1$ vector containing the desired low-sidelobe angle response, and let \mathbf{t}_b be the $M \times 1$ vector of the desired Doppler response. The vector

$$\mathbf{t} = \mathbf{t}_b \otimes \mathbf{t}_a \quad (107)$$

is a separable space-time window sequence. The vector

$$\mathbf{g}_t = \mathbf{t} \odot \mathbf{v}_t \quad (108)$$

will produce a low-sidelobe adapted pattern. The term “tapered fully adaptive” will be used when the steering vector is chosen in this way. The tapered fully adaptive algorithm is not strictly optimum in any sense. In a strong interference environment, however, the weight vector from Equation (106) approximates the best least squares fit to the desired response \mathbf{g}_t , subject to its having response nulls on the interference.⁶ This fact is the primary motivation for use of the tapered fully adaptive approach.

A block diagram of fully adaptive STAP is given in Figure 22. Whatever the choice of steering vector, fully adaptive STAP requires the solution of an MN -dimensional linear system of equations.

The size of the linear system grows linearly with the size of the array or the length of the coherent processing interval. For many radar systems, the product MN is likely to range from several hundred to several thousand. The implementation of a fully adaptive approach is beyond current capabilities in real-time computing. The performance of fully adaptive STAP, however, will be considered here for two reasons. First, it allows illustration of a number of performance metrics that will be useful in the study of reduced-dimension algorithms. Second, the performance of optimum fully adaptive STAP will serve as a baseline to which the performance of suboptimum approaches can be compared.

⁶This argument follows from a decomposition of the covariance matrix into interference and noise subspaces. As the interference-to-noise ratio approaches infinity, the weight vector reduces to a projection of the desired response vector onto the vector space orthogonal to the interference subspace.

3.4 STAP Performance Metrics and Fully Adaptive Performance

To demonstrate fully adaptive STAP performance, return to the example airborne early warning system described in Table 2. The flight parameters and interference scenario assumed are given in Table 4. Both clutter and jamming are present; these parameters amount to a JNR per element of 38 dB and a CNR per element per pulse of 47 dB. In this baseline scenario, assume the clutter is exactly unambiguous in Doppler and that there is no velocity misalignment or intrinsic clutter motion. More stressing environments will be considered in Chapter 7. With this scenario, the performance of both optimum fully adaptive and tapered fully adaptive STAP is considered under a variety of performance metrics. These metrics will also be used in subsequent chapters to compare various suboptimum STAP algorithms.

3.4.1 Adapted Patterns

Given the weight vector produced by a space-time processor, its response as a function of angle and Doppler is one indicator of the processor performance. This response is called the adapted pattern, defined by

$$P_{\mathbf{w}}(\vartheta, \varpi) = \left| \mathbf{w}^H \mathbf{v}(\vartheta, \varpi) \right|^2. \quad (109)$$

The adapted pattern is a two-dimensional angle-Doppler frequency response. For a uniform linear array and a fixed pulse repetition interval, the pattern is a two-dimensional (inverse) Fourier transform of the weight vector. Ideally, the adapted pattern has nulls in the directions of interference sources and high gain at the angle and Doppler of the presumed target direction. The shape and sidelobe levels of the adapted pattern are also of interest. In the absence of interference (noise only), the adapted pattern is commonly called the *quiescent* pattern.

Figure 23 shows an adapted pattern for optimum STAP. The presumed target is located at 0° azimuth with Doppler frequency 100 Hz. The pattern's mainlobe is at the target location. The jamming is suppressed by the deep vertical pattern nulls at the jammer azimuths. The pattern has a slanted null that spans the clutter ridge, thereby suppressing both mainlobe and sidelobe clutter. These pattern nulls suppress the interference at the output to well below thermal noise. Figure 23(b) shows the two principal cuts of this pattern. The azimuth pattern at the target Doppler represents the receive beamformer; it exhibits nulls at both the jammer azimuths and the azimuth where the sidelobe clutter has the same Doppler as the target. Note the high azimuth sidelobes due to the implied uniform taper of the target steering vector. The second cut shows the Doppler response at the target azimuth. The deep clutter null evident at zero Doppler suppresses mainlobe clutter. Jammer and clutter nulls present at other azimuths do not appear in this pattern cut. Note that the Doppler response also has high sidelobes due to the implied uniform taper of the temporal component of the target steering vector.

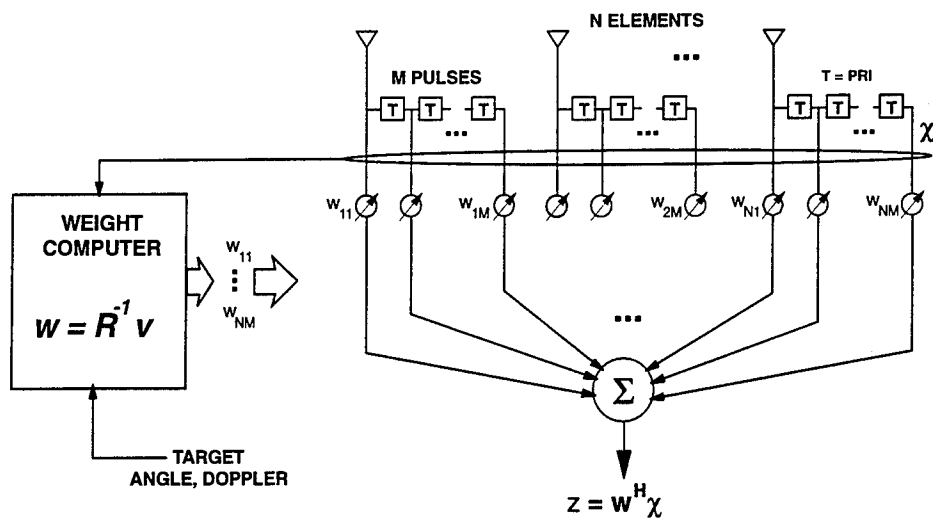


Figure 22. Fully adaptive space-time processing

TABLE 4
Platform and Interference Scenario for Baseline Scenario.

Platform Parameters

Platform altitude	9000 m
Platform velocity	50 m/s
Number of clutter foldovers	$\beta = 1$
Velocity misalignment angle	0°

Interference Scenario

Jamming	
Number of jammers	2
Azimuth angles	$-40^\circ, 25^\circ$
Elevation angles	$0^\circ, 0^\circ$
ERPD	1000 W/MHz
Range	370 km
Clutter	
Number of patches	360
Range	130 km
Reflectivity γ	-3 dB
Intrinsic velocity σ_v	0 m/s

The corresponding patterns for tapered fully adaptive STAP are given in Figure 24. The algorithm uses a separable space-time taper comprising a spatial 30-dB Chebyshev taper and a temporal 40-dB Chebyshev taper. The jammer and clutter nulls are again clearly evident. The pattern mainlobe is widened in both angle and Doppler because of the taper. The pattern cuts show the reduced sidelobe levels in both angle and Doppler. The price paid for the lower sidelobes is a widened mainlobe and a slight loss of SNR gain (1.79 dB in this case). When considering reduced-dimension algorithms, one consideration in the algorithm design will be to maintain low sidelobe adapted and quiescent patterns.

3.4.2 SINR

A common measure of processor performance is the **output signal-to-interference-plus-noise ratio, or SINR**. Divide the output signal into target and interference-plus-noise components,

$$z = z_t + z_u = \alpha_t \mathbf{w}^H \mathbf{v}_t + \mathbf{w}^H \chi_u. \quad (110)$$

Let $p_t = E\{|z_t|^2\}$ and $p_u = E\{|z_u|^2\}$ be the output target power and output interference-plus-noise power. The **SINR is then defined as**

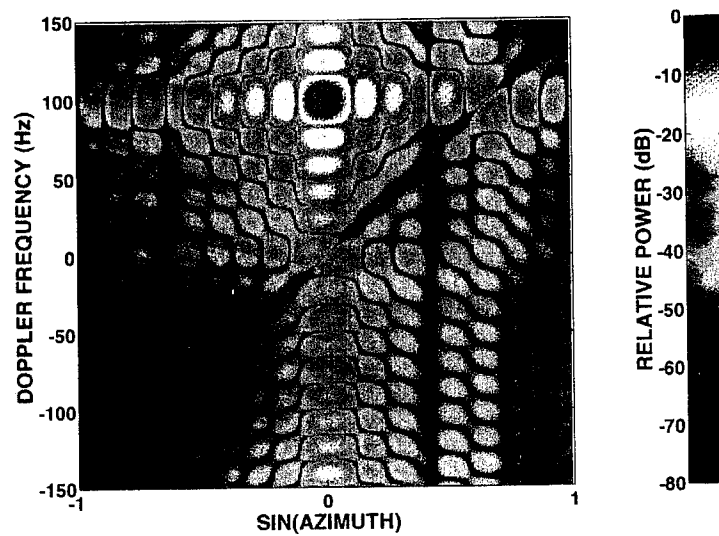
$$SINR = \frac{p_t}{p_u} = \frac{\sigma^2 \xi_t |\mathbf{w}^H \mathbf{v}_t|^2}{\mathbf{w}^H \mathbf{R}_u \mathbf{w}}, \quad (111)$$

where ξ_t is the target SNR on a single pulse for a single array element. Since σ^2 is the noise power per element, the quantity $\sigma^2 \xi_t$ is the input target power per element per pulse. Substitution of the optimum weight vector into Equation (111) leads to the **optimum SINR given by**

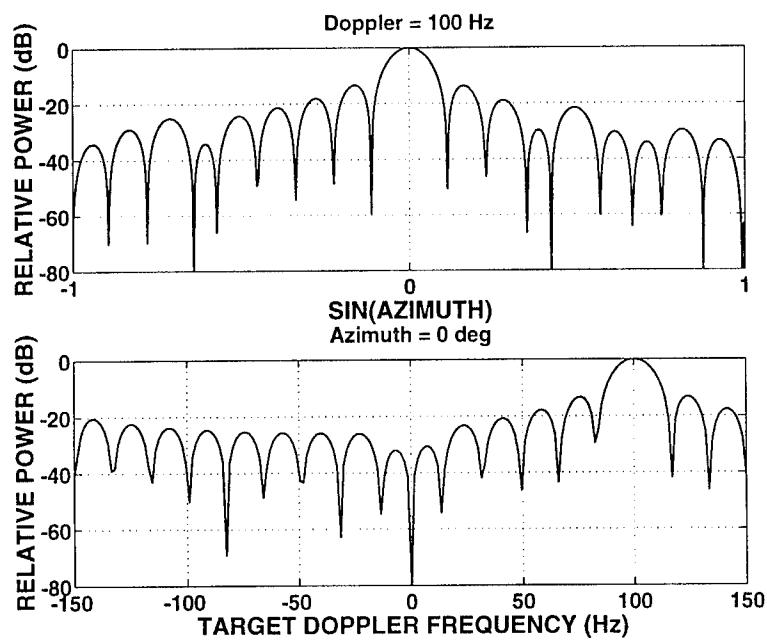
$$SINR_o = \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}_u^{-1} \mathbf{v}_t. \quad (112)$$

Similarly, tapered fully adaptive produces a suboptimum SINR given by

$$SINR_c = \frac{\sigma^2 \xi_t |\mathbf{g}_t^H \mathbf{R}_u^{-1} \mathbf{v}_t|^2}{\mathbf{g}_t^H \mathbf{R}_u^{-1} \mathbf{g}_t}. \quad (113)$$

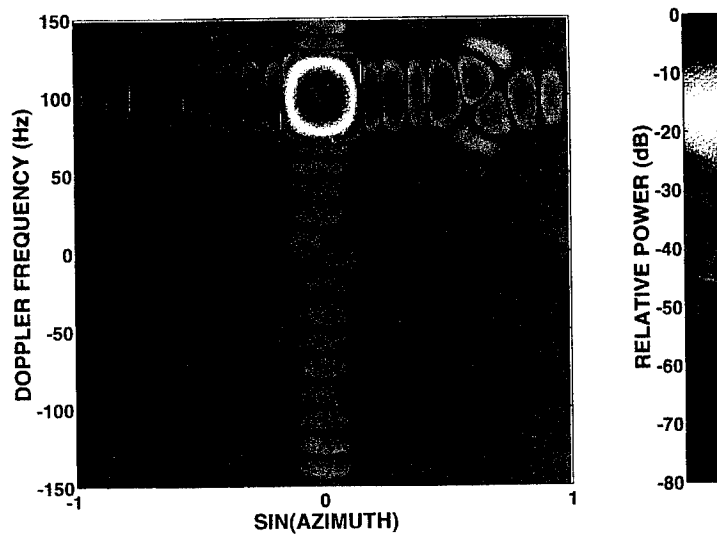


(a)

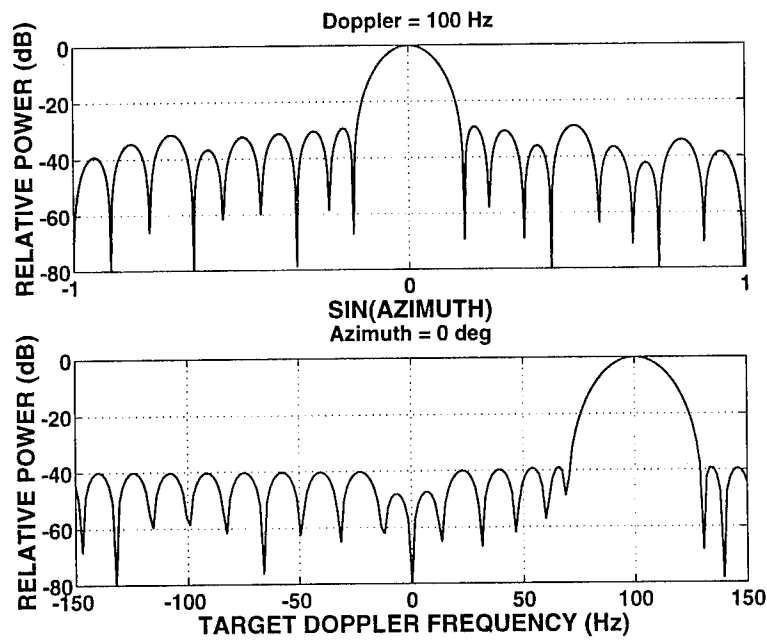


(b)

Figure 23. Example scenario: optimum fully adaptive STAP. (a) Adapted pattern. (b) Principal cuts at target azimuth and Doppler.



(a)



(b)

Figure 24. Example scenario: tapered fully adaptive STAP. (a) Adapted pattern. (b) Principal plane cuts at target azimuth and Doppler.

The above equations yield performance at a single angle and Doppler. Since the target velocity is unknown, **the interest in SINR performance is as a function of target Doppler.** The Doppler space performance is considered by holding the target angle fixed and varying the target Doppler, computing a separate adaptive weight vector for each Doppler.⁷ Let $\mathbf{v}_t(\varpi) = \mathbf{b}(\varpi) \otimes \mathbf{a}(\vartheta_t)$ and $\mathbf{g}_t(\varpi) = \mathbf{t} \odot \mathbf{v}_t(\varpi)$ be the untapered and tapered steering vectors for Doppler ϖ . Therefore,

$$\text{SINR}_o(\varpi) = \sigma^2 \xi_t \mathbf{v}_t(\varpi)^H \mathbf{R}_u^{-1} \mathbf{v}_t(\varpi), \quad (114)$$

and

$$\text{SINR}_c(\varpi) = \frac{\sigma^2 \xi_t |\mathbf{g}_t(\varpi)^H \mathbf{R}_u^{-1} \mathbf{v}_t(\varpi)|^2}{\mathbf{g}_t(\varpi)^H \mathbf{R}_u^{-1} \mathbf{g}_t(\varpi)}. \quad (115)$$

Figure 25 shows fully adaptive performance for the example scenario. The input SNR is taken to be 0 dB. Optimum fully adaptive STAP achieves about 25 dB SINR over the majority of the Doppler space. A perfect matched filter in an interference-free environment provides an SINR of $10 \log_{10} MN = 25.1$ dB. In the presence of interference, optimum fully adaptive STAP is providing near maximum gain on target while suppressing both clutter and jamming to well below thermal noise. When the target is near 0 Hz or 300 Hz the SINR is very low, because in this case the target is close to the mainlobe clutter in both angle and Doppler. Performance degrades as the target falls into the response null that the processor has placed on the mainlobe clutter.

The same figure also shows the performance of tapered fully adaptive STAP. A 30-dB Chebyshev spatial taper and a 40-dB Chebyshev Doppler taper were taken as the desired response. The tapered fully adaptive performance is always slightly lower than the optimum. Otherwise, the shape of the curve is similar to the fully adaptive case. Over most of the Doppler space, the SINR achieved with tapered fully adaptive STAP is about 1.8 dB less than optimum. This difference is approximately equal to the combined angle and Doppler taper losses. This loss, and the widened response mainlobe, is the price paid for lower adapted pattern sidelobes.

⁷To plot SINR versus target *angle* or versus angle and Doppler, one must also account for the angle dependence of the input SNR ξ_t , which follows the shape of the transmit pattern.

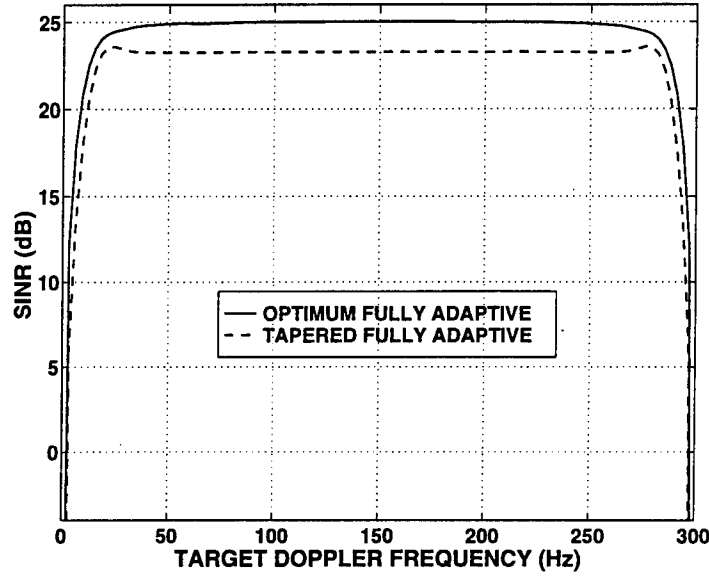


Figure 25. SINR for the optimum and tapered fully adaptive STAPs.

In the fully adaptive SINR plots of Figure 25, a separate weight vector was computed for every potential target Doppler. The resultant optimum SINR curve is a smooth upper bound on the performance achievable with any suboptimum STAP algorithm. In practice, however, a set of weight vectors would be computed to form a bank of space-time filters that cover the Doppler space. The number of filters is typically equal to M , the number of pulses in the CPI. Let ϖ_m be the Doppler to which the m th filter is tuned, and let \mathbf{w}_m be the corresponding weight vector. For example, with optimum fully adaptive STAP,

$$\mathbf{w}_m = \mathbf{R}_u^{-1} \mathbf{v}(\varpi_m) . \quad (116)$$

Furthermore, let $SINR_m(\varpi)$ be the SINR achieved with the m th filter against a target at Doppler ϖ . The algorithm SINR is defined to be the SINR of the filter with maximum SINR,

$$SINR(\varpi) = \max_m SINR_m(\varpi) . \quad (117)$$

For targets whose Doppler is not at the center frequency of one of the Doppler filters, an additional straddling loss is incurred.

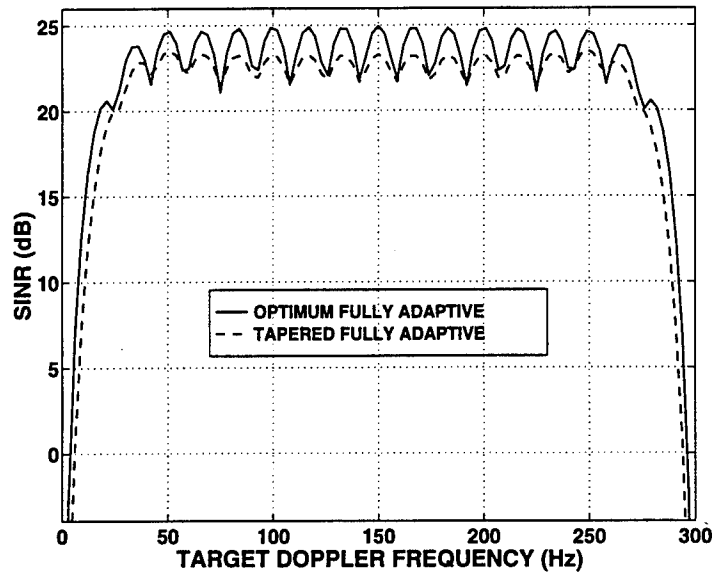


Figure 26. SINR for the optimum and tapered fully adaptive STAPs, including Doppler straddling losses.

Figure 26 shows fully adaptive STAP performance including straddling losses. For each curve a bank of $M = 18$ filters is formed. Optimum STAP, with its implied uniform taper, results in space-time filters with narrower Doppler responses, and therefore more straddling losses, than tapered fully adaptive.

3.4.3 SINR Loss

It is often very useful to express STAP performance relative to what could be obtained in the absence of interference. In a noise-only environment, the optimum processor is simply

$$\mathbf{w} = \mathbf{v}_t, \quad (118)$$

which is a space-time matched filter. The optimum output signal-to-noise ratio, denoted SNR_o , is then

$$SNR_o = MN\xi_t. \quad (119)$$

The gain of MN represents coherent spatial and temporal integration over N elements and M pulses.

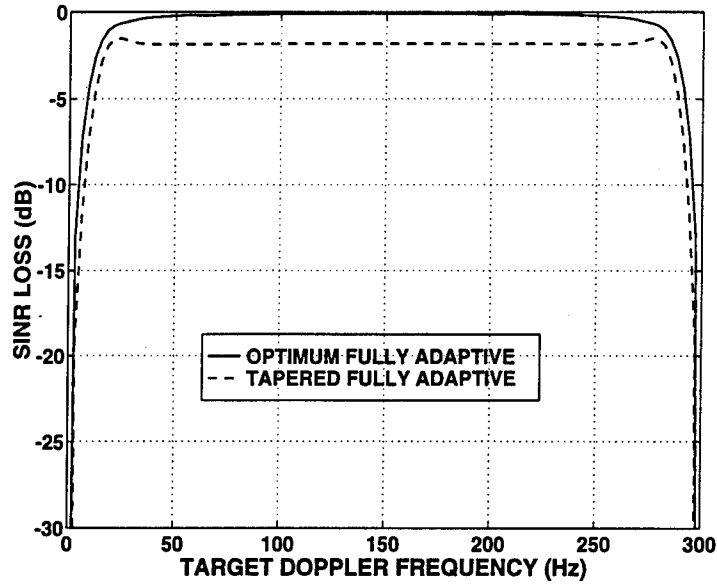


Figure 27. SINR loss for the fully adaptive STAP.

The SINR loss, L_{SINR} , of a space-time processing algorithm is defined to be its performance relative to the matched filter SNR in an interference-free environment. Thus,

$$L_{SINR}(\varpi) = \frac{SINR(\varpi)}{SNR_o}. \quad (120)$$

Note that L_{SINR} lies between 0 and 1. SINR loss is useful because it incorporates many of the factors contributing to performance loss in a single quantity. In addition to the losses incurred by suppressing the interference, SINR loss as defined above includes both taper losses and filter straddling losses. It will be shown in Section 3.5 how the effect of covariance estimation can also be incorporated into the SINR loss metric. SINR loss can be a useful way to translate space-time processor performance to a radar system's detection performance by including it as an additional loss factor in the radar equation. For example, if R_{\max} is the radar's noise-limited detection range, the system would achieve better than half this detection range for all scenarios where the SINR loss is greater than 12 dB.

Figure 27 shows the SINR loss for the fully adaptive STAP algorithms considered above. As expected, the optimum STAP produces an SINR loss of near 0 dB for most of the Doppler space. These curves are just scaled versions of the SINR plot of Figure 25, and their explanation is the same. SINR loss will be one of the primary metrics used to compare various suboptimum algorithms.

3.4.4 SINR Improvement Factor

Another way to present STAP algorithm performance is as a gain in SINR relative to that which exists on a single element and a single pulse. Let $SINR_{in}$ be the SINR on a single element for a single pulse. This is given by

$$SINR_{in} = \frac{\sigma^2 \xi_t}{R_u(1, 1)} = \frac{\xi_t}{1 + \xi_c + \xi_j} , \quad (121)$$

where ξ_c and ξ_j are the input CNR and JNR on a single element for a single pulse. Let $\xi_i = \xi_c + \xi_j$ be the input interference-to-noise ratio. Typically the clutter and jamming are very large, so $SINR_{in}$ is a very small quantity.

The SINR improvement factor, I_{SINR} , is defined as

$$I_{SINR}(\varpi) = \frac{SINR(\varpi)}{SINR_{in}} . \quad (122)$$

A convenient general rule can be derived for cases where the interference is large and the STAP algorithm provides near optimum performance. In this case,

$$I_{SINR} = MN(1 + \xi_i) \approx MN\xi_i . \quad (123)$$

Thus the SINR improvement factor is typically large, and it increases as the interference becomes stronger. SINR improvement factor includes not only the amount of interference rejection but also the coherent gain on target due to both receive beamforming and Doppler filtering.

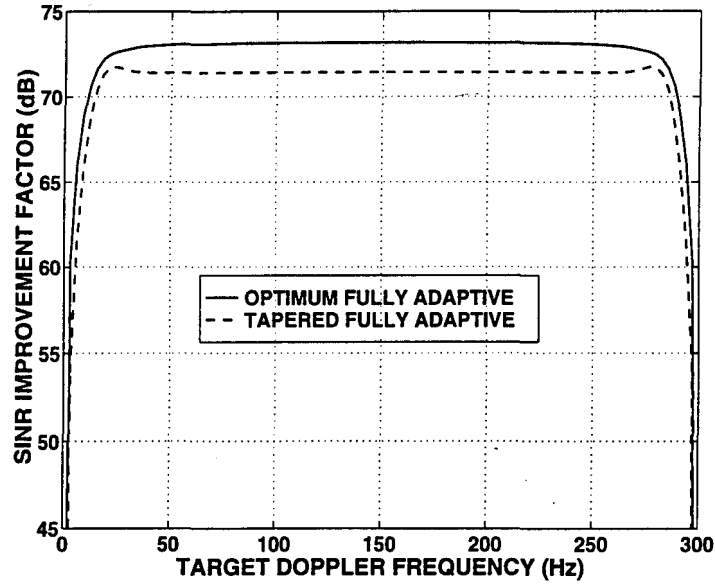


Figure 28. SINR improvement factor for the optimum and tapered fully adaptive STAP.

Figure 28 shows the SINR improvement factors for the example scenario. The input interference-to-noise ratio is 48.1 dB, with 47 dB CNR and 38 dB JNR. Over the center of the Doppler space the optimum SINR achieves an SINR improvement of 73.2 dB. Note that the approximate relationship does apply ($73.2 \approx 48.1 + 25.1$). These curves are scaled replicas of the SINR plots of Figure 25.

Alternative definitions of SINR improvement factor are also possible. For instance, one may wish to express STAP performance with respect to some nonadaptive or reference processing scheme. In this case the corresponding improvement factor is given by

$$I_{SINR}(\varpi) = \frac{SINR(\varpi)}{SINR_{ref}(\varpi)} , \quad (124)$$

where $SINR_{ref}(\varpi)$ denotes the performance of the reference processing scheme.

3.4.5 Minimum Detectable Velocity and Usable Doppler Space Fraction

The SINR performance as a function of Doppler can be used to derive figures of merit that describe the velocity coverage provided by a space-time processing algorithm. To begin, define the acceptable SINR performance to be an SINR loss of $L_{SINR} = x$. With $x = -12$ dB, acceptable performance is at least 50% of the maximum detection range. Similarly, $x = -5$ dB defines acceptable to be at least 75% of the radar's noise-limited detection range.

The minimum detectable velocity (MDV) is defined as the velocity closest to that of the mainlobe clutter at which acceptable SINR loss is achieved. Let $f_L(x)$ and $f_U(x)$ be the Doppler frequencies below and above the mainlobe clutter Doppler at which the acceptable SINR loss is achieved. Define the minimum detectable Doppler f_{\min} as

$$f_{\min}(x) = \frac{1}{2}(f_U(x) - f_L(x)) , \quad (125)$$

which is equal to one-half of the width of the mainlobe clutter notch. The minimum detectable velocity is then defined as

$$MDV(x) = \frac{\lambda}{2} f_{\min} . \quad (126)$$

A target whose velocity differs from the mainlobe clutter velocity by less than the MDV falls far enough into the clutter notch that acceptable performance is not achieved.

Another important quantity is the fraction of the Doppler space over which acceptable performance is achieved. This metric will be called the *usable Doppler space fraction* (UDSF). This metric is useful in choosing multiple PRFs to cover a velocity regime larger than the unambiguous velocity of a single CPI. The presence of antenna crab and backlobe clutter may cause the space-time processor to place nulls at Doppler frequencies other than that of mainlobe clutter. In this case, the MDV alone is insufficient to determine the amount of Doppler space where SINR performance is still acceptable. The UDSF may be thought of as the probability that the SINR loss is greater than the acceptable value, assuming that the target Doppler frequency is uniformly distributed over the Doppler space. Now define $F_L(x) = \Pr\{L_{SINR}(\varpi) \leq x\}$ to be the cumulative distribution of the SINR loss. The UDSF is then given by

$$UDSF(x) = 1 - F_L(x) . \quad (127)$$

For many cases where the SINR loss exhibits only a null on the mainlobe clutter, the simpler expression

$$UDSF(x) = 1 - \frac{2f_{\min}}{f_r} \quad (128)$$

applies.

TABLE 5
MDV and UDSF for fully adaptive STAP

	Optimum	Tapered Fully Adaptive
MDV(−12 dB)	1.2 m/s	2.0 m/s
UDSF(−12 dB)	96.0 %	94.9 %
MDV(−5 dB)	2.7 m/s	3.8 m/s
UDSF(−5 dB)	94.1 %	92.1 %

For the example scenario, the MDV and UDSF can be derived from Figure 27. Table 5 shows the results for two levels of acceptable SINR loss. The requirement for low sidelobes causes a loss in MDV in addition to the loss in SINR performance.

3.5 Sample Support and SMI

In the examples thus far the weights were derived with assumed knowledge of the covariance matrix \mathbf{R}_u . In practice, \mathbf{R}_u must be estimated from the finite data available. Sample matrix inversion (SMI) algorithms are considered [24] where K_e snapshots are used to form the *sample covariance matrix* estimate of \mathbf{R}_u :

$$\hat{\mathbf{R}}_u = \frac{1}{K_e} \sum_{l=1}^{K_e} \chi_l \chi_l^H . \quad (129)$$

Typically, the training samples χ_l cover a range interval surrounding but not including the range gate of interest. The SMI weight vector is then computed from

$$\mathbf{w} = \hat{\mathbf{R}}_u^{-1} \mathbf{g}_t . \quad (130)$$

Because of the covariance *estimate*, the SMI weight vector is suboptimum. An additional performance loss is incurred due to covariance estimation; this loss depends both on the number of samples and on the relationship between \mathbf{g}_t and the target steering vector.

The theory of covariance estimation and its effect on SMI is well developed. In their defining work on SMI [24], Reed, Mallett, and Brennan (RMB) quantified the expected performance of SMI as a function of the number of samples, given a matched steering vector $\mathbf{g}_t = \mathbf{v}_t$. Mismatches between the target steering vector and the vector \mathbf{g}_t used for weight computation may occur by design (because of tapering), because of imperfect knowledge of the target direction (angle and Doppler), or both. Boroson [31] and Kelly [30] have extended the analysis to a general mismatched steering vector. Let $SINR_c$ be the SINR assuming known covariance, from Equation (113). The SINR obtained with the SMI weight vector, Equation (130), is

$$SINR_a = \frac{\sigma^2 \xi_t |\mathbf{g}_t^H \hat{\mathbf{R}}_u^{-1} \mathbf{v}_t|^2}{\mathbf{g}_t^H \hat{\mathbf{R}}_u^{-1} \mathbf{R}_u \hat{\mathbf{R}}_u^{-1} \mathbf{g}_t}, \quad (131)$$

a random quantity depending on the snapshots used for the covariance estimate. Define a new random variable

$$\rho = \frac{SINR_a}{SINR_c} \quad (132)$$

to be the loss in performance, relative to that with perfect covariance knowledge, due to covariance estimation. Under the assumptions of i.i.d. Gaussian snapshots, it has been shown by Boroson and by Kelly that the expected value of this loss is

$$E\{\rho\} = \frac{1}{K_e + 1} \left[K_e + 1 - N_{\text{dof}} + \frac{1}{\rho_c} \right], \quad (133)$$

where N_{dof} is the weight vector dimension. The quantity ρ_c above is defined by

$$\rho_c = \frac{|\mathbf{g}_t^H \mathbf{R}_u^{-1} \mathbf{v}_t|^2}{(\mathbf{g}_t^H \mathbf{R}_u^{-1} \mathbf{g}_t)(\mathbf{v}_t^H \mathbf{R}_u^{-1} \mathbf{v}_t)} = \frac{SINR_c}{SINR_o}. \quad (134)$$

and is a measure of the mismatch between \mathbf{g}_t and \mathbf{v}_t ; it is also the loss in performance relative to optimum given perfect covariance knowledge. With a matched steering vector, $\rho_c = 1$ and ρ is a beta random variable [24] whose expected value is given by

$$E\{\rho\} = \frac{K_e + 2 - N_{\text{dof}}}{K_e + 1} . \quad (135)$$

In this case the expected loss is independent of the interference scenario and depends only on the number of samples K_e and the weight vector dimension. From Equation (135), it is found that for effective performance in a stationary environment, between $2N_{\text{dof}}$ and $5N_{\text{dof}}$ independent samples are required for covariance estimation. In the general mismatched steering vector case, the expected loss is also dependent on the interference scenario through ρ_c , although for small taper or direction mismatches, this dependence is weak.

The above theory can be used to include the effects of covariance estimation in any of the SINR-derived performance metrics. For example, the SINR loss of Equation (120) may be modified as follows:

$$\begin{aligned} L_{\text{SINR}}(\varpi, K_e) &= E \left\{ \frac{\text{SINR}_a}{\text{SINR}_{\text{mf}}} \right\} = E \left\{ \frac{\text{SINR}_a}{\text{SINR}_c} \right\} \left(\frac{\text{SINR}_c}{\text{SINR}_{\text{mf}}} \right) \\ &= \rho(K_e) L_{\text{SINR}}(\varpi, \infty) , \end{aligned} \quad (136)$$

where $L_{\text{SINR}}(\varpi, \infty)$ is the SINR loss with known covariance [Equation (120)].

To illustrate the impact of covariance estimation, consider a fully adaptive SMI approach with the example system, where $N_{\text{dof}} = MN = 324$. Figure 29 shows the expected SINR loss from Equation (136). A matched steering vector (optimum fully adaptive) is assumed. With only $K_e = 324$ samples, a performance loss of more than 20 dB is incurred due to insufficient data for good covariance estimation. At least $2N_{\text{dof}} - 3 = 645$ samples are required to achieve within 3 dB of the known covariance result.

By reducing the adaptive weight dimensionality, the performance with a given amount of data can be dramatically improved. The sample support issue is made more difficult by the nonstationary nature of real radar clutter and jamming. Nonhomogeneity in range, combined with the clutter power and elevation angle dependence on range, all effectively reduce the number of range gates over which the clutter scenario is effectively stationary. The range interval covered by a given number of samples depends on the instantaneous bandwidth of the radar. As the bandwidth becomes smaller, there are fewer samples available to span a given range interval or, conversely, the larger the range extent covered by a given number of samples. The need for adequate covariance estimation is one major factor motivating reduced-dimension STAP algorithms. Most of the analysis of this report is done assuming known covariance but with the underlying purpose of reducing the dimensionality as much as possible. This will ensure that performance with limited data is close to optimum.

It should be understood that in Equation (133), no target signal was assumed present in the snapshots used to form the SMI weight vector. This assumption is critical to the convergence result of Equation (133). It is valid for most radar system scenarios, where the target signals are small and confined to a single range gate, and especially where the training method excludes the range gate under test from the training set. The presence of a strong target signal can dramatically increase the number of samples required for a specified level of performance. This issue was first studied by Miller [22] and more recently investigated by Feldman [32].

3.6 Computational Complexity

The fully adaptive weight vector of Equation (104) is the solution to a linear system of equations whose dimension is MN . For modest-size radars, the product MN may vary from several hundred to several thousand, depending on the antenna array length and the CPI size. Computing a single weight vector requires on the order of $(MN)^3$ operations. Typically, M separate weight vectors are computed to form a bank of filters to cover the Doppler space at each receive beam pointing angle. For radars with a narrow transmit beam, two or three (sum, difference, blanking) receive beams are typically formed with each CPI, yielding a total of $3M$ space-time weight vectors for each CPI. This process is repeated for each training set used to cover the range interval of the radar.

The sheer computational load required to solve the large systems of linear equations required by fully adaptive STAP, in the time necessary for real-time radar operation, is simply beyond the capabilities of current digital processor technology. This is another factor that makes fully adaptive STAP impractical and, along with sample support considerations, motivates research into reduced-dimension STAP algorithms. The next chapters investigate ways to break the fully adaptive problem into smaller dimension adaptive problems that can be solved with both reasonably sized processors and a limited amount of training data.

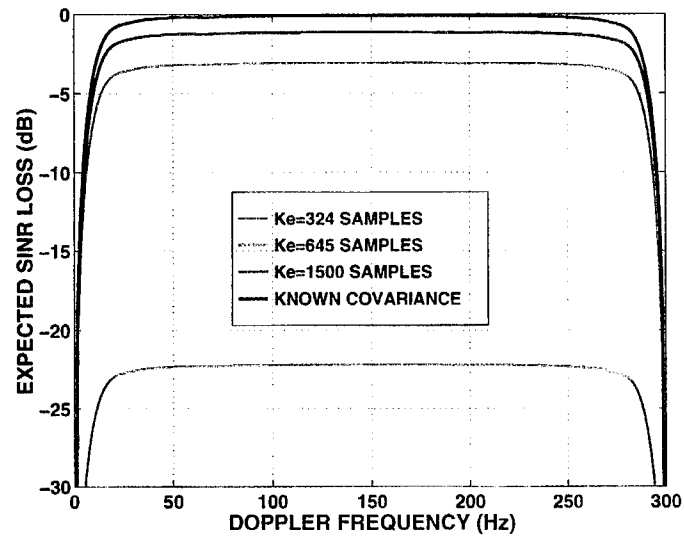


Figure 29. Expected SINR loss for SMI with matched steering vector.

4. PARTIALLY ADAPTIVE STAP

4.1 Introduction

The fully adaptive approaches considered in the last chapter are impractical for reasons of computational complexity and the sample support required for weight training.⁸ This chapter considers reduced-dimension or partially adaptive STAP algorithms. The objective is to break a prohibitively large problem down into a number of smaller, more manageable adaptive problems while achieving near-optimum performance. A generic partially adaptive structure is presented, consisting of a data transformation or preprocessor matrix followed by reduced-dimension adaptive processing. The theory of partially adaptive nulling in the presence of low rank interference is reviewed. Conditions for which a reduced-dimension processor, with known covariance, provides performance equal to fully adaptive STAP are derived. **The reduced-dimension processor may actually provide better performance with limited data because it will incur much less estimation loss.** These results are then applied to the STAP problem. A taxonomy is developed of partially adaptive algorithms that are classified according to the domain in which adaptive weighting occurs or, equivalently, by the type of preprocessing done first.

4.2 A Generic Architecture

A partially adaptive processor takes a large set of input signals, transforms them to a relatively small number of signals, and then solves a reduced-dimension adaptive filtering problem with the transformed data. A general architecture that describes this process is shown in Figure 30. The input data is an MN -dimensional space-time snapshot.

This data is transformed to a new $D \times 1$ vector $\tilde{\chi}$ by means of an $MN \times D$ preprocessor matrix \mathbf{T} ,

$$\tilde{\chi} = \mathbf{T}^H \chi. \quad (137)$$

Throughout this chapter the tilde symbol, \sim , denotes a quantity (data, steering vector, covariance, etc.) in the reduced-dimension space of the data after preprocessing. In general, the transformed data may be decomposed as

$$\tilde{\chi} = \alpha \tilde{\mathbf{v}}_t + \tilde{\chi}_u, \quad (138)$$

⁸The development in this section has benefited greatly from discussions with A. Steinhardt, S. Krich, E. Baranoski, and D. Marshall.

where $\tilde{\mathbf{v}}_t = \mathbf{T}^H \mathbf{v}_t$ is the transformed target steering vector and $\tilde{\chi}_u = \mathbf{T}^H \chi_u$ is the interference-plus-noise component.

After data transformation, the $D \times 1$ adaptive weight vector

$$\tilde{\mathbf{w}} = \tilde{\mathbf{R}}_u^{-1} \tilde{\mathbf{g}}_t \quad (139)$$

is computed, where

$$\tilde{\mathbf{R}}_u = E \{ \tilde{\chi}_u \tilde{\chi}_u^H \} = \mathbf{T}^H \mathbf{R}_u \mathbf{T} \quad (140)$$

is the $D \times D$ covariance matrix of the transformed data and $\tilde{\mathbf{g}}_t$ is a $D \times 1$ desired response or target steering vector. Given a desired response \mathbf{g}_t for a fully adaptive processor, the new desired response is defined as

$$\tilde{\mathbf{g}}_t = \mathbf{T}^H \mathbf{g}_t. \quad (141)$$

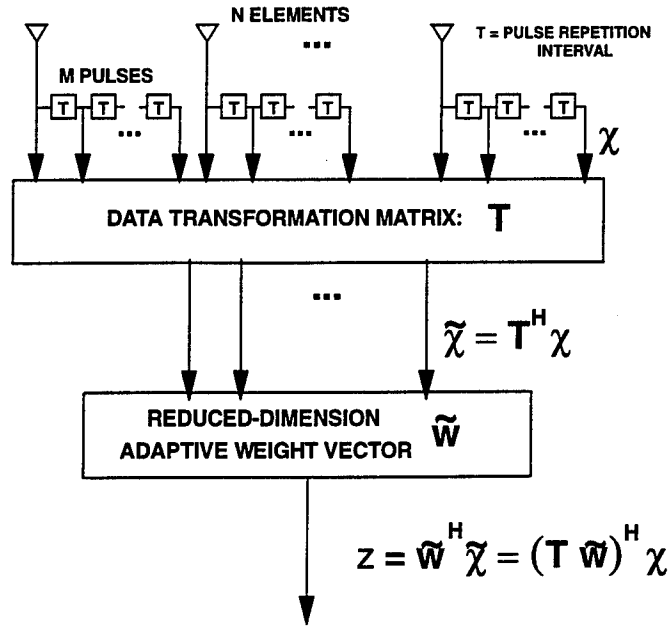


Figure 30. A generic partially adaptive STAP architecture.

Applying the computed weights yields the final output

$$z = \tilde{\mathbf{w}}^H \tilde{\chi} = (\mathbf{T}\tilde{\mathbf{w}})^H \chi . \quad (142)$$

The partially adaptive weight vector $\tilde{\mathbf{w}}$ may be projected back to the full dimensionality to produce the *composite* $MN \times 1$ vector

$$\mathbf{w}_{\text{pa}} = \mathbf{T}\tilde{\mathbf{w}} , \quad (143)$$

which represents the cascade of both the preprocessor and the adaptive weighting. The composite weight vector is useful for comparing the performance of fully and partially adaptive processors.

The major issue to be addressed is the design of \mathbf{T} . If \mathbf{T} is invertible, there is no loss of performance [23] but no reduction in dimensionality, either. It is desirable to design \mathbf{T} with $D \ll MN$ as small as possible while at the same time achieving as close to fully adaptive (optimum) performance as possible.

4.3 Exploiting Low-Rank Interference

It was shown in Chapter 2 that the interference covariance matrix is or in many cases may be approximated as low rank. This section will show that low-rank interference may be exploited to achieve significant reduction in the adaptive problem dimensionality with little or no sacrifice in performance relative to the fully adaptive case. The general theory of adaptive nulling with reduced rank interference will be presented, followed by its application to the STAP problem.

Assume that the interference-plus-noise covariance matrix consists of two components,

$$\mathbf{R}_u = \mathbf{R}_1 + \mathbf{R}_2 . \quad (144)$$

Assume that \mathbf{R}_1 is full rank and is either known a priori or can be estimated separately from \mathbf{R}_2 . This latter condition requires the existence of data without the \mathbf{R}_2 component present. Second, assume that \mathbf{R}_2 is low rank, i.e.,

$$r_2 = \text{rank}(\mathbf{R}_2) < MN , \quad (145)$$

and has spectral decomposition

$$\mathbf{R}_2 = \mathbf{E}_2 \mathbf{\Lambda}_2 \mathbf{E}_2^H, \quad (146)$$

where $\mathbf{\Lambda}_2$ is an $r_2 \times r_2$ diagonal matrix of the nonzero eigenvalues of \mathbf{R}_2 and \mathbf{E}_2 is the $MN \times r_2$ matrix of the corresponding eigenvectors. In the simplest case, think of \mathbf{R}_1 being noise and \mathbf{R}_2 as being low rank interference. In general, the preprocessor \mathbf{T} may decrease the interference rank but it cannot increase it.

Under these assumptions, the fully adaptive weight vector from Equation (106) is derived by substituting (146) into Equation (144) and applying the matrix inversion lemma:

$$\mathbf{w}_{fa} = \mathbf{R}_1^{-1} \mathbf{g}_t - \mathbf{R}_1^{-1} \mathbf{E}_2 \left(\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \mathbf{\Lambda}_2^{-1} \right)^{-1} \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{g}_t. \quad (147)$$

Close inspection of Equation (147) shows that \mathbf{w}_{fa} may also be written as the matrix-vector product

$$\mathbf{w}_{fa} = \begin{bmatrix} \mathbf{R}_1^{-1} \mathbf{g}_t, & \mathbf{R}_1^{-1} \mathbf{E}_2 \end{bmatrix} \mathbf{c}, \quad (148)$$

where \mathbf{c} is the $(r_2 + 1) \times 1$ coefficient vector

$$\mathbf{c} = \begin{bmatrix} 1; & - \left(\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \mathbf{\Lambda}_2^{-1} \right)^{-1} \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{g}_t \end{bmatrix}. \quad (149)$$

Equation (148) shows that the fully adaptive weight vector lies entirely in an $(r_2 + 1)$ -dimensional subspace comprising the desired response and the \mathbf{E}_2 interference subspace:⁹

$$\mathbf{w}_{fa} \in \text{span} \left\{ \mathbf{R}_1^{-1} [\mathbf{g}_t, \mathbf{E}_2] \right\}. \quad (150)$$

For example, with $\mathbf{R}_1 = \mathbf{I}$ and the matched steering vector $\mathbf{g}_t = \mathbf{v}_t$, Equation (150) shows that the optimum weight vector lies in the target-plus-interference subspace [32].

⁹ Assume the desired response vector (target steering vector) has a component orthogonal to the \mathbf{R}_2 interference subspace; if not, no algorithm will work.

This observation suggests the partially adaptive architecture of Figure 31(a), where the data is first sent through transform

$$\mathbf{T} = \mathbf{R}_1^{-1} \mathbf{B} , \quad (151)$$

whose outputs are then adaptively combined with a weight vector $\tilde{\mathbf{w}}$. Here \mathbf{B} is an $MN \times D$ “beam” matrix to be determined. The following theorem provides a condition for the optimum design of \mathbf{B} .

Theorem 2 *If the matrix \mathbf{B} is designed to satisfy*

$$\text{span} \{[\mathbf{g}_t, \mathbf{E}_2]\} \subset \text{span} \{\mathbf{B}\} , \quad (152)$$

then the partially adaptive processor of Figure 31(a) achieves performance, with known covariance, equal to that of the fully adaptive processor.

Another way of stating the condition Equation (152) is to require

$$\mathbf{P}_B [\mathbf{g}_t, \mathbf{E}_2] = [\mathbf{g}_t, \mathbf{E}_2] , \quad (153)$$

where $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$ is a projection matrix that projects a vector onto the subspace spanned by the columns of \mathbf{B} .

This result should not be surprising given Equation (150); a rigorous proof is provided in Appendix 2. To illustrate the power of Theorem 2, consider $\mathbf{R}_1 = \mathbf{I}$ and $\mathbf{g}_t = \mathbf{v}_t$. By choosing \mathbf{B} to span both the target and interference subspaces, optimum performance can be achieved with only $r_2 + 1$ adaptive degrees of freedom. The matrix \mathbf{B} can be viewed as a target “beam” plus a set of interference cancellation “beams.” The interference cancellation beams, because they span the interference subspace, produce output signals with maximum interference power. Therefore the interference is suppressed with minimum distortion to the target (quiescent) beam: optimum SINR results! Practically, the advantages are that one only needs sufficient data to support a $D \times D$ covariance estimation, where $D = (r_2 + 1)$. Given a fixed amount of data, the performance of the partially adaptive approach may in fact be better than fully adaptive, where losses due to imperfect covariance estimation will be much greater. Theorem 2 also applies when $\mathbf{g}_t \neq \mathbf{v}_t$ and shows that by including the desired response in the beam matrix, tapered fully adaptive performance (near-optimum SINR and adapted pattern) can be achieved with a reduced dimension processor.

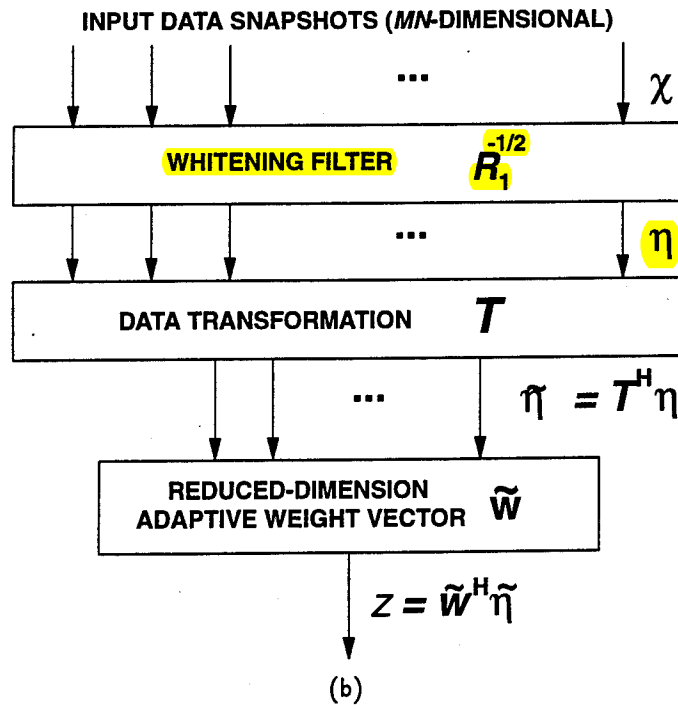
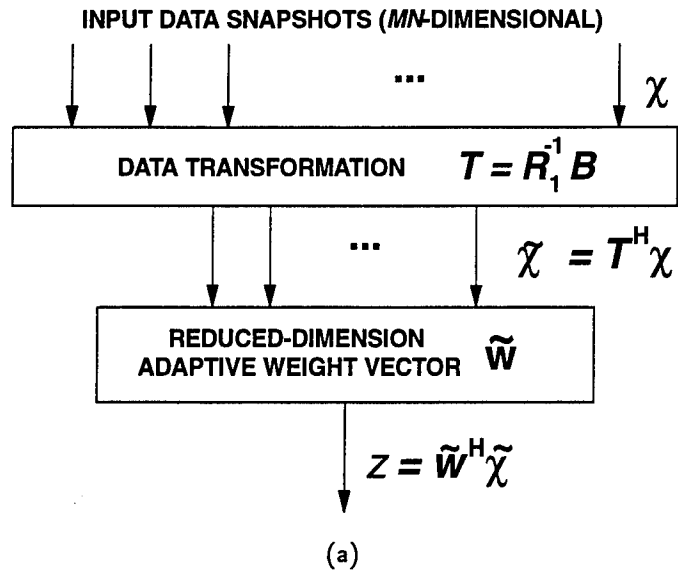


Figure 31. (a) Reduced dimension adaptive processing with low-rank interference. (b) Whitening filter viewpoint.

An alternate form for a partially adaptive processor, shown in Figure 31(b), is often useful. Here the preprocessor is divided into two stages. The first portion is a whitening filter that cleanses the \mathbf{R}_1 component of the data. The matrix \mathbf{T} then projects the whitened data into a lower dimension space, and the projected data is adaptively weighted with $\tilde{\mathbf{w}}$ to produce the output.

The whitening filter is chosen to be the matrix $\mathbf{R}_1^{-1/2}$, which exists because \mathbf{R}_1 is nonsingular.¹⁰ The whitening filter is known because it is assumed that \mathbf{R}_1 is known or can be estimated. Let the whitening filter output be $\boldsymbol{\eta} = \mathbf{R}_1^{-1/2} \boldsymbol{\chi}$ and let $\mathbf{h}_t = \mathbf{R}_1^{-1/2} \mathbf{g}_t$ be the whitened desired response. The covariance matrix (interference plus noise) of the whitened data is then

$$\mathbf{S}_u = \mathbf{I} + \mathbf{S}_2 , \quad (154)$$

$$\mathbf{S}_2 = \mathbf{R}_1^{-1/2} \mathbf{R}_2 \mathbf{R}_1^{-1/2} . \quad (155)$$

Since the whitening filter is invertible, it does not alter the rank of any data component. The interference \mathbf{S}_2 has rank r_2 and spectral decomposition

$$\mathbf{S}_2 = \mathbf{F}_2 \boldsymbol{\Gamma}_2 \mathbf{F}_2^H . \quad (156)$$

The eigenvectors before and after whitening are related by

$$\text{span} \{ \mathbf{F}_2 \} = \text{span} \{ \mathbf{R}_1^{-1/2} \mathbf{E}_2 \} , \quad (157)$$

which follows by examining the nullspace of \mathbf{S}_2 from Equations (155) and (156). The main issue is how to design the preprocessor \mathbf{T} after whitening. The following variation of Theorem 2 provides a condition for the optimum design of \mathbf{T} in Figure 31(b).

Theorem 3 *If the preprocessor \mathbf{T} is designed to satisfy*

$$\text{span} \{ [\mathbf{h}_t , \mathbf{F}_2] \} \subset \text{span} \{ \mathbf{T} \} , \quad (158)$$

then the partially adaptive processor of Figure 31(b) achieves performance, with known covariance, equal to that of the fully adaptive processor.

¹⁰From a computational viewpoint, it may be better to construct the whitening filter from the Cholesky factor of \mathbf{R}_1 .

Another way of stating the condition in Equation (158) is to require

$$\mathbf{P}_T [\mathbf{h}_t, \mathbf{F}_2] = [\mathbf{h}_t, \mathbf{F}_2], \quad (159)$$

where $\mathbf{P}_T = \mathbf{T}(\mathbf{T}^H \mathbf{T})^{-1} \mathbf{T}^H$ is the projection matrix for \mathbf{T} .

The proof takes two steps. First, the whitening filter is invertible and therefore does not alter the performance. Second, the processor after whitening is a special case of Figure 31(a), with $\mathbf{R}_1 = \mathbf{I}$. Theorem 2 is applied and the proof is complete. An important result from Theorem 3 is that in the presence of colored noise, the pertinent interference subspace is the subspace of the whitened data. The condition in Equation (158) is in terms of the whitened eigenvectors and response vector. It can also be stated, utilizing Equation (157), as

$$\text{span} \{ \mathbf{R}_1^{-1/2} [\mathbf{g}_t, \mathbf{E}_2] \} \subset \text{span} \{ \mathbf{T} \}. \quad (160)$$

This observation also follows by noting that the processors of Figure 31 are identical if, in Figure 31(b), $\mathbf{T} = \mathbf{R}_1^{-1/2} \mathbf{B}$.

Theorems 2 and 3 are very powerful. They provide the foundation for many partially adaptive approaches. Beamspace nulling and subspace-based adaptive processing are examples of approaches that can be derived from this theory. Unfortunately, the interference subspace is typically unknown a priori, so the optimum a priori design of the preprocessor is impossible. Limited prior knowledge may allow one to design a nonadaptive transform that encompasses the range of expected interference scenarios. In this case the dimensionality is reduced, but not to the degree possible with perfect information. With no a priori knowledge, the interference subspace must be estimated data-adaptively. For large problem sizes, the process of subspace estimation may also be computationally prohibitive. For this reason it is useful to also view the transform \mathbf{T} in Figure 30 as a filter that nonadaptively suppresses a portion of the interference. In so doing, both a reduction in data dimensionality and a reduction in effective interference rank may be achieved. Theorems 2 or 3 may then be applied to gain further dimensionality reduction prior to the final SMI weight computation.

4.4 Application to STAP

The interference-plus-noise component of the data typically has three mutually uncorrelated components: clutter, jamming, and noise. It was shown in Chapter 2 that the space-time covariance matrix is given by

$$\mathbf{R}_u = \mathbf{R}_c + \mathbf{R}_j + \mathbf{R}_n. \quad (161)$$

The noise, being typically dependent only on the receiver characteristics, is well known a priori and full rank. Much less is known a priori about the clutter and jamming, except that they are of low rank. Therefore, one application of the above theory is to identify

$$\mathbf{R}_1 = \mathbf{R}_n \quad \text{and} \quad \mathbf{R}_2 = \mathbf{R}_c + \mathbf{R}_j \quad (162)$$

and to use an architecture such as Figure 31(a). With reasonable assumptions such as those of Chapter 2, \mathbf{R}_n is white both spatially and temporally white, so that the processors of Theorems 2 and 3 default to the same architecture. In this case adaptive nulling simultaneously rejects both clutter and jamming. **Theorem 2 states that fully adaptive performance can be achieved if the preprocessor matrix contains the span of the combined clutter-plus-jamming subspace.**

In some cases, limited a priori information may be exploited to obtain clutter-free data from which an estimate of the jamming-plus-noise component can be made. For example, it was seen in Chapter 2 that depending on the radar waveform, there may be clutter-free range gates (beyond the horizon with low PRF) or clutter-free Doppler frequencies (with medium or high PRF). Another way to derive clutter-free samples of the signal environment is for the radar to operate in a passive receive-only mode for a brief time prior to the transmission of the first pulse of a CPI [33]. If the jamming can be assumed stationary enough so that a jammer-plus-noise estimate obtained in these ways is accurate for the time interval that also contains the target and clutter, a second application of the above theory results. In this case, the association

$$\mathbf{R}_1 = \mathbf{R}_j + \mathbf{R}_n \quad \text{and} \quad \mathbf{R}_2 = \mathbf{R}_c \quad (163)$$

is made. **The partially adaptive structures of Theorems 2 and 3 split the interference cancellation into two steps. The first step optimally nulls the jamming and projects the data into a lower dimension, then the second stage adaptively suppresses the clutter.** Because the jamming is typically not known a priori, the first step will also require adaptivity. An architecture of this form is called a **two-step nulling (TSN), or sequential nulling architecture** [34, 37]. At first glance it appears that a large dimensional problem must be solved in the first step, as the space-time jamming covariance matrix is of dimension MN . However, both the block diagonal nature of \mathbf{R}_j and the low-rank nature of the spatial jammer covariance matrix may be used to advantage; in essence, jammer nulling requires only spatial adaptation (small dimension) whose weights will at most need to be recomputed for each PRI. When it is feasible, two-step nulling may produce significant reductions in required adaptive degrees of freedom.

The large dimensionality of the fully adaptive space-time processing problem and the fact that the interference is mostly unknown a priori lead toward an architecture whose first step provides nonadaptive filtering to reduce the dimensionality prior to adaptive processing. This process is illustrated in Figure 32, which provides a final block diagram for a partially adaptive space-time processor. After nonadaptive filtering, assume the dimensionality has been reduced to a point that weight computation is feasible. It may be beneficial to utilize Theorems 2 and 3, in which case the next step is a subspace projection that provides additional dimensionality reduction. Finally, a weight vector is computed by solving an SMI problem with the final reduced-dimension data. This architecture encompasses both simultaneous and two-step nulling approaches. With some approaches, a final stage of nonadaptive processing may be required, as indicated by the dashed block. With this framework in mind, a number of approaches will be considered, classified by the type of transformation performed in the T_1 block of Figure 32.

Many preprocessor types are possible. The input data originates from a sampling in time from the pulses of the CPI and in space at the locations of the antenna array elements. Combinations of spatial and temporal filtering compose the different types of transforms. A convenient classification, in terms of four basic categories, is shown in Figure 33. Each quadrant shows a box representing the data for a single range gate after a different type of transform. For example, with no transformation, a matrix of element by PRI data composes a space-time snapshot. One way to reduce the problem dimensionality in this domain would be to utilize all elements but only a small number of PRIs. Algorithms of this type are referred to as *element-space pre-Doppler*, because full element adaptivity is retained and Doppler processing will be done after adaptation.

Spatial filtering (beamforming or subarraying) may be performed on the element outputs of each PRI prior to adaptation. In so doing the space-time snapshot is transformed into a new snapshot consisting of angle beam-by-PRI data. Algorithms that perform beamforming prior to adaptation are referred to as *beam-space algorithms*. Beam-space algorithms have long been utilized for reducing DOF in spatial nulling and angle estimation applications. Beam-space approaches can also reduce the required degrees of freedom for canceling the clutter seen by an airborne radar. Beam-space approaches are also amenable to two-step nulling. STAP approaches that adapt with beamformed data from a small subset of PRIs will be termed *beam-space pre-Doppler* algorithms. This class is represented by the lower left block of Figure 33.

In an analogous fashion, temporal filtering (Doppler processing) may be performed on the data from each array element prior to adaptive processing. This operation transforms the space-time snapshot into a snapshot of Doppler bin and element data. STAP algorithms that operate on a subset of this data are termed *element-space post-Doppler* algorithms, as the adaptation occurs after Doppler processing. *Post-Doppler algorithms will require solving a separate adaptive problem for each Doppler bin.* Because the clutter Doppler frequency depends on angle, Doppler filtering serves to localize clutter in angle. Modern radar systems are very stable from pulse to pulse, making possible nonadaptive Doppler filters with extremely low sidelobes (> 80 dB). Doppler filtering can therefore isolate the clutter interfering with a target to just those angular regions of the clutter ridge corresponding to the Doppler filter mainlobe. For this reason, *post-Doppler algorithms can significantly reduce the number of adaptive degrees of freedom required.*

Finally, the lower right corner of Figure 33 represents algorithms where both spatial and temporal preprocessing of the data is done prior to adaptation. Think of such processing as cascading a beamformer on each PRI with a Doppler filter bank on each beam. *A subset of the resultant angle-Doppler filter outputs will then be adaptively processed to produce the final outputs.* This class is therefore called *beam-space post-Doppler algorithms*. One example of a beam-space post-Doppler algorithm is to perform a two-dimensional DFT of the input data prior to adaptations. Many other examples exist. With beam-space post-Doppler, the idea is to localize the interference both temporally and spatially prior to adaptation so that very few outputs need to be combined adaptively.

The next two chapters will study the algorithms of each of these four classes. Chapter 5 considers element space approaches, and Chapter 6 covers beam-space algorithms.

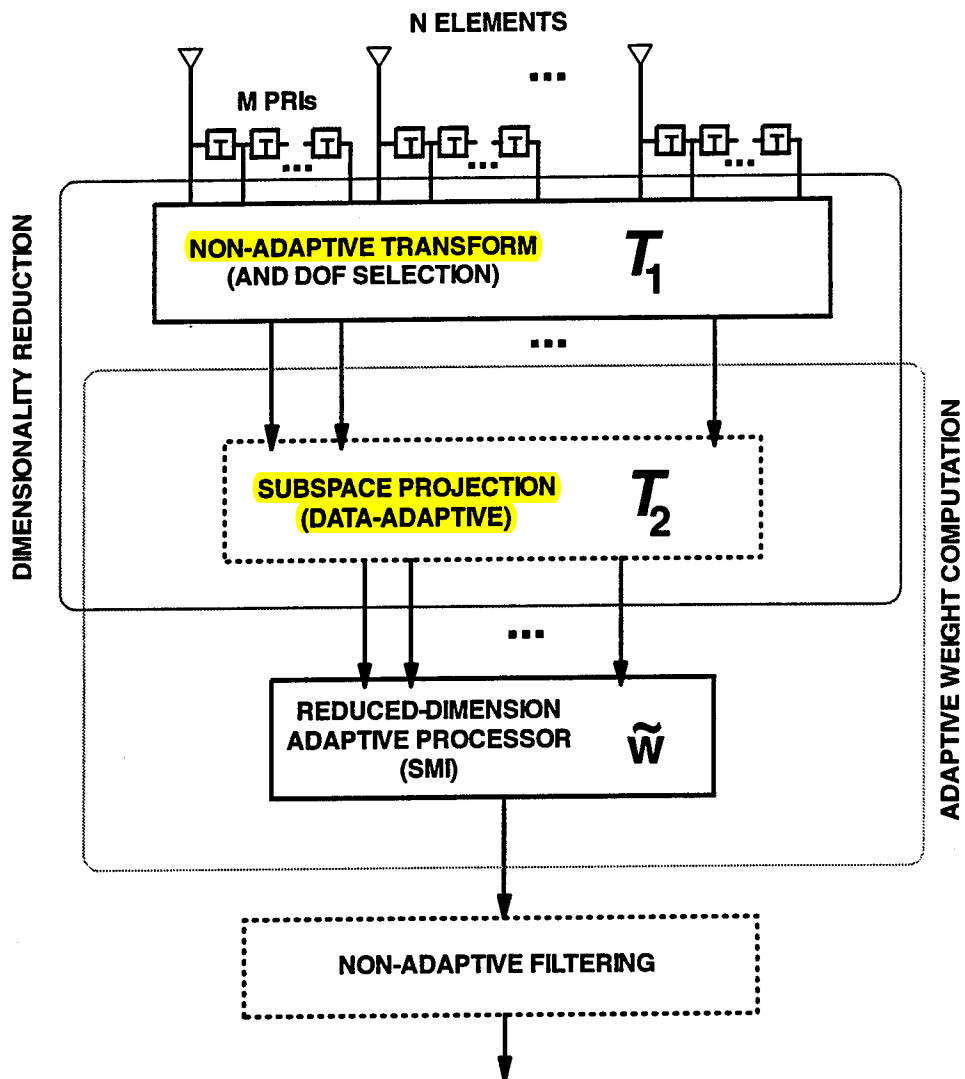


Figure 32. Another block diagram for reduced-dimension space-time adaptive processing.

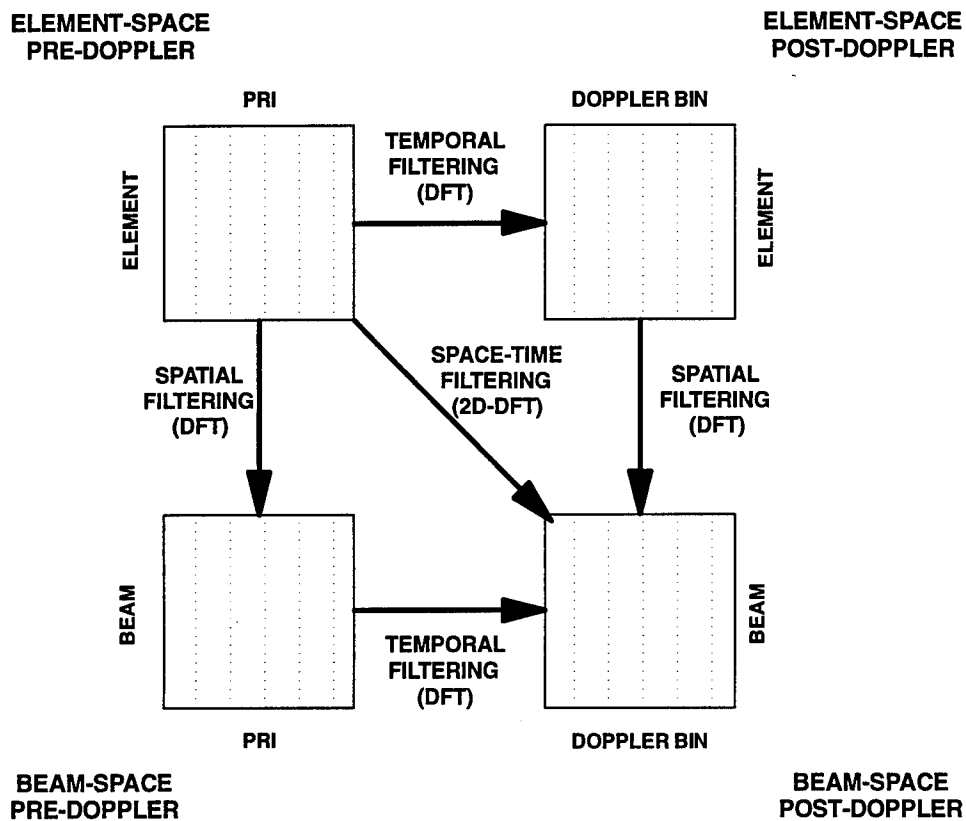


Figure 33. A taxonomy of reduced-dimension STAP algorithms, classified by the type of nonadaptive transformation applied to the CPI data.

5. ELEMENT-SPACE STAP

5.1 Introduction

This chapter considers element-space STAP architectures whereby every element is adaptively weighted. Element-space approaches retain full spatial dimensionality but reduce the overall problem size by reducing the number of temporal degrees of freedom prior to adaptation. Full element-space adaptivity provides the flexibility to handle a completely unknown jamming environment and also the potential for effective clutter cancellation at all angles. Because they retain full spatial adaptivity, element-space approaches are also robust to channel mismatch errors. Typically, element-space architectures aim to suppress clutter and jamming simultaneously. By their nature, element-space techniques may be impractical for very large arrays. Element-space pre-Doppler STAP will be considered first. In this approach, adaptive weighting is done a few pulses at a time, with final Doppler filtering performed after space-time adaptation. Then, element-space post-Doppler techniques will be considered, where a separate Doppler filter bank is applied to the pulses from each element. The outputs of a small number of filters on each element form the degrees of freedom for low dimension space-time adaptation. The issue of Doppler filter design is addressed, and a number of different post-Doppler architectures are discussed.

5.2 Element-Space Pre-Doppler STAP

The first algorithm class considered is element-space pre-Doppler STAP. The basic idea is to adaptively combine the data from only a few (typically 2 or 3) pulses at a time rather than all the pulses of the CPI. Utilizing more than one pulse provides the temporal adaptivity required for clutter cancellation, while retaining full spatial adaptivity provides a means to handle jamming simultaneously. Adaptive processing is then followed by a fixed (nonadaptive) Doppler filter bank that provides coherent integration over the full CPI and the means for velocity estimation. This architecture was first put forth by Brennan [18], who termed it the “adapt then filter” approach.

To be more precise, define a *sub-CPI* to be a subset of the CPI comprising K successive pulses. Given M pulses in a CPI, there are $M' = M - K + 1$ sub-CPIs. Number them from 0 to $M' - 1$, with the 0th sub-CPI consisting of pulses $0, \dots, K - 1$, the p th sub-CPI consisting of pulses $p, \dots, p + K - 1$, and so on. A KN -dimensional space-time adaptive weight vector is computed and applied to each sub-CPI. The M' output pulses are subsequently passed through a standard Doppler filter bank. Because successive sub-CPIs utilize overlapping sets of pulses, outputs from successive sub-CPIs will be temporally correlated.

Figure 34 shows two block diagrams for element-space pre-Doppler: Figure 34(a) shows a sub-CPI portion, while Figure 34(b) shows a full CPI depiction. Now, consider the 0th sub-CPI, consisting of pulses $0, \dots, K - 1$. Recall that \mathbf{x}_m specifies the $N \times 1$ spatial snapshot for the m th PRI.

The sub-CPI snapshot vector $\tilde{\chi}$ is the $KN \times 1$ vector

$$\tilde{\chi} = (\mathbf{J} \otimes \mathbf{I}_N)^H \chi, \quad (164)$$

where

$$\mathbf{J} = \begin{bmatrix} \mathbf{I}_K \\ \mathbf{0}_{(M-K) \times K} \end{bmatrix} \quad (165)$$

is an $M \times K$ selection matrix that chooses the first K pulses from the CPI. The notation $\mathbf{0}_{l \times m}$ refers to an $l \times m$ matrix of zeros. For this architecture, the matrix $\mathbf{J} \otimes \mathbf{I}_N$ is the preprocessor of Figure 30. The sub-CPI snapshot contains an interference-plus-noise component $\tilde{\chi}_u = (\mathbf{J} \otimes \mathbf{I}_N)^H \chi_u$ and, if a target is present, the target component

$$\tilde{\chi}_t = \alpha_t (\mathbf{J} \otimes \mathbf{I}_N)^H \mathbf{v}_t = \alpha_t \mathbf{J}^H \mathbf{b}_t \otimes \mathbf{a}_t \quad (166)$$

$$= \alpha_t \tilde{\mathbf{b}}_t \otimes \mathbf{a}_t = \alpha_t \tilde{\mathbf{v}}_t, \quad (167)$$

where $\tilde{\mathbf{v}}_t = \tilde{\mathbf{b}}_t \otimes \mathbf{a}_t$, $\tilde{\mathbf{b}}_t = \tilde{\mathbf{b}}(\varpi_t)$, and $\tilde{\mathbf{b}}(\varpi) = \mathbf{b}(\varpi; K)$ is a K -pulse temporal steering vector. The sub-CPI adaptive weight vector is computed as

$$\tilde{\mathbf{w}} = \tilde{\mathbf{R}}_u^{-1} \tilde{\mathbf{g}}_t, \quad (168)$$

where $\tilde{\mathbf{R}}_u$ is the $KN \times KN$ interference-plus-noise covariance matrix and $\tilde{\mathbf{g}}_t$ is a $KN \times 1$ sub-CPI desired response. Note that the problem size has been reduced by a factor of M/K relative to fully adaptive STAP.

The sub-CPI covariance matrix has all the properties described in Chapter 2, but with K substituted for M . It may be broken up into clutter, jamming, and thermal noise components:

$$\tilde{\mathbf{R}}_u = \tilde{\mathbf{R}}_c + \tilde{\mathbf{R}}_j + \tilde{\mathbf{R}}_n. \quad (169)$$

The sub-CPI jammer covariance matrix has rank, from Equation (50),

$$\tilde{r}_j = K \rho_j. \quad (170)$$

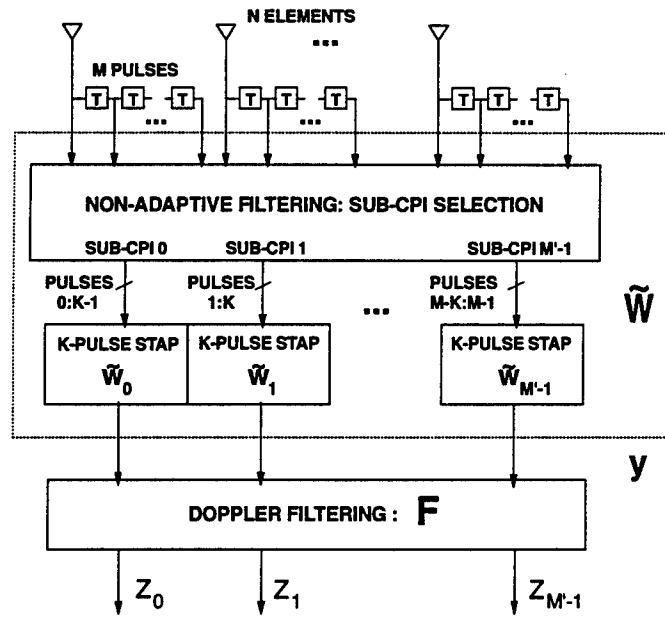
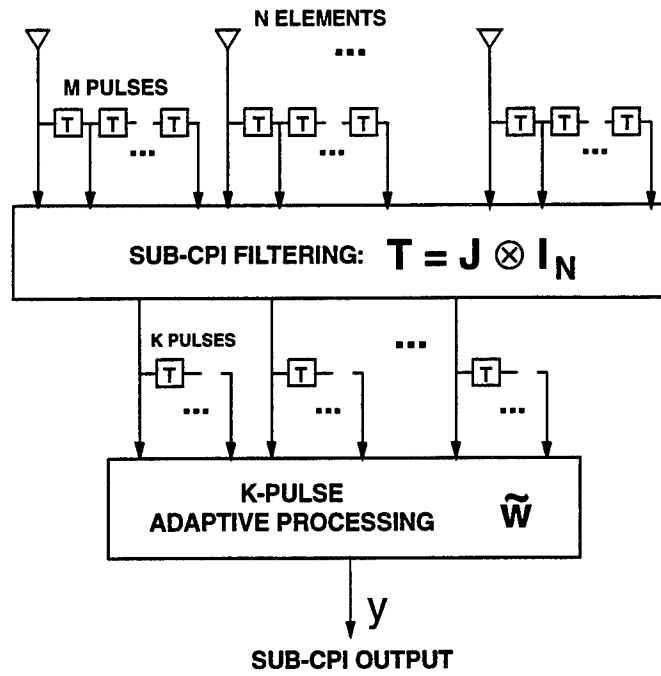


Figure 34. Element-space pre-Doppler STAP. (a) Sub-CPI processing. (b) Full CPI processing.

Similarly, under the conditions for which Brennan's rule applies, the rank of $\tilde{\mathbf{R}}_c$ is

$$\tilde{r}_c = N + (K - 1)\beta . \quad (171)$$

Finally, the sub-CPI noise covariance matrix is both spatially and temporally white. Although the interference within the sub-CPI is still less than full rank, its rank (mainly the clutter) becomes a larger fraction of the snapshot dimension as K becomes smaller.

The desired response for pre-Doppler STAP is chosen with the subsequent Doppler processing in mind. As the outputs of the sub-CPI adaptations will be Doppler filtered, it is desirable for the temporal component of the sub-CPI response to maximally cover the Doppler space. Therefore, a sensible choice is a binomial taper steered to the center of the Doppler space. This choice concentrates all the nulls of the quiescent Doppler response to the Doppler frequency of the mainlobe clutter, and prevents extraneous nulls in Doppler from being placed at other Dopplers where target detection is desired. With this motivation, the sub-CPI desired response is taken to be

$$\tilde{\mathbf{g}}_t = (\tilde{\mathbf{t}}_b \odot \tilde{\mathbf{b}}(\varpi_b)) \otimes (\mathbf{t}_a \odot \mathbf{a}(\vartheta_t)) . \quad (172)$$

In this equation, $\tilde{\mathbf{t}}_b$ is the $K \times 1$ binomial taper and ϖ_b denotes the Doppler frequency in the center of the Doppler space, relative to the mainlobe clutter Doppler. That is, if mainlobe clutter has Doppler ϖ_c , the binomial response will be steered to $\varpi_b = \varpi_c + 1/2$. For example, with $K = 3$ and the mainlobe clutter at zero Doppler, $\varpi_b = 1/2$ and

$$\tilde{\mathbf{t}}_b \odot \tilde{\mathbf{b}}(\varpi_b) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \odot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad (173)$$

is just a 3-pulse MTI canceller impulse response. A spatial taper \mathbf{t}_a is included in Equation (172) for completeness.

Now consider the full CPI. There are several ways in which one might train and apply the weights. In the simplest case, a single KN -dimensional adaptive weight vector is computed and this weight vector is applied to each sub-CPI. Training samples for this single weight computation may be derived from within a single sub-CPI or from multiple sub-CPIs. More generally, one could compute and apply a separate adaptive weight vector for each sub-CPI. After the first sub-CPI, the weights are updated with each subsequent PRI of data. This more rapid updating may be useful to combat a changing interference environment (due to antenna rotation, for example) but also requires that the computations be done more rapidly. Assume the general case of a separate adaptive problem for each sub-CPI from here on. A full CPI block diagram of the architecture was given in Figure 34(b).

The p th sub-CPI snapshot consisting of pulses $p, \dots, p + K - 1$ may be written

$$\tilde{\chi}_p = (\mathbf{J}_p \otimes \mathbf{I}_N)^H \chi, \quad (174)$$

where

$$\mathbf{J}_p = \begin{bmatrix} \mathbf{0}_{p \times K} \\ \mathbf{I}_K \\ \mathbf{0}_{(M-K-p) \times K} \end{bmatrix}. \quad (175)$$

The p th sub-CPI weight vector is given by

$$\tilde{\mathbf{w}}_p = \tilde{\mathbf{R}}_{up}^{-1} \tilde{\mathbf{g}}_t, \quad (176)$$

where $\tilde{\mathbf{R}}_{up}$ is the p th sub-CPI covariance matrix. The desired response $\tilde{\mathbf{g}}_t$ does not change from one sub-CPI to the next. The sub-CPI weight vectors can be decomposed into their spatial beamformer components for each pulse. Let

$$\tilde{\mathbf{w}}_p = \begin{bmatrix} \tilde{\mathbf{w}}_{p,0}; & \tilde{\mathbf{w}}_{p,1}; & \dots; & \tilde{\mathbf{w}}_{p,K-1} \end{bmatrix}, \quad (177)$$

so that $\tilde{\mathbf{w}}_{p,k}$ represents the spatial weight vector for the k th pulse of p th sub-CPI. The output signal from the p th sub-CPI is then

$$y_p = \tilde{\mathbf{w}}_p^H \tilde{\chi}_p = \sum_{k=0}^{K-1} \tilde{\mathbf{w}}_{p,k}^H \mathbf{x}_{p+k}, \quad p = 0 : M' - 1. \quad (178)$$

Assemble the outputs from all sub-CPIs into an $M' \times 1$ vector

$$\mathbf{y} = [y_0; y_1; \dots; y_{M'-1}] = \tilde{\mathbf{W}}^H \chi, \quad (179)$$

where $\tilde{\mathbf{W}}$ is an $MN \times M'$ matrix containing the weights from all sub-CPIs. This matrix is given by

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{w}}_{0,0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \tilde{\mathbf{w}}_{0,1} & \tilde{\mathbf{w}}_{1,0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \tilde{\mathbf{w}}_{0,K-1} & & & \ddots & \mathbf{0} \\ \mathbf{0} & \ddots & & & \tilde{\mathbf{w}}_{M-K,0} \\ \vdots & \ddots & \ddots & & \tilde{\mathbf{w}}_{M-K,1} \\ \mathbf{0} & & \ddots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{w}}_{M-K,K-1} \end{bmatrix} . \quad (180)$$

If the weight vectors for each sub-CPI are equal, $\tilde{\mathbf{W}}$ has the block-Toeplitz structure characteristic of a temporal convolution with a K -pulse space-time filter.

The sub-CPI output signals are processed by a length M' Doppler filter bank. Let $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{M'}]$ be an $M' \times M'$ DFT matrix, and let \mathbf{t}_d be an $M' \times 1$ taper for the Doppler filters. The Doppler filter bank is represented by the matrix

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}_0 & \mathbf{f}_1 & \cdots & \mathbf{f}_{M'-1} \end{bmatrix} = \text{diag}(\mathbf{t}_d) \mathbf{U}^* , \quad (181)$$

where the m th Doppler filter is given by $\mathbf{f}_m = \mathbf{t}_d \odot \mathbf{u}_m^*$. With this convention, the Doppler filter output is

$$\mathbf{z} = \begin{bmatrix} z_0 & z_1 & \cdots & z_{M'-1} \end{bmatrix} = \mathbf{F}^H \mathbf{y} , \quad (182)$$

where the signal $z_m = \mathbf{f}_m^H \mathbf{y}$ is the final output for the m th Doppler bin. The output signal can also be expressed in terms of the full dimension snapshot as

$$z_m = \mathbf{f}_m^H \tilde{\mathbf{W}}^H \boldsymbol{\chi} = \mathbf{w}_m^H \boldsymbol{\chi} , \quad (183)$$

where

$$\mathbf{w}_m = \tilde{\mathbf{W}} \mathbf{f}_m \quad (184)$$

is the m th bin composite weight vector that represents all of the processing steps (adaptive and fixed) involved in producing the final output.

The algorithm's SINR performance may now be computed. Assume the target is at angle ϑ_t and Doppler ϖ . By utilizing the composite weight vectors in Equation (111), the SINR in each Doppler bin is given by the same formulas used for fully adaptive STAP,

$$SINR_m(\varpi) = \frac{\sigma^2 \xi_t |\mathbf{w}_m^H \mathbf{v}(\vartheta_t, \varpi)|^2}{\mathbf{w}_m^H \mathbf{R}_u \mathbf{w}_m} . \quad (185)$$

For each potential target Doppler, take the maximum over all Doppler filters:

$$SINR(\varpi) = \max_m SINR_m(\varpi) . \quad (186)$$

It may be more convenient to compute SINR in the lower dimension space. Let $\mathbf{R}_{uy} = \tilde{\mathbf{W}}^H \mathbf{R}_u \tilde{\mathbf{W}}$ be the temporal covariance matrix of the interference-plus-noise signals into the Doppler filter bank, and let

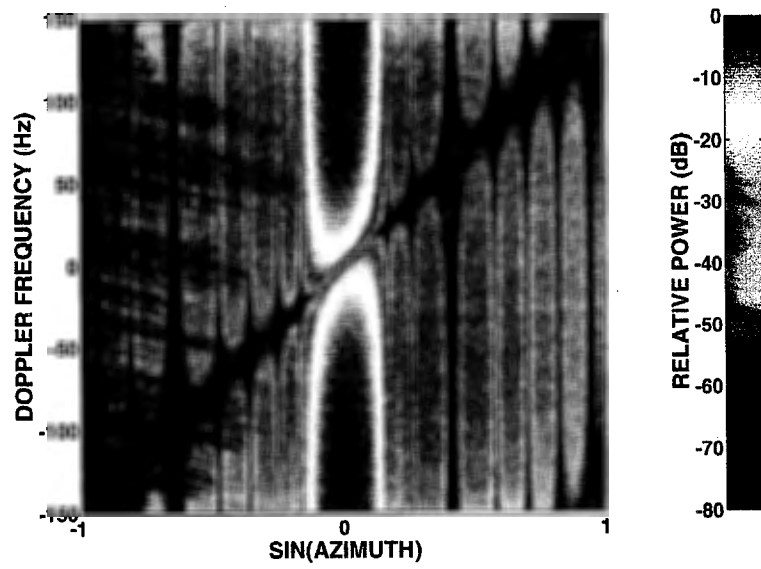
$$\mathbf{v}_y(\vartheta, \varpi) = \tilde{\mathbf{W}}^H \mathbf{v}(\vartheta, \varpi) \quad (187)$$

be the target vector at the same place. The SINR can also be found from

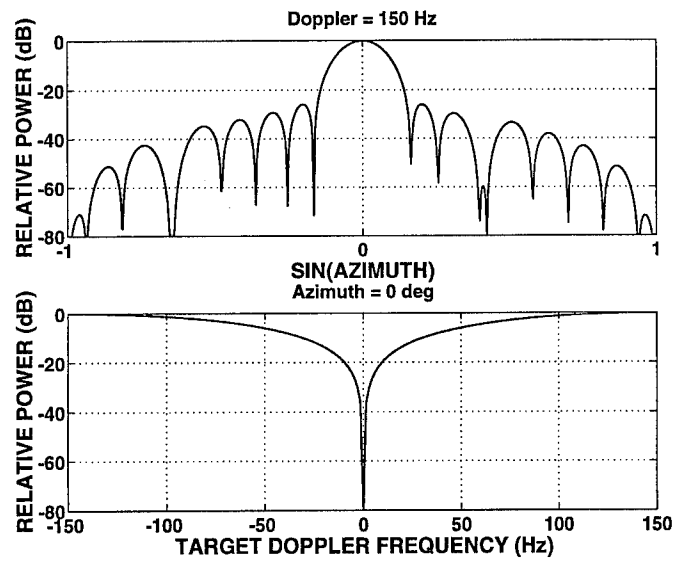
$$SINR_m(\varpi) = \frac{\sigma^2 \xi_t |\mathbf{f}_m^H \mathbf{v}_y(\vartheta_t, \varpi)|^2}{\mathbf{f}_m^H \mathbf{R}_{uy} \mathbf{f}_m} . \quad (188)$$

This latter expression involves only M' -dimensional matrix-vector products, as opposed to the MN -dimensional computations involved in Equation (185).

To illustrate the performance of element-space pre-Doppler STAP, consider again the radar system example of Table 2 and the interference scenario of Table 4. Let $K = 2$ so that with each sub-CPI, a $2N = 36$ -dimensional weight vector is computed. Figure 35(a) and (b) show the adapted pattern from a single sub-CPI and the principal pattern cuts in angle and Doppler. The pattern exhibits deep nulls at the two jammer azimuths and a null along the clutter ridge as expected. The azimuth cut corresponds to a 150-Hz Doppler frequency and shows reduced pattern response at $\sin \phi = \pm 1$, where the clutter has this same Doppler and is suppressed. The Doppler cut shows that at the target azimuth, the only null is that at 0 Hz required to suppress the clutter.



(a)



(b)

Figure 35. $K = 2$ element-space pre-Doppler STAP. (a) Sub-CPI adapted pattern. (b) Principal cuts in angle and Doppler.

Another view of the sub-CPI weight vector is shown in Figure 36. Here $\tilde{\mathbf{w}}$ is broken down into the spatial weight vector components for each PRI, from Equation (177). The individual weight magnitudes are plotted versus element number. For PRI #1, the weight on element #0 is zeroed, and the receive aperture is shifted right of the array physical center. For PRI #2 the receive aperture is shifted left by zeroing the weight for the rightmost element of the array. The shapes of the two responses are identical but shifted from one another by an interelement spacing. Thus, the 2-pulse STAP adaptively implements DPCA processing [2] with spatially adaptive beamformers to cancel the clutter while *simultaneously* suppressing the two jamming signals. With DPCA processing, the outputs of fully formed beams are combined temporally to cancel the clutter; no inherent capability to suppress jamming below the nominal sidelobe level exists.

The sub-CPI response plotted in Figure 35(b) provides coverage over much of the Doppler space. It also colors the thermal noise into the subsequent Doppler filter bank. For this example, assume a 40-dB Chebyshev Doppler filter bank. The composite adapted pattern for Doppler bin 5 (93.75 Hz), from Equation (184), is plotted in Figure 37. The clutter and jammer nulls are still present, and the response has its maximum focused at both the angle and Doppler of a potential target. The corresponding SINR loss is shown in Figure 38. Both the $K = 2$ and $K = 3$ cases are shown, and optimum fully adaptive STAP is included for reference. Because of the angle and Doppler tapers assumed, the pre-Doppler curves exhibit about a 3-dB loss over most of the Doppler space. Because the noise is temporally colored, the low-order Doppler filters contain less noise than those for bins in the middle of the Doppler space. As a result, the SINR is significantly better close to the mainlobe clutter (MDV) than might be expected from examining a sub-CPI response alone. For this case of zero intrinsic clutter motion, $K = 3$ -pulse pre-Doppler exhibits a wider notch at the mainlobe clutter Doppler because of the wider notch implied by the 3-pulse binomial steering vector. When intrinsic clutter motion is present, $K = 2$ pulses are insufficient and the $K = 3$ will provide substantially better performance. When pre-Doppler is implemented with a separate adaptation for each sub-CPI, differences between sub-CPI weight vectors will modulate the target (and any interference residue) prior to Doppler filtering. These differences are likely to be more predominant at Doppler frequencies close to mainlobe clutter. Simulations have shown that diagonal loading or subspace projection may be required to maintain the desired MDV performance. Finally, any clutter residue left after adaptive processing will be coherently integrated by the Doppler filter bank. If the CPI is very long, the potential for this residue to be integrated back above thermal noise merits additional consideration.

The results of this section illustrate one way of achieving near-optimum performance with a partially adaptive space-time processing architecture.

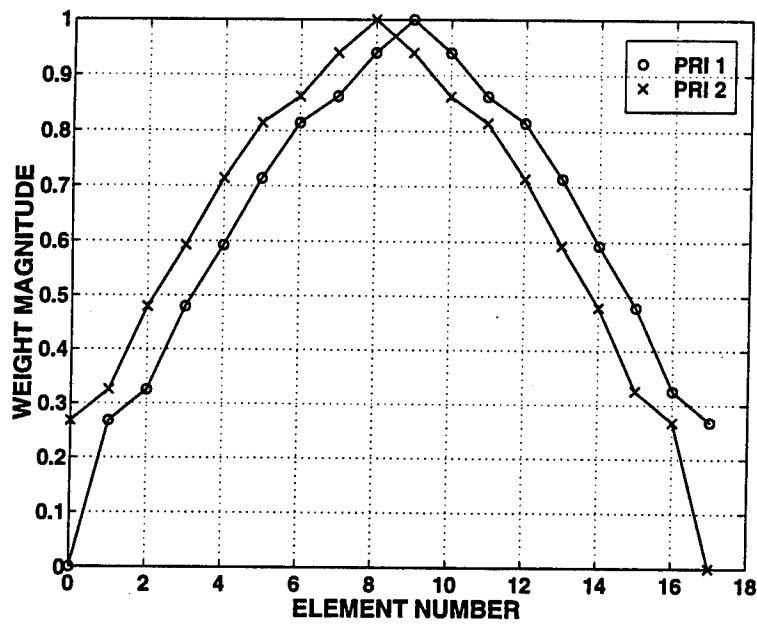
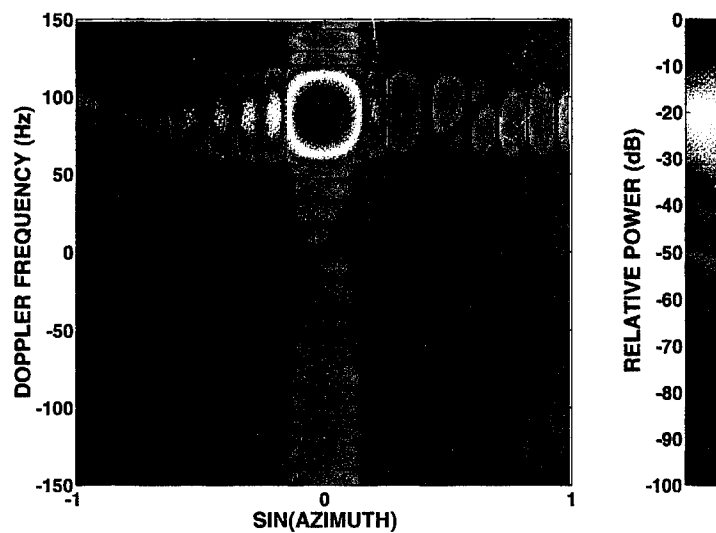
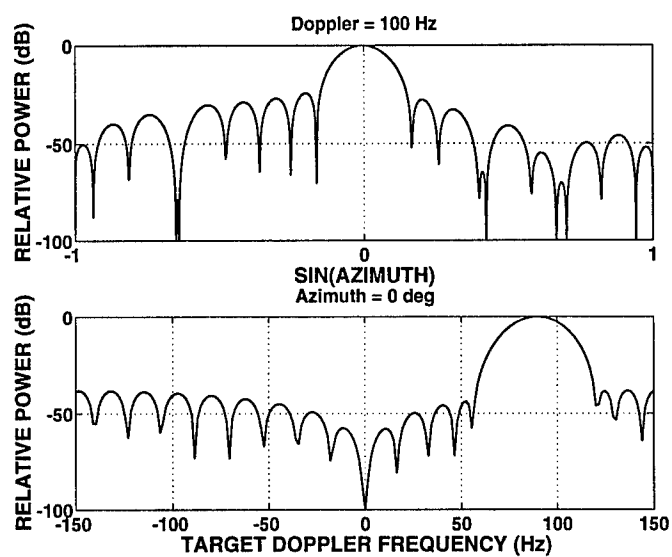


Figure 36. Sub-CPI weight vector decomposed into spatial beamformer components, illustrating the DPCA effect.



(a)



(b)

Figure 37. Composite pattern for Doppler bin 5. (a) Full pattern. (b) Principal cuts in angle and Doppler.

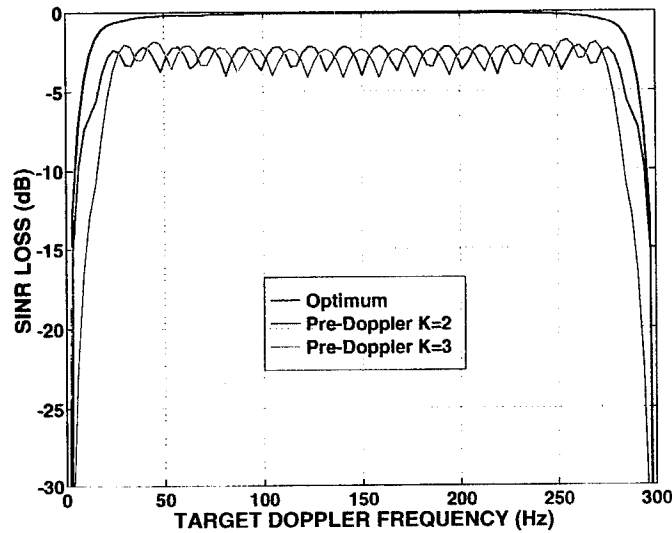


Figure 38. SINR loss for element space pre-Doppler. Zero intrinsic clutter motion.

5.3 Element-Space Post-Doppler

Whereas pre-Doppler STAP reduces the dimensionality by processing only a few pulses at a time, this section considers techniques where all the pulses of the CPI are filtered prior to adaptation. A conventional radar does Doppler filtering after beamforming. Here, Doppler filtering is performed separately on the signals from each array element. A different adaptive problem will be solved in each target Doppler bin, utilizing the signals from all elements. This class of algorithms, represented in the upper right corner of Figure 33, is called *element-space post-Doppler STAP*. The basic idea is that Doppler filtering, with the capability for very low Doppler sidelobes, can localize the clutter in angle and thereby reduce the required number of adaptive degrees of freedom. First, a post-Doppler technique will be considered that utilizes only spatial adaptive processing; then approaches that employ space-time adaptation after Doppler processing will be looked at.

5.3.1 Post-Doppler Adaptive Beamforming

The first post-Doppler algorithm considered utilizes a single Doppler filter bank on each element. Adaptive spatial beamforming is then performed separately within each Doppler bin.¹¹ This approach, termed *post-Doppler adaptive beamforming*, is shown in Figure 39. Because the spatial and temporal processing steps are separated in this architecture, it has also been referred to as “factored post-Doppler” or just “factored STAP.” It is assumed that Doppler filtering suppresses mainlobe clutter nonadaptively and localizes the competing sidelobe clutter in angle. Within each Doppler filter, the adaptive processing places spatial nulls both at the angles of jamming signals and at the angles where the sidelobe clutter Doppler falls within the Doppler passband. Doppler filters containing mainlobe clutter are lost. If performance is acceptable, this is an excellent approach because a significant reduction in dimensionality has been achieved. This approach only requires solving M separate N -dimensional adaptive problems.

Let the Doppler filter bank on each element be represented by the $M \times M$ matrix $\mathbf{F} = [\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{M-1}]$. Filters with very low Doppler sidelobes are required to suppress mainlobe clutter nonadaptively, so assume a tapered DFT filter bank

$$\mathbf{F} = \text{diag}(\mathbf{t}_d) \mathbf{U}^* , \quad (189)$$

where \mathbf{t}_d is an $M \times 1$ low sidelobe taper and \mathbf{U} is the $M \times M$ DFT matrix.

Focus on the adaptive processing for the m th Doppler bin. Let $\tilde{\chi}_m$ be the $N \times 1$ spatial snapshot obtained by collecting the m th Doppler bin outputs from each element,

$$\tilde{\chi}_m = (\mathbf{f}_m \otimes \mathbf{I}_N)^H \chi . \quad (190)$$

Note that for this architecture, the preprocessor of Figure 30 is simply $(\mathbf{f}_m \otimes \mathbf{I}_N)$. The filtered spatial snapshot consists of the interference-plus-noise component

$$\tilde{\chi}_{um} = (\mathbf{f}_m^H \otimes \mathbf{I}_N) \chi_u \quad (191)$$

¹¹Strictly speaking, this approach is not really a space-time adaptive algorithm, as the adaptive weights are spatial only.

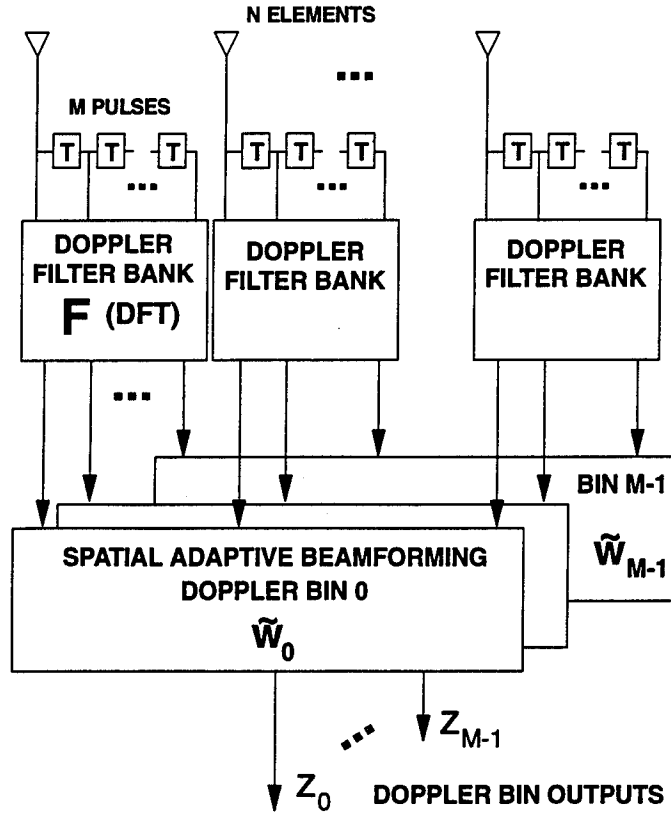


Figure 39. Block diagram of post-Doppler adaptive beamforming (factored post-Doppler).

and, if a target is present, a component

$$\tilde{\chi}_{tm} = (\mathbf{f}_m^H \otimes \mathbf{I}_N)^H (\alpha_t \mathbf{v}_t) \quad (192)$$

$$= (\alpha_t \mathbf{f}_m^H \mathbf{b}_t) \mathbf{a}_t \quad (193)$$

$$= \tilde{\alpha}_{tm} \mathbf{a}_t, \quad (194)$$

where $\tilde{\alpha}_{tm} = \alpha_t \mathbf{f}_m^H \mathbf{b}_t$ is the Doppler-filtered target amplitude and \mathbf{a}_t is the target spatial steering vector.

Compute the adaptive weight vector

$$\tilde{\mathbf{w}}_m = \tilde{\mathbf{R}}_{um}^{-1} \tilde{\mathbf{g}}_t, \quad (195)$$

where the m th bin spatial covariance matrix is given by

$$\tilde{\mathbf{R}}_{um} = E \{ \tilde{\chi}_{um} \tilde{\chi}_{um}^H \} = (\mathbf{f}_m \otimes \mathbf{I}_N)^H \mathbf{R}_u (\mathbf{f}_m \otimes \mathbf{I}_N). \quad (196)$$

The vector $\tilde{\mathbf{g}}_t = \mathbf{t}_a \odot \mathbf{a}_t$ is the desired spatial response, where \mathbf{t}_a is an $N \times 1$ spatial taper. Equation (195) requires the solution of an N -dimensional linear system, and this procedure is repeated for each of the M Doppler bins.

To see the effect of Doppler filtering, derive the component interference covariance matrices by substituting the results of Chapter 2 into Equation (196). The noise component is given by

$$\tilde{\mathbf{R}}_{nm} = \mathbf{f}_m^H \mathbf{f}_m \sigma^2 \mathbf{I}_N. \quad (197)$$

Similarly,

$$\tilde{\mathbf{R}}_{jm} = \mathbf{f}_m^H \mathbf{f}_m \Phi_j \quad (198)$$

is the jammer component, provided that the jamming is stationary across the CPI. Since both noise and jamming are temporally white, their spatial covariance matrices are simply scaled by the Doppler filter white noise gain $\mathbf{f}_m^H \mathbf{f}_m$. The noise is full-rank N and the jamming has rank ρ_j . The clutter spatial covariance matrix (for the m th Doppler bin) is

$$\begin{aligned} \tilde{\mathbf{R}}_{cm} &= (\mathbf{f}_m \otimes \mathbf{I}_N)^H \mathbf{V}_c \mathbf{\Xi}_c \mathbf{V}_c^H (\mathbf{f}_m \otimes \mathbf{I}_N) \\ \tilde{\mathbf{R}}_{cm} &= \sum_{k=1}^{N_c} \sigma^2 \xi_k |\mathbf{f}_m^H \mathbf{b}_k|^2 \mathbf{a}_k \mathbf{a}_k^H. \end{aligned} \quad (199)$$

The contribution from the k th scatterer to the m th bin output may be expressed in terms of the filtered clutter-to-noise-ratio

$$\tilde{\xi}_{mk} = \xi_k |\mathbf{f}_m^H \mathbf{b}_k|^2. \quad (200)$$

Equation (199) can then be expressed compactly as

$$\tilde{\mathbf{R}}_{cm} = \mathbf{A}_c \tilde{\Xi}_{cm} \mathbf{A}_c^H, \quad (201)$$

where $\tilde{\Xi}_{cm} = \sigma^2 \text{diag}(\tilde{\xi}_{m1}, \dots, \tilde{\xi}_{mN_c})$ and \mathbf{A}_c is the $N \times N_c$ matrix of the clutter *spatial* steering vectors. Although clutter exists over all azimuth, the clutter covariance will be of low numerical rank if the Doppler processing is effective in suppressing most of the clutter nonadaptively. The approximate clutter rank will then depend on both the clutter power spectral density and on the Doppler filter frequency response. Thus the effective clutter rank may vary with Doppler bin.

Once the adaptive weights are computed from Equation (195), the output of the m th bin is given by $z_m = \tilde{\mathbf{w}}_m^H \tilde{\mathbf{x}}_m$ or equivalently, $z_m = \mathbf{w}_m^H \mathbf{x}$, where the composite weight vector is given by

$$\mathbf{w}_m = (\mathbf{f}_m \otimes \mathbf{I}_N) \tilde{\mathbf{w}}_m = \mathbf{f}_m \otimes \tilde{\mathbf{w}}_m. \quad (202)$$

This latter form of the composite weight vector, being the Kronecker product of a temporal filter (fixed) and a spatial filter (adaptive), is characteristic of a factored space-time processor. Once the adaptive weights are computed, algorithm performance is computed in the usual way.

The performance of post-Doppler adaptive beamforming is illustrated against the example system and scenario of Tables 2 and 4. Performance will depend on the clutter left after Doppler filtering, so results are shown for different Doppler filter sidelobe levels, assuming Chebyshev windowed filters. Consider first a single Doppler bin; **bin 6 centered at 100 Hz is chosen**. With $\beta = 1$, this Doppler bin corresponds to a sidelobe clutter patch centered at 42° azimuth. Assuming a 40-dB sidelobe Doppler filter, Figure 40 shows the clutter power spectrum and the resultant adapted pattern. Figure 40(a) shows that the largest clutter input to the adaptive processor arrives from a region centered at 42° . However, a significant amount of mainbeam clutter leaks through the Doppler sidelobes at or slightly above thermal noise. Nulling this clutter spatially causes loss of gain on target and poor performance, as evidenced by the pattern of Figure 40(b). Although the dominant sidelobe clutter region and the two jamming signals are nulled, so is the target. Figure 41 shows the results with 80-dB Chebyshev Doppler filters. In this case the Doppler filter sidelobes suppress mainlobe clutter well below thermal noise prior to adaptation. The adaptive beamformer need only react to the competing sidelobe clutter, which is nulled while maintaining nearly full gain on target. The resultant pattern is shown in Figure 41(b). Note also the two point nulls placed on the jamming signals at -40° and 25° . Although performance is much better with a very low-sidelobe Doppler filter, the heavier taper results in additional taper loss and a slightly wider region of sidelobe clutter that must be nulled spatially.

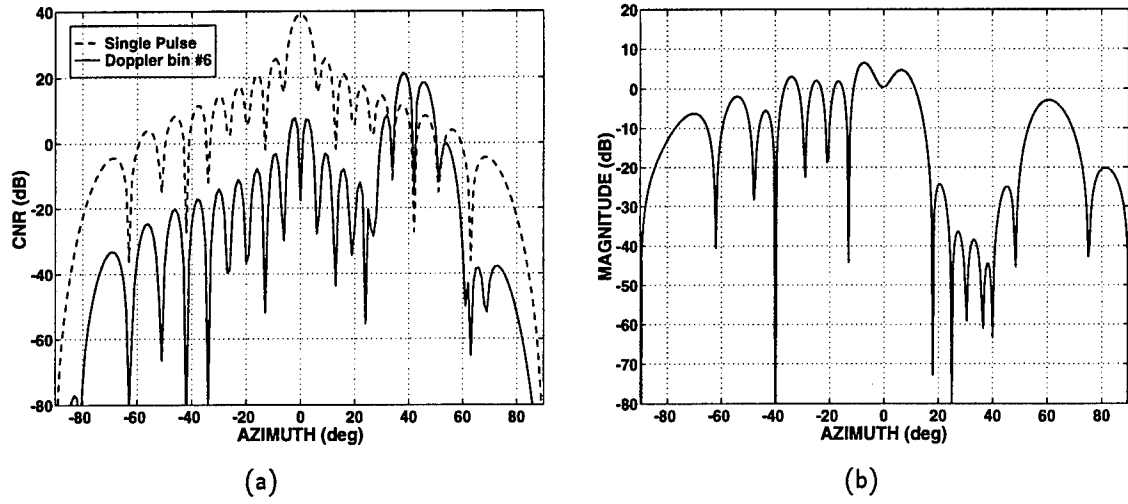


Figure 40. Example performance for Doppler bin 6 (100 Hz) with 40-dB Chebyshev Doppler filters. (a) Clutter power spectral density. (b) Adapted pattern.

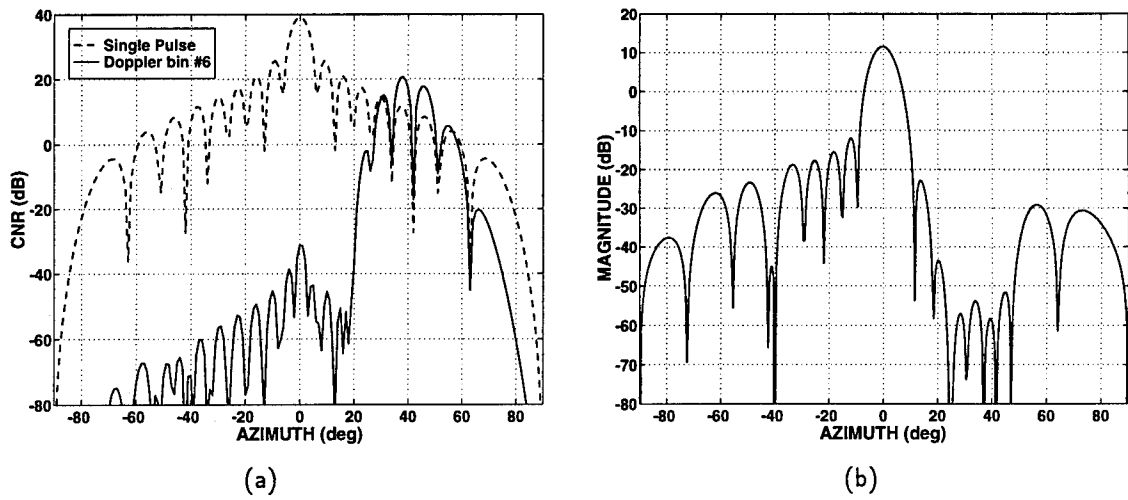


Figure 41. Example performance for Doppler bin 6 (100 Hz) with 80-dB Chebyshev Doppler filters. (a) Clutter power spectral density. (b) Adapted pattern.

Finally, SINR loss performance is plotted in Figure 42. Doppler sidelobes of 80 dB or lower are required to achieve reasonable performance over any of the Doppler space.

With sufficient taper the algorithm achieves about 4 dB of loss over the middle of the Doppler space, which is mostly attributable to the heavy taper. Note also the very wide notch over which performance is low, even with heavy taper. This poor performance at Dopplers approaching that of the mainbeam clutter is due to mainlobe clutter leakage through the skirts of the wide Doppler filter mainlobe. At least one-third of the Doppler space is lost with this approach and the corresponding minimum detectable velocity is poor.

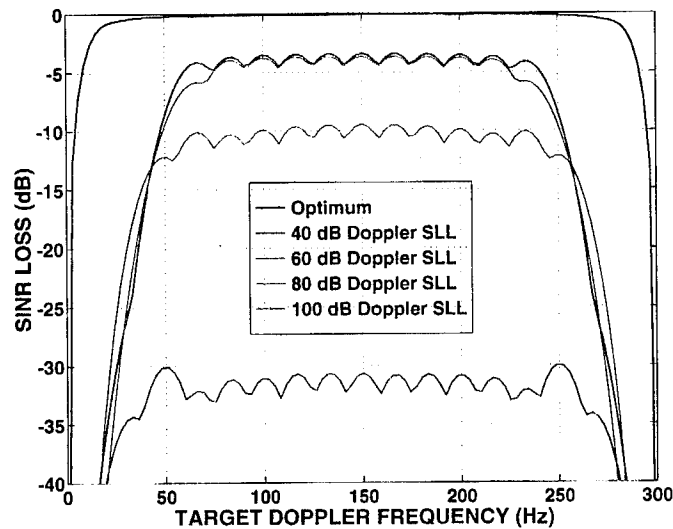


Figure 42. SINR loss as a function of Doppler filter sidelobe level.

The reasons for the relatively poor performance of post-Doppler adaptive beamforming are a consequence of parameters of the example radar system, particularly the CPI length (Doppler resolution) and the aperture size (number of elements). Post-Doppler adaptive beamforming is predicated on the ability of Doppler filtering to both reject mainbeam clutter and to confine the competing sidelobe clutter within a narrow angular region. The relatively poor Doppler resolution available with a short CPI is made worse by the need for low Doppler sidelobes. As a result, the portion of the Doppler space truly free of mainlobe clutter is limited, and in those regions a wide angular region of sidelobe clutter must be suppressed adaptively. This requires many spatial degrees of freedom. These realities have been illustrated with an unambiguous clutter scenario; the results are worse when the clutter is ambiguous in Doppler. For these reasons, the post-Doppler adaptive beamforming approach is better suited to radar systems with longer CPIs and many spatial degrees of freedom.

5.3.2 Multiwindow Post-Doppler STAP

The poor performance of factored post-Doppler can be improved by providing the architecture with temporal adaptivity. This is done by considering the adaptive combination of *multiple* Doppler filters from each element. Each filter may be thought of as a different windowing of the M pulses of data, so such an architecture is called *multiwindow post-Doppler STAP*. A block diagram of the processing for a single target Doppler bin is given in Figure 43. **For the given Doppler bin, each element has an identical bank of K filters.** The default $K = 1$ case is simply factored post-Doppler. Typically, K need only be two or three, and the adaptive problem dimension is reduced by a factor of M/K . The filtered KN signals are adaptively weighted to produce the output for this Doppler bin, and the process is repeated for each Doppler bin.

The analysis of multiwindow post-Doppler follows in much the same way as for post-Doppler adaptive beamforming. **Let m denote the target Doppler bin of interest, and let $\tilde{\mathbf{F}}_m$ be an $M \times K$ matrix whose columns form a set of M -pulse FIR filters that are applied to the signals on each element.** The matrix $\tilde{\mathbf{F}}_m$ is assumed to be of full column rank K , but the columns of $\tilde{\mathbf{F}}_m$ are not necessarily orthogonal. The filtered signals on each element are collected to form a $KN \times 1$ space-time snapshot for the m th bin

$$\tilde{\boldsymbol{\chi}}_m = (\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N)^H \boldsymbol{\chi}. \quad (203)$$

This snapshot contains an interference-plus-noise component $\tilde{\boldsymbol{\chi}}_{um}$ and, under the target-present hypothesis, a target component given by

$$\tilde{\boldsymbol{\chi}}_{tm} = \alpha_t (\tilde{\mathbf{F}}_m^H \otimes \mathbf{I}_N) \mathbf{v}_t = \alpha_t \tilde{\mathbf{F}}_m^H \mathbf{b}_t \otimes \mathbf{a}_t. \quad (204)$$

The m th bin adaptive weight vector is computed from

$$\tilde{\mathbf{w}}_m = \tilde{\mathbf{R}}_{um}^{-1} \tilde{\mathbf{g}}_t, \quad (205)$$

where

$$\tilde{\mathbf{R}}_{um} = E\{\tilde{\chi}_{um} \tilde{\chi}_{um}^H\} \quad (206)$$

is the $KN \times KN$ interference-plus-noise covariance matrix. Given an $MN \times 1$ desired response \mathbf{g}_t optimized for the m th bin Doppler frequency ϖ_m , $\tilde{\mathbf{g}}_t$ is chosen according to Equation (141):

$$\tilde{\mathbf{g}}_t = (\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N)^H \mathbf{g}_t. \quad (207)$$

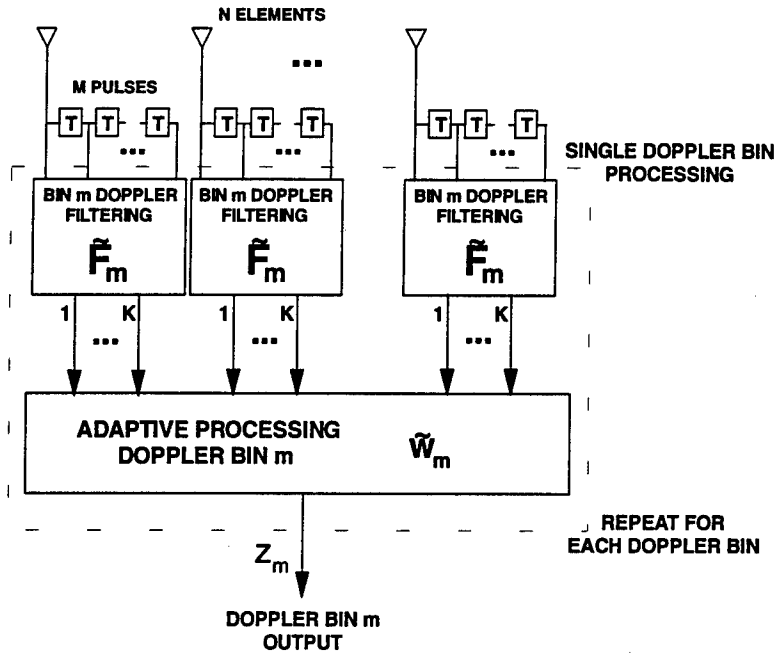


Figure 43. Block diagram for multiwindow post-Doppler STAP

Finally, the output for the m th bin is

$$z_m = \tilde{\mathbf{w}}_m^H \tilde{\chi} = \mathbf{w}_m^H \chi , \quad (208)$$

where

$$\mathbf{w}_m = \left(\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N \right) \tilde{\mathbf{w}}_m \quad (209)$$

is the m th bin composite weight vector.

The major issue is the design of the Doppler filter matrix $\tilde{\mathbf{F}}_m$. There are two requirements. First $\tilde{\mathbf{F}}_m$ must pass targets with Doppler frequencies in the band centered at ω_m . Second, $\tilde{\mathbf{F}}_m$ should be chosen in a way that minimizes the number of adaptive degrees of freedom required for interference cancellation. To this end, consider the covariance matrix after multiwindow Doppler processing. The thermal noise component has covariance

$$\tilde{\mathbf{R}}_{nm} = \sigma^2 \tilde{\mathbf{F}}_m^H \tilde{\mathbf{F}}_m \otimes \mathbf{I}_N . \quad (210)$$

Because $\tilde{\mathbf{F}}_m$ is full rank, the rank of the noise covariance after filtering is also full;

$$\tilde{r}_n = \text{rank}(\tilde{\mathbf{R}}_{nm}) = KN . \quad (211)$$

The noise after filtering, however, may be temporally correlated if columns of $\tilde{\mathbf{F}}_m$ are not orthogonal. Similarly, the jammer component has covariance matrix

$$\tilde{\mathbf{R}}_{jm} = \tilde{\mathbf{F}}_m^H \tilde{\mathbf{F}}_m \otimes \Phi_j \quad (212)$$

and the rank of the jammer interference is

$$\tilde{r}_j = K\rho_j . \quad (213)$$

The jamming component is also temporally correlated according to the correlation between different filters. Neither \tilde{r}_n nor \tilde{r}_j depends on the detailed structure of $\tilde{\mathbf{F}}_m$ aside from the assumption of full rank.

The clutter component of $\tilde{\mathbf{R}}_{um}$ is given by

$$\begin{aligned}\tilde{\mathbf{R}}_{cm} &= (\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N)^H \mathbf{V}_c \mathbf{\Xi}_c \mathbf{V}_c^H (\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N) \\ &= \sum_{k=1}^{N_c} \sigma^2 \xi_k \left(\tilde{\mathbf{F}}_m^H \mathbf{b}_k \mathbf{b}_k^H \tilde{\mathbf{F}}_m \right) \otimes \mathbf{a}_k \mathbf{a}_k^H .\end{aligned}\quad (214)$$

Unlike the noise and jamming covariance matrices, the clutter covariance has a rank that is a function of the Doppler filter matrix $\tilde{\mathbf{F}}_m$. Therefore, $\tilde{\mathbf{F}}_m$ will be designed such that the clutter rank after filtering is minimum.

Chapter 2 showed that, under the assumptions of no velocity misalignment and zero intrinsic clutter motion, the clutter rank is given by Brennan's rule, Equation (76). A similar rule for the clutter rank after Doppler filtering provides a condition for the design of the Doppler filter matrix $\tilde{\mathbf{F}}_m$:

Theorem 4 *If the assumptions of Brennan's rule are satisfied, and if there exists a $K \times K$ non-singular matrix \mathbf{Q} and a length $M' = M - K + 1$ vector $\mathbf{f} = [f_0; f_1; \dots; f_{M'-1}]$ such that*

$$\tilde{\mathbf{F}}_m \mathbf{Q} = \begin{bmatrix} f_0 & & & \mathbf{0} \\ f_1 & \ddots & & \\ \vdots & & & f_0 \\ f_{M'-1} & & & f_1 \\ & \ddots & & \vdots \\ \mathbf{0} & & & f_{M'-1} \end{bmatrix} = \text{Toeplitz} \left(\begin{bmatrix} \mathbf{f}; \mathbf{0}_{(K-1) \times 1} \end{bmatrix}, \begin{bmatrix} f_0, \mathbf{0}_{1 \times (K-1)} \end{bmatrix} \right), \quad (215)$$

then

$$\tilde{r}_c = \text{rank}(\tilde{\mathbf{R}}_{cm}) = N + (K - 1)\beta. \quad (216)$$

This theorem will be referred to as the post-Doppler version of Brennan's rule. If Theorem 4 is satisfied, the Doppler filtering is equivalent to performing a convolution with an M' -pulse filter \mathbf{f} . On each element, M pulses are filtered with \mathbf{f} and the K meaningful output samples are retained. Each output corresponds to filtering a different M' -pulse sub-CPI. Figure 44 illustrates the idea for $N = 4$, $M = 3$, $K = 2$, and $\beta = 1$. The two outputs formed from each element are labeled "Filter 1" and "Filter 2." In accordance with Equation (215) for $\mathbf{Q} = \mathbf{I}_K$, each filter processes a different $M' = 2$ pulse sub-CPI. As long as Filter 1 and Filter 2 have the same response, clutter observations are repeated by different filters on different elements. The clutter rank is equal to the number of independent clutter observations, which can be seen from the figure to be given by Equation (216).

The special case

$$\mathbf{f} = [1; \mathbf{0}_{(M'-1) \times 1}] \quad (217)$$

has already been encountered; this is just pre-Doppler STAP, where no filtering is performed prior to adaptation.

Theorem 4 is proven by first considering only $\mathbf{Q} = \mathbf{I}_K$. In this case, if Equation (215) is satisfied, the filtered steering vector from Equation (204) has form

$$(\tilde{\mathbf{F}}_m \otimes \mathbf{I}_N)^H \mathbf{v}(\vartheta, \varpi) = \tilde{\mathbf{F}}_m^H \mathbf{b}(\varpi; M) \otimes \mathbf{a}(\vartheta) \quad (218)$$

$$= (\mathbf{f}^H \mathbf{b}(\varpi; M')) \mathbf{b}(\varpi; K) \otimes \mathbf{a}(\vartheta) \quad (219)$$

$$= (\mathbf{f}^H \mathbf{b}(\varpi; M')) \tilde{\mathbf{v}}(\vartheta, \varpi) , \quad (220)$$

where

$$\tilde{\mathbf{v}}(\vartheta, \varpi) = \mathbf{b}(\varpi; K) \otimes \mathbf{a}(\vartheta) \quad (221)$$

is a K -pulse space-time steering vector. The quantity $\mathbf{f}^H \mathbf{b}(\varpi; M')$ is the response of filter \mathbf{f} to a signal at Doppler ϖ . The clutter covariance, Equation (214), can now be written as

$$\tilde{\mathbf{R}}_{cm} = \sum_{k=1}^{N_c} \sigma^2 \xi_k |\mathbf{f}^H \mathbf{b}(\varpi; M')|^2 \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^H , \quad (222)$$

or, in more compact form,

$$\tilde{\mathbf{R}}_{cm} = \tilde{\mathbf{V}}_c \tilde{\Xi}_{cm} \tilde{\mathbf{V}}_c^H , \quad (223)$$

where $\tilde{\mathbf{V}}_c$ is a $KN \times N_c$ matrix of clutter space-time steering vectors,

$$\tilde{\Xi}_{cm} = \sigma^2 \text{diag}(\tilde{\xi}_{m1}, \dots, \tilde{\xi}_{mN_c}) , \quad (224)$$

and

$$\tilde{\xi}_{mk} = \xi_k |\mathbf{f}^H \mathbf{b}(\varpi_k; M')|^2 \quad (225)$$

is the filtered CNR for the k th clutter patch. Assume that after filtering, there is nonzero power along the full clutter ridge. Application of Brennan's rule to Equation (223) yields

$$\tilde{r}_c = N + (K - 1)\beta, \quad (226)$$

thus proving the $\mathbf{Q} = \mathbf{I}_K$ case. The generalization to an arbitrary nonsingular \mathbf{Q} follows directly, as multiplication by a nonsingular matrix does not change the rank, and the proof is complete.

Consider two implementations of the Doppler filter matrix. The first architecture corresponds to $\mathbf{Q} = \mathbf{I}_K$ and is shown in Figure 45.

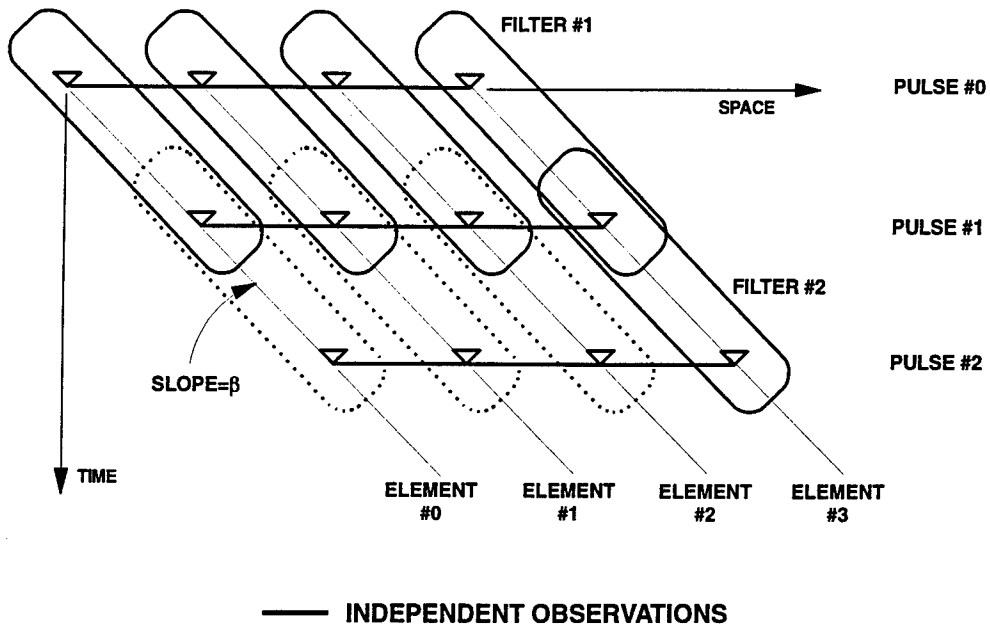


Figure 44. Multiwindow post-Doppler form of Brennan's rule. Example for $N = 4$, $M = 3$, $K = 2$, and $\beta = 1$. The grouping of pulses for the two filters satisfies Theorem 3.

Each element has a bank of M' -pulse Doppler filters that produces K output pulses for each filter (Doppler bin). For each Doppler bin, an adaptive processor combines the K pulses from each element to produce the output signal for that bin. This architecture was first studied by Brennan [18], who used the name filter then adapt. One may also view the Doppler filter bank as a set of K length M' FFTs, with each FFT processing a different M' -pulse sub-CPI. For example, with $K = 2$, one FFT would process pulses $0 : M - 2$, and the second would process pulses $1 : M - 1$. This viewpoint of staggered sub-CPIs leads to the alternative name *PRI-staggered post-Doppler*¹² that will be used. PRI-staggered post-Doppler satisfies conditions of Equation (215) with either uniformly weighted or tapered Doppler filters. Let $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{M-1}]$ be an $M' \times M$ matrix formed from the first M' rows of the $M \times M$ DFT matrix, and let \mathbf{t}_f be an $M' \times 1$ Doppler filter taper. The Doppler filter bank for PRI-staggered is given by

$$\mathbf{F} = [\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{M-1}] = \text{diag}(\mathbf{t}_f)\mathbf{U}, \quad (227)$$

where

$$\mathbf{f}_m = \mathbf{t}_f \odot \mathbf{u}_m \quad (228)$$

is the m th bin Doppler filter impulse response. The Doppler filter matrix $\tilde{\mathbf{F}}_m$ for PRI-staggered post-Doppler is then

$$\tilde{\mathbf{F}}_m = \text{Toeplitz} \left([\mathbf{f}_m; \mathbf{0}_{(K-1) \times 1}], [f_{m0}, \mathbf{0}_{1 \times (K-1)}] \right). \quad (229)$$

The second architecture considered is shown in Figure 46. Here a single length M Doppler filter bank is present on each element. The m th bin output is formed by adaptively combining the spatial samples from a cluster of K adjacent Doppler bins centered at the m th bin center frequency. This approach is called *adjacent-bin post-Doppler*. DiPietro [12] has examined this technique; he refers to it as “extended factored STAP.” To specify the form of $\tilde{\mathbf{F}}_m$ for adjacent-bin post-Doppler, let \mathbf{U} be the $M \times M$ DFT matrix and let \mathbf{t}_f be an $M \times 1$ taper. The Doppler filter bank is represented by the matrix

$$\mathbf{F} = \text{diag}(\mathbf{t}_f)\mathbf{U}^*. \quad (230)$$

¹²The author believes that this name was originally coined by Dr. William Ballance of Hughes Aircraft Company.

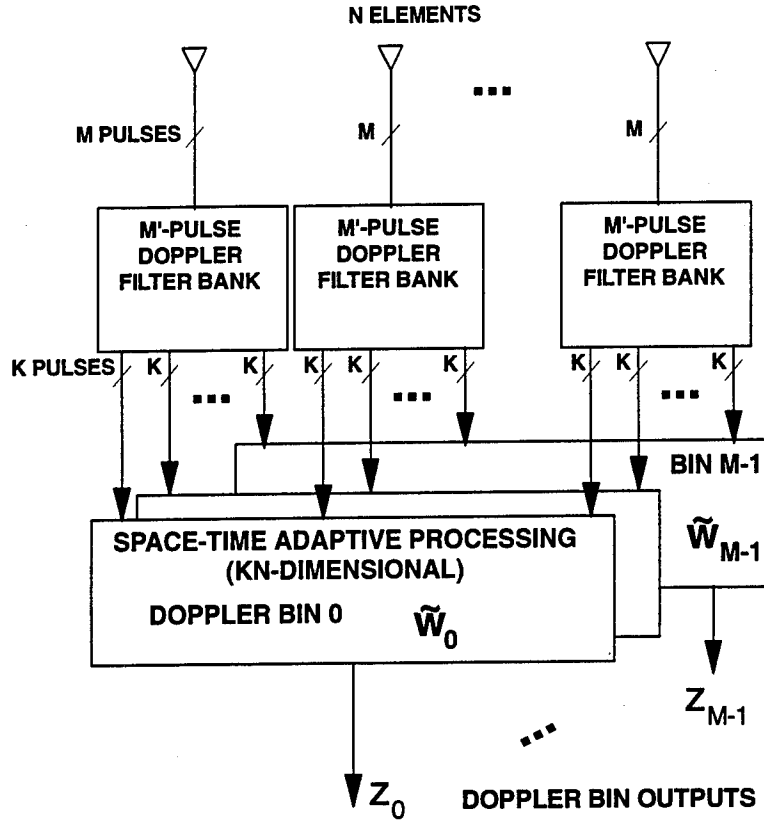


Figure 45. Block diagram for PRI-staggered post-Doppler STAP.

For example, when K is odd, i.e., $K = 2P + 1$, the m th bin output adaptively combines signals from Doppler bins $m - P, \dots, m + P$,

$$\tilde{\mathbf{F}}_m = [\mathbf{f}_{m-P}, \dots, \mathbf{f}_m, \dots, \mathbf{f}_{m+P}] . \quad (231)$$

The adjacent filters are defined to wrap around the Doppler space edges. When K is even, the output bin frequencies are typically defined to lie halfway between the filter bank center frequencies so that the filter clusters are symmetric about the output bin frequencies.

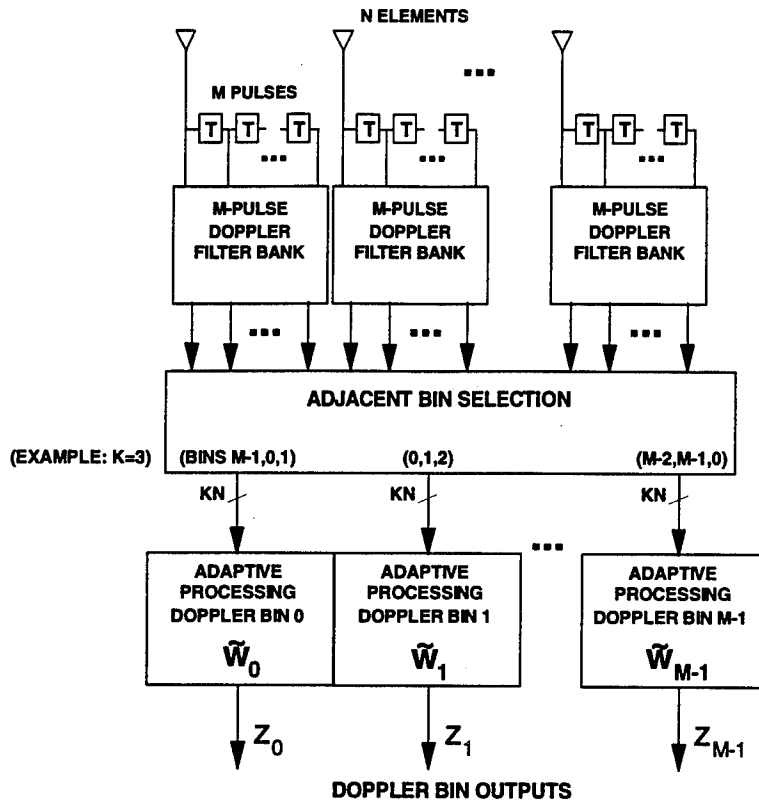


Figure 46. Block diagram for adjacent-bin post-Doppler STAP.

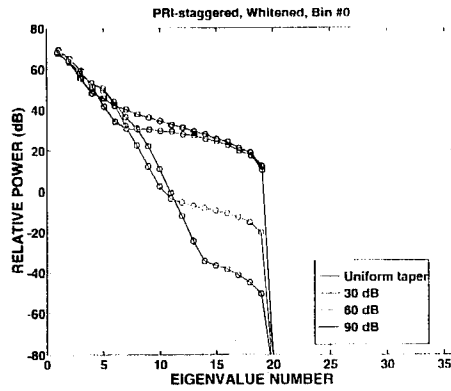
If no Doppler tapering is employed, adjacent-bin post-Doppler can be shown to satisfy Equation (215). However, if the Doppler filters are tapered, or if the frequency spacing between filters is reduced through zero-padding, the minimum rank condition, Equation (215), of Theorem 3 is generally not achieved. In these cases, adjacent-bin post-Doppler may require more adaptive degrees of freedom than PRI-staggered post-Doppler. Examples will be shown for adjacent bin post-Doppler with and without a low-sidelobe taper. Aside from the required number of degrees of freedom, adjacent-bin post-Doppler has some computational advantages over PRI-staggered post-Doppler. Fewer Doppler FFTs and therefore less data buffering are required with adjacent-bin post-Doppler.

Further examination of Theorem 4 reveals an interesting duality between multiwindow post-Doppler STAP and displaced-phase-center-antenna (DPCA) processing. DPCA performs non-adaptive processing spatially, with the aim of forming identical beams that utilize spatially shifted subapertures of the array. These beams are then delayed with respect to each other and subtracted to compensate for the platform motion. Multiwindow post-Doppler, on the other hand, performs nonadaptive processing temporally; the outputs from each element are subsequently beamformed. Theorem 4 says that if shifted temporal subapertures (sub-CPIs) are processed with identical filters, minimum clutter rank is achieved. PRI-staggered post-Doppler may be thought of as standard DPCA, while adjacent-bin post-Doppler is more analogous to a DPCA based on sum and difference beams [2]. Because modern radars are extremely stable temporally, multiwindow post-Doppler STAP performance is not limited to the same degree that element pattern errors limit DPCA performance. Being adaptive by design, multiwindow post-Doppler can suppress sidelobe clutter, as well as suppressing clutter and jamming simultaneously.

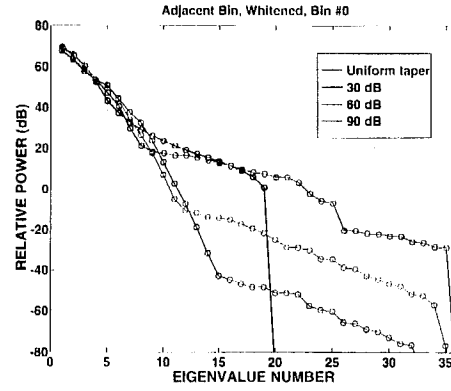
5.3.3 Performance Results

Consider once again the example scenario of Table 4 to illustrate multiwindow post-Doppler STAP performance. The clutter in this case is Doppler-unambiguous with $\beta = 1$. Let $K = 2$ so that a 36-dimensional adaptive problem is solved in each Doppler bin. First, Figure 47 shows a number of clutter covariance eigenspectra. The left column shows results for PRI-staggered post-Doppler, while the right column shows adjacent-bin post-Doppler. Each row corresponds to the covariance matrix for a different Doppler bin. Each plot has several curves that correspond to different Doppler filter tapers. In these plots the eigenspectra are computed after temporally whitening the data to account for the noise correlation induced by the nonorthogonal Doppler filters.

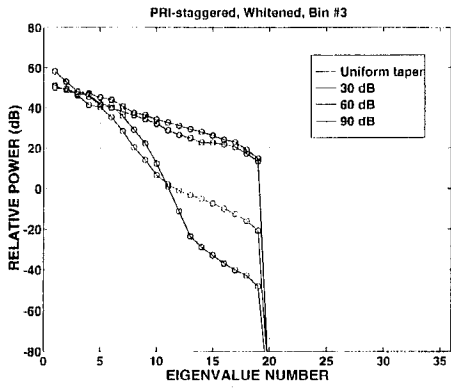
In all cases, the clutter rank for PRI-staggered post-Doppler is $\tilde{r}_c = 19$, as predicted by Theorem 4. The exact shape of the eigenspectra depends on the filter shape and the clutter power spectral density in Doppler (angle). Generally, as the Doppler sidelobes are lowered, the eigenspectrum shows a wider region of larger eigenvalues followed by a plateau of smaller eigenvalues. The level of this plateau generally falls as the Doppler sidelobe level is reduced. Moving from Doppler bin zero toward the center of the Doppler space, the amount of mainbeam clutter that gets through the filter is reduced so the largest eigenvalue becomes smaller. Turning now to the adjacent bin post-Doppler plots, the clutter eigenspectra exhibit a sharp drop only when a uniform taper is used. In this case Theorem 4 is satisfied and the clutter rank is 19. With tapered Doppler filters, adjacent bin post-Doppler no longer satisfies the theorem and the clutter rank is larger (and nearly full). Still, as the Doppler sidelobe level is reduced, the number of eigenvalues that are significant with respect to thermal noise is also reduced. Based on these plots, it can be expected that with only $K = 2$, PRI-staggered will provide better performance for moderate tapers.



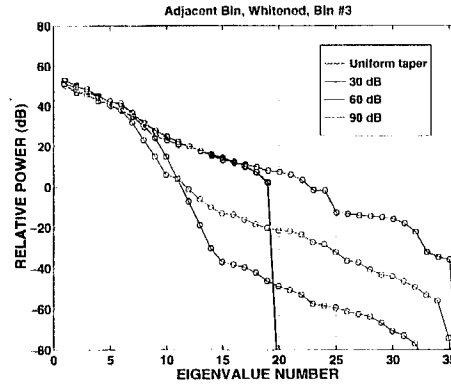
(a)



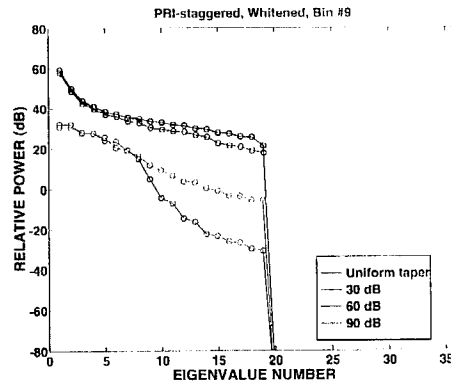
(d)



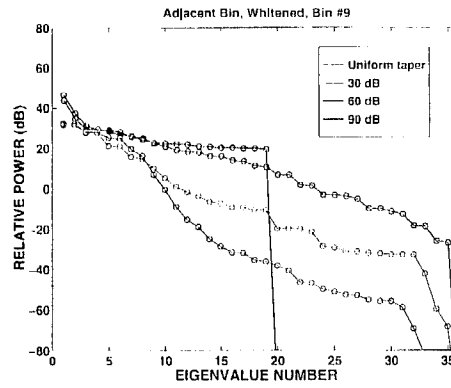
(b)



(e)



(c)



(f)

Figure 47. **Clutter eigenspectra** for multiwindow post-Doppler approaches with $K = 2$. (a) PRI-staggered post-Doppler, covariance matrix for Doppler bin 0. (b) PRI-staggered post-Doppler, Doppler bin 3. (c) PRI-staggered post-Doppler, Doppler bin 9. (d) Adjacent-bin post-Doppler, Doppler bin 0. (e) Adjacent-bin, Doppler bin 3. (f) Adjacent-bin, Doppler bin 9.

SINR loss performance is shown in Figure 48. With uniform taper, performance of PRI-staggered and adjacent-bin post-Doppler are nearly identical. With 30-dB taper, the increased clutter rank for adjacent-bin post-Doppler manifests itself as a severe performance loss over most of the Doppler space. PRI-staggered, as expected, maintains excellent performance over the whole Doppler space. The difference between the two is lessened as the Doppler taper increases, but is still noticeable with 60-dB Doppler sidelobes. At 80- dB Doppler sidelobes, the two approaches again perform similarly. The conclusion is that PRI-staggered post-Doppler generally outperforms adjacent-bin post-Doppler given the same number of degrees of freedom and Doppler sidelobes (and resultant taper loss). Adjacent-bin post-Doppler is more sensitive to the Doppler taper, as its degrees of freedom are not ideally oriented to canceling the clutter ridge. Additional examples, including ambiguous clutter, velocity misalignment, and intrinsic clutter motion, support this conclusion. However, as more degrees of freedom are used (K increases) or as the CPI length increases, the performance differences between PRI-staggered and adjacent-bin become less pronounced.

The above examples illustrate post-Doppler performance in an environment of combined clutter and jamming, where it was assumed that the jamming was stationary over the CPI. Because post-Doppler algorithms combine all pulses prior to adaptation, a jamming scenario that changes from pulse to pulse will require more degrees of freedom than will stationary jamming. Pre-Doppler approaches where the weights are updated every pulse are better suited to a rapidly varying jamming environment. On the other hand, post-Doppler approaches that reject a large portion of the clutter ridge nonadaptively may have more degrees of freedom available to suppress jamming would than a pre-Doppler approach, which must suppress the full clutter ridge.

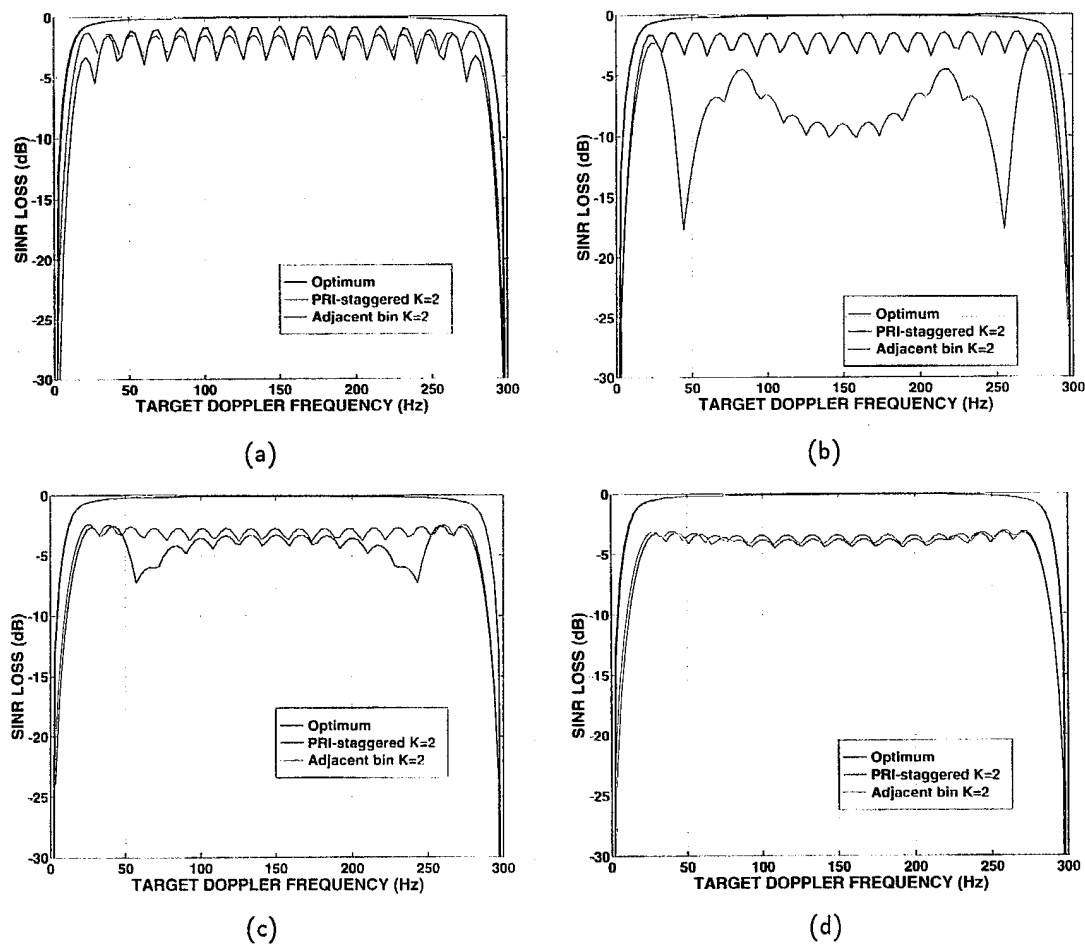


Figure 48. SINR loss performance for PRI-staggered and adjacent-bin post-Doppler STAP. (a) Untapered (uniform) Doppler filters. (b) 30-dB Chebyshev Doppler filters. (c) 60-dB Chebyshev Doppler filters. (d) 90-dB Chebyshev Doppler filters.

6. BEAMSPACE STAP

6.1 Introduction

So far, this report has only considered approaches that adaptively combine signals from all the antenna array elements. Dimensionality reduction was achieved through adaptively processing only a few pulses at a time or by adaptively combining a small number of filtered outputs on each element. In many cases, additional dimensionality reduction can be achieved by beamforming the signals on each element prior to adaptation. Beamforming in this context is a spatial-only operation. Beamforming may reduce the dimensionality by localizing the significant interference to a few signals (beams) and providing additional suppression of the interference outside of the angular region of interest. Architectures that adaptively combine signals after some initial beamforming will be called *beam-space* architectures.

Within the STAP context, beam-space adaptation may reduce the dimensionality required for clutter and jammer suppression. The clutter return is typically largest in the direction of the transmit pattern mainlobe. Thus, a cluster of receive beams centered at the transmit direction will localize the clutter in this area and provide additional sidelobe clutter suppression. The mainlobe clutter portion of the ridge can then be handled adaptively with fewer degrees of freedom than a full element-space technique. Jamming is typically localized to a few discrete angles. Beam-space jammer nulling works best if one has the ability to steer beams in the approximate jammer directions. This in turn requires approximate knowledge of the jammer locations, which may or may not be available a priori. If it is, beam selection may be done a priori; if not, it may be possible to select beams data-adaptively.

If both clutter and jamming are present, the clutter may make the data-adaptive selection of beams for jammer nulling difficult. However, if clutter-free data (ranges or Doppler bins) are available, a beam-space architecture may be implemented using the two-step nulling approach described in Section 4.4. In the first step, the clutter-free data is used to derive an adaptive beam-space transform that cancels jamming as part of the beam-space preprocessing. The resulting output signals are then adaptively combined in a second step to suppress clutter and produce the final output. The jammer nulling first step can be a separate beam-space nulling problem in itself. When two-step nulling is feasible, it may provide significant reduction in adaptive problem size, along with commensurate reductions in computational complexity and training requirements.

In practice, the potential for beam-space processing is dependent on the radar system parameters, the sidelobe levels that can be achieved, and on the a priori knowledge available.

Two sections divide the beam-space discussion: pre- and post-Doppler. Beam-space pre-Doppler algorithms will be considered first. This class of techniques is the dual of the multiwindow post-Doppler algorithms that were discussed in the last section.

Beamforming is simply the application of spatial windows to the element data. The exact dual would be to retain full temporal adaptivity after beamforming; instead, the report will move directly to sub-CPI adaptation seeking greater dimensionality reduction. Beam-space post-Doppler algorithms are then presented. In this architecture, clutter is very effectively localized through both spatial and temporal processing prior to adaptation. With each class, there are both “displaced-filter” and “adjacent-filter” counterparts to the multiwindow post-Doppler approaches discussed in the last chapter.

6.2 Beam-space Pre-Doppler STAP

This section considers beam-space pre-Doppler STAP, represented by the lower left corner of the chart in Figure 33. This approach is illustrated in the block diagrams of Figure 49. Here the problem dimensionality is reduced in two ways. First, the element data is preprocessed with an $N \times K_s$ beamformer matrix $\tilde{\mathbf{G}}$ to produce a small number K_s of beam outputs. Second, only the beam outputs from a small, K_t -pulse, sub-CPI are adaptively processed at one time. The sub-CPI processing is shown in Figure 49(a). The adaptive problem dimensionality is $K = K_s K_t$. Typically $K_t \ll M$ and $K_s \ll N$ so that a significant reduction in problem size is achieved. As with its element-space counterpart, a separate adaptive problem is solved for each sub-CPI, and the sub-CPI outputs are coherently processed with an $M' = M - K_t + 1$ -pulse Doppler filter bank. A full CPI depiction of beam-space pre-Doppler is shown in Figure 49(b). An analysis will be presented in terms of a general beamformer matrix and then the issues of how to select beams in practice discussed.

6.2.1 Basic Analysis

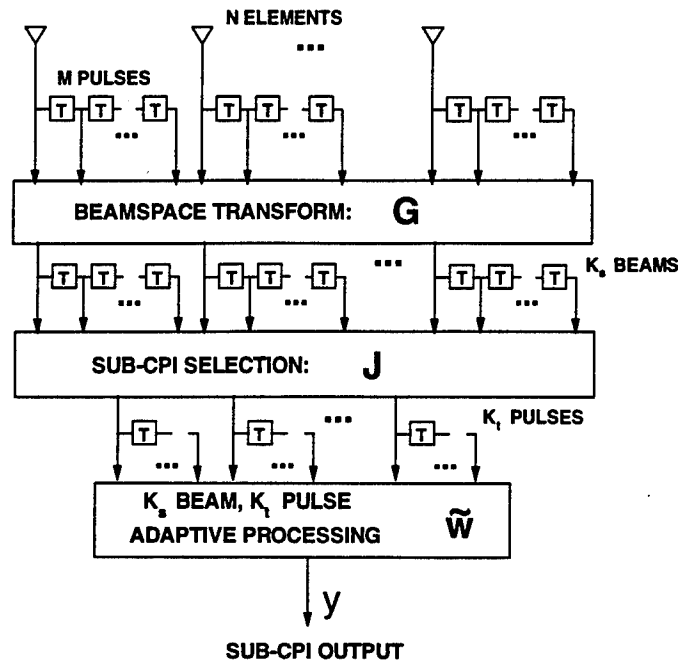
The analysis proceeds in much the same way as for element-space pre-Doppler. To begin, consider only the first sub-CPI containing the returns from pulses $0 : K_t - 1$. Let $\tilde{\mathbf{x}}$ denote the $K_t K_s \times 1$ beam-space sub-CPI snapshot,

$$\tilde{\mathbf{x}} = (\mathbf{J} \otimes \tilde{\mathbf{G}})^H \mathbf{x}. \quad (232)$$

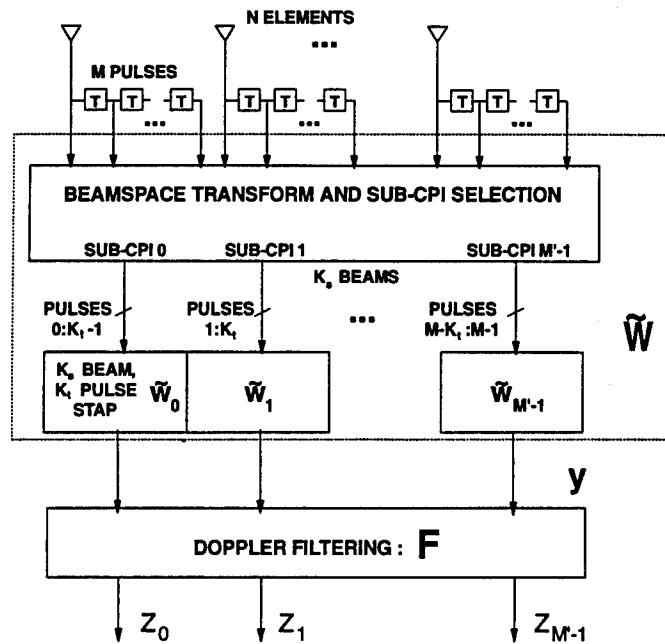
Recall that the matrix \mathbf{J} , defined in Equation (165), is an $M \times K_t$ selection matrix that chooses the first K_t pulses from the CPI. This snapshot consists of an interference-plus-noise component $\tilde{\mathbf{x}}_u$ and, if a target is present, the target component

$$\tilde{\mathbf{x}}_t = (\mathbf{J} \otimes \tilde{\mathbf{G}})^H (\alpha_t \mathbf{v}_t) = \alpha_t \tilde{\mathbf{b}}_t \otimes \tilde{\mathbf{G}}^H \mathbf{a}_t, \quad (233)$$

where $\tilde{\mathbf{b}}_t = \mathbf{b}(\varpi_t; K_t)$ is the target’s sub-CPI temporal steering vector.



(a)



(b)

Figure 49. Beamspace pre-Doppler STAP. (a) Sub-CPI processing. (b) Full CPI processing.

The sub-CPI adaptive weight vector is computed from

$$\tilde{\mathbf{w}} = \tilde{\mathbf{R}}_u^{-1} \tilde{\mathbf{u}}_t, \quad (234)$$

where $\tilde{\mathbf{R}}_u = E\{\tilde{\chi}_u \tilde{\chi}_u^H\}$ is the $K_t K_s \times K_t K_s$ beamspace sub-CPI interference-plus-noise covariance matrix and $\tilde{\mathbf{u}}_t$ is a $K_t K_s \times 1$ desired response. The desired response for beamspace pre-Doppler is chosen as follows. Let $\tilde{\mathbf{g}}_t$ be the desired sub-CPI response for element space pre-Doppler, from Equation (172). The beamspace desired response is taken to be

$$\tilde{\mathbf{u}}_t = (\mathbf{I}_{K_t} \otimes \tilde{\mathbf{G}})^H \tilde{\mathbf{g}}_t \quad (235)$$

$$= (\tilde{\mathbf{t}}_b \odot \tilde{\mathbf{b}}(\varpi_b)) \otimes \tilde{\mathbf{G}}^H(\mathbf{t}_a \odot \mathbf{a}_t). \quad (236)$$

The binomial-tapered temporal component of $\tilde{\mathbf{u}}_t$ is identical to that of element space pre-Doppler. The spatial component is modified by the beamspace transform in a way, Equation (141), that preserves the quiescent spatial response.

To consider the full CPI, let the p th sub-CPI snapshot be

$$\tilde{\chi}_p = (\mathbf{J}_p \otimes \tilde{\mathbf{G}})^H \chi, \quad (237)$$

where \mathbf{J}_p is the selection matrix, Equation (175). The p th sub-CPI weight vector

$$\tilde{\mathbf{w}}_p = \tilde{\mathbf{R}}_{up}^{-1} \tilde{\mathbf{u}}_t \quad (238)$$

is computed and applied to produce the sub-CPI output $y_p = \tilde{\mathbf{w}}_p^H \tilde{\chi}_p$. It is useful to think of the sub-CPI weight vector in terms of its constituent spatial components. Let $\tilde{\mathbf{w}}_{p,k}$ be the $K_s \times 1$ vector of beam weights for the k th pulse of the p th sub-CPI. The sub-CPI weight vector is rearranged to form the $K_s \times K_t$ weight matrix

$$\tilde{\mathbf{W}}_p = [\tilde{\mathbf{w}}_{p,0}, \tilde{\mathbf{w}}_{p,1}, \dots, \tilde{\mathbf{w}}_{p,K_t}] \quad (239)$$

so that

$$\tilde{\mathbf{w}}_p = \text{vec}(\tilde{\mathbf{W}}_p). \quad (240)$$

The sub-CPI output can then be written as

$$y_p = [(\mathbf{J}_p \otimes \tilde{\mathbf{G}}) \tilde{\mathbf{w}}_p]^H \boldsymbol{\chi} = [(\mathbf{I}_M \otimes \tilde{\mathbf{G}}) \text{vec}(\tilde{\mathbf{W}}_p \mathbf{J}_p^T)]^H \boldsymbol{\chi} . \quad (241)$$

The sub-CPI outputs are assembled into a vector \mathbf{y} ,

$$\mathbf{y} = [(\mathbf{I}_M \otimes \tilde{\mathbf{G}}) \tilde{\mathbf{W}}]^H \boldsymbol{\chi} , \quad (242)$$

where

$$\tilde{\mathbf{W}} = \left[\text{vec}(\tilde{\mathbf{W}}_0 \mathbf{J}_0^T), \text{vec}(\tilde{\mathbf{W}}_1 \mathbf{J}_1^T), \dots, \text{vec}(\tilde{\mathbf{W}}_{M'-1} \mathbf{J}_{M'-1}^T) \right] \quad (243)$$

$$= \begin{bmatrix} \tilde{\mathbf{w}}_{0,0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \tilde{\mathbf{w}}_{0,1} & \tilde{\mathbf{w}}_{1,0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \tilde{\mathbf{w}}_{0,K_t-1} & & & \ddots & \mathbf{0} \\ \mathbf{0} & \ddots & & & \tilde{\mathbf{w}}_{M-K_t,0} \\ \vdots & \ddots & \ddots & & \tilde{\mathbf{w}}_{M-K_t,1} \\ \mathbf{0} & & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{w}}_{M-K_t,K_t-1} \end{bmatrix} \quad (244)$$

is of size $MK_s \times M'$. Finally, the sub-CPI outputs are passed through a length M' Doppler filter bank denoted by $\mathbf{F} = [\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{M'}]$. The final output of the m th Doppler bin is given by

$$z_m = \mathbf{f}_m^H \mathbf{y} = \mathbf{w}_m^H \boldsymbol{\chi} , \quad (245)$$

where the beamspace pre-Doppler composite weight vector is given by

$$\mathbf{w}_m = (\mathbf{I}_M \otimes \tilde{\mathbf{G}}) \tilde{\mathbf{W}} \mathbf{f}_m . \quad (246)$$

The SINR performance can then be computed from Equations (185) and (186) in the usual ways.

6.2.2 Beamformer Matrix Design

This section considers the design of the beamformer matrix $\tilde{\mathbf{G}}$. As usual, the goal is maximum performance with minimal effort; i.e., use of as few beams as possible. Assume that $\tilde{\mathbf{G}}$ is of full column rank K_s , but the columns need not be orthogonal. To understand the effect of beamforming, examine the interference and noise components of the beamspace sub-CPI snapshot. Without loss of generality, the discussion will focus on the first sub-CPI from Equation (232). The results of Chapter 2 are used freely.

The thermal noise component of the data has covariance matrix

$$\tilde{\mathbf{R}}_n = (\mathbf{J} \otimes \tilde{\mathbf{G}})^H \mathbf{R}_n (\mathbf{J} \otimes \tilde{\mathbf{G}}) \quad (247)$$

$$= \sigma^2 \mathbf{J}^H \mathbf{J} \otimes \tilde{\mathbf{G}}^H \tilde{\mathbf{G}} = \sigma^2 \mathbf{I}_{K_t} \otimes \tilde{\mathbf{G}}^H \tilde{\mathbf{G}}. \quad (248)$$

As $\tilde{\mathbf{G}}$ is full rank, the noise covariance $\tilde{\mathbf{R}}_n$ is also full rank,

$$\tilde{r}_n = K_t K_s. \quad (249)$$

The noise may not be spatially white, however, due to the coloring induced by the beamformer matrix. Similarly, the sub-CPI jammer covariance matrix is given by

$$\tilde{\mathbf{R}}_j = \mathbf{J} \otimes \tilde{\mathbf{G}}^H \mathbf{R}_j (\mathbf{J} \otimes \tilde{\mathbf{G}}) \quad (250)$$

$$= \mathbf{I}_{K_t} \otimes \tilde{\mathbf{G}}^H \Phi_j \tilde{\mathbf{G}} \quad (251)$$

$$= \mathbf{I}_{K_t} \otimes (\tilde{\mathbf{G}}^H \mathbf{A}_j) \Xi_j \mathbf{A}_j^H \tilde{\mathbf{G}}, \quad (252)$$

where Equation (45) has been used. The jammer rank is at most $\tilde{r}_j = K_t \rho_j$, but it may be less if the beamformer nulls some of the jamming signals.

The clutter component is again the most interesting. From Equation (61),

$$\tilde{\mathbf{R}}_c = (\mathbf{J} \otimes \tilde{\mathbf{G}})^H \mathbf{R}_c (\mathbf{J} \otimes \tilde{\mathbf{G}}) \quad (253)$$

$$= \sum_{k=1}^{N_c} \sigma^2 \xi_k (\mathbf{J}^H \mathbf{b}_k \mathbf{b}_k^H \mathbf{J}) \otimes (\tilde{\mathbf{G}}^H \mathbf{a}_k \mathbf{a}_k^H \tilde{\mathbf{G}}) \quad (254)$$

$$= \sum_{k=1}^{N_c} \sigma^2 \xi_k (\tilde{\mathbf{b}}_k \tilde{\mathbf{b}}_k^H) \otimes (\tilde{\mathbf{G}}^H \mathbf{a}_k \mathbf{a}_k^H \tilde{\mathbf{G}}). \quad (255)$$

Notice that the expression above is very similar to Equation (214) from the multiwindow post-Doppler discussion. The clutter covariance rank will depend on the beamformer matrix. The minimum clutter rank is achieved with a beamformer matrix that satisfies the following analogue of Theorem 4, referred to as the beamspace version of Brennan's rule.

Theorem 5 *If the assumptions of Brennan's rule are satisfied, and if there exists a $K_s \times K_s$ nonsingular matrix \mathbf{Q} and a length $N' = N - K_s + 1$ vector $\mathbf{g} = [g_0; g_1; \dots; g_{N'-1}]$ such that*

$$\tilde{\mathbf{G}}\mathbf{Q} = \begin{bmatrix} g_0 & & & \mathbf{0} \\ g_1 & & \ddots & \\ \vdots & & & g_0 \\ g_{N'-1} & & & g_1 \\ & & \ddots & \vdots \\ \mathbf{0} & & & g_{N'-1} \end{bmatrix} = \text{Toeplitz} \left(\begin{bmatrix} \mathbf{g}; \mathbf{0}_{(K_s-1) \times 1} \end{bmatrix}, \begin{bmatrix} g_0, \mathbf{0}_{1 \times (K_s-1)} \end{bmatrix} \right), \quad (256)$$

then the clutter rank is minimum and equal to

$$\tilde{r}_c = \text{rank}(\tilde{\mathbf{R}}_c) = K_s + (K_t - 1)\beta, \quad (257)$$

The proof of Theorem 5 follows the exact same steps as the proof of Theorem 4 and is not included here. The Toeplitz form of Equation (256) represents the identical beamforming (with \mathbf{g}) of overlapping N' -element subapertures, with each subaperture shifted by one interelement spacing from the previous one. This processing is depicted in Figure 50 for the case $N = 4$, $K_s = 2$, $K_t = 3$. Note that beam #2 on pulse 0 and beam #1 on pulse 1 are observations from effectively the same point in space. If the beamformer matrix obeys the condition of Theorem 5, these two observations are identical, and only one contributes to the clutter rank. The number of independent clutter observations is four, which equals that given by Equation (257).

The conditions for which minimum clutter rank is achieved are equivalent to the conditions for which perfect DPCA clutter cancellation is possible. The beamformer matrix of Equation (256) with $\mathbf{Q} = \mathbf{I}_{K_s}$ is precisely that required by DPCA. In fact, DPCA is simply a special case of beamspace pre-Doppler space-time processing, whereby the weights are fixed as opposed to being computed data-adaptively. Appendix 3 contains a discussion of DPCA processing for reference.

Because of this equivalence, the $\mathbf{Q}_s = \mathbf{I}_K$ form of beamspace pre-Doppler will be referred to as *displaced phase center pre-Doppler* or *displaced-beam pre-Doppler*. Displaced-beam pre-Doppler is the dual of PRI-staggered post-Doppler STAP. Theorem 5 also gives the minimum number of beams required for clutter cancellation. Partially adaptive STAP theory says the number of adaptive degrees of freedom must be greater than the clutter rank,

$$K \geq \tilde{r}_c + 1. \quad (258)$$

By utilizing Equation (257) and $K = K_s K_t$, the equivalent condition

$$K_s \geq \beta + \frac{1}{K_t - 1} \quad (259)$$

results. The required number of beams (or effective phase centers) increases linearly with the effective platform motion (or Doppler ambiguity).

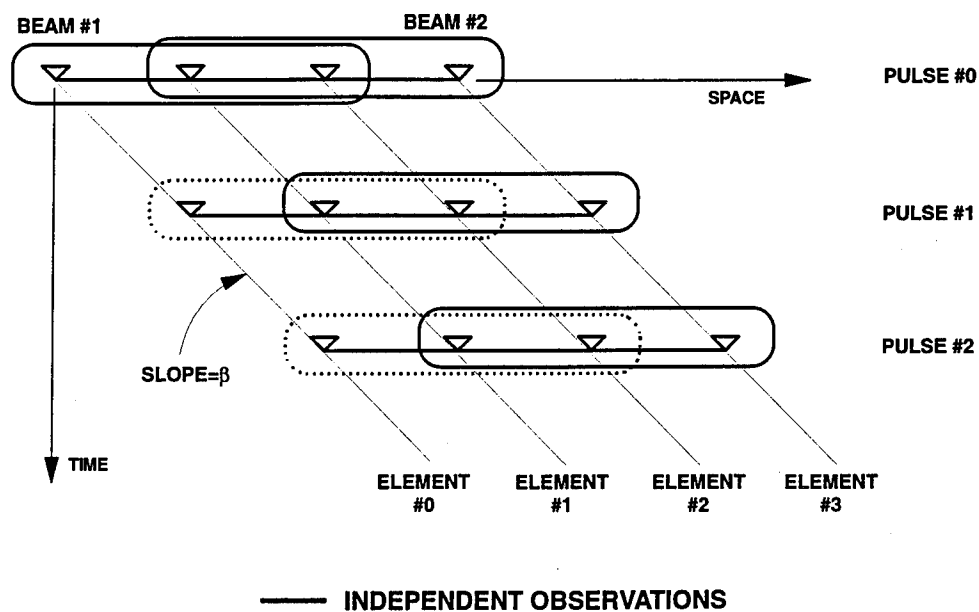


Figure 50. Beamspace version of Brennan's rule for the case $N = 4$, $K_t = 3$, $K_s = 2$, and $\beta = 1$.

In practice, constructing a set of beamformers to satisfy Theorem 5 is much more difficult than constructing a set of Doppler filters to satisfy Theorem 4. The reason is that the spatial errors due to antenna element mismatches are typically relatively high (say 20 – 50 dB). In contrast, the temporal errors due to pulse-to-pulse instability can be kept to very low (80 – 100 dB) levels in modern solid-state radars. Moreover, the need to suppress jamming spatially may result in adaptive beamformers that are altered from the form required for minimum clutter rank. However, Theorems 4 and 5 provide a great deal of insight into the clutter characteristics and to the duality between different STAP architectures.

The beamspace analog of adjacent-bin post-Doppler, called *adjacent-beam pre-Doppler*, utilizes a set of adjacent beams, each of which uses the full aperture. Let \mathbf{G} be an $N \times N$ matrix beamformer whose columns are beamformers steered to different angles, and let \mathbf{J} be an $N \times K_s$ selection matrix that chooses the columns of \mathbf{G} corresponding to a cluster of adjacent beams centered at the transmit direction. Adjacent-beam pre-Doppler utilizes the beamformer matrix

$$\tilde{\mathbf{G}} = \mathbf{G}\mathbf{J} \quad (260)$$

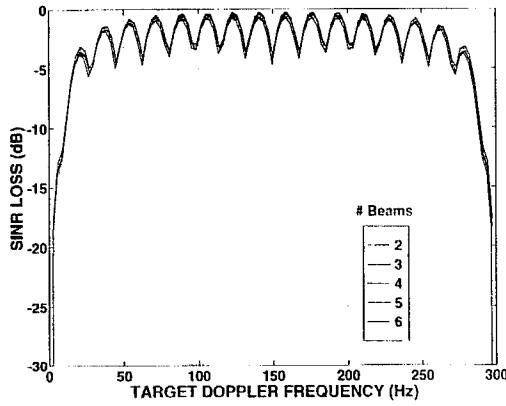
for clutter cancellation. With no tapering, adjacent-beam post-Doppler satisfies Equation (256) with $\mathbf{Q}_s \neq \mathbf{I}_{K_s}$. When tapering on receive is employed, the adjacent-beam approach may not satisfy Theorem 5. As with element-space post-Doppler algorithms, performance with adjacent beams (for the same number of beams) may be worse than that achievable with displaced beams.

6.2.3 Performance Examples

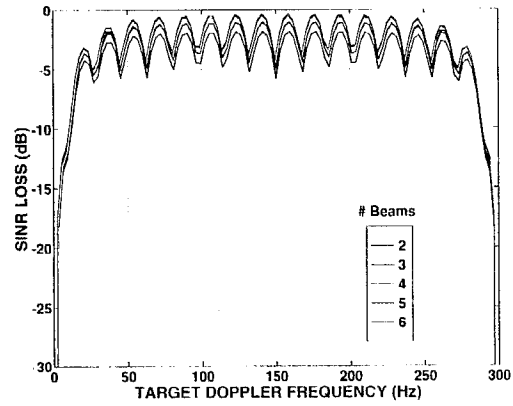
The performance of beamspace pre-Doppler will be illustrated first in a clutter-only environment. The example system of Table 2 and the example clutter scenario of Table 4 are assumed. With $K_t = 2$, the number of beams is varied. The resulting SINR loss performance is shown in Figure 51. With displaced beams, performance is excellent for as few as two beams, as predicted by Equation (259). Because of this, performance changes very little as additional beams are added.

Figure 51(c) shows a similar result with tapered beamformer. As Theorem 5 is still satisfied, performance is excellent again as the number of DOF is always larger than the clutter rank. The adjacent beam case is shown in Figures 51(b) and (d). With no taper, the clutter rank is given by Equation (257) and two beams are sufficient. With a taper, however, the adjacent-beam approach does not work well with only two beams. In this case performance suffers at Dopplers close to the mainlobe clutter. The adjacent beam approach requires at least four beams before its performance is equivalent to that of the displaced-beam case with only two beams.

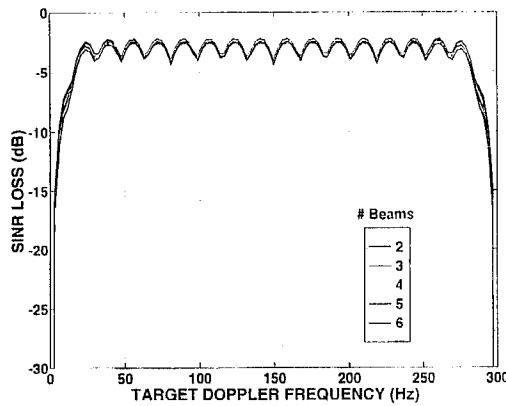
Performance against ambiguous clutter is shown in Figure 52. With $\beta = 2$ there are two azimuths at which clutter has a given Doppler. At least $K_s = 3$ beams are required. Similarly, when $\beta = 3$, at least four beams are needed to attain reasonable performance over most of the Doppler space. These results are in agreement with Equation (259). Note also how the mainlobe clutter notch widens as β increases.



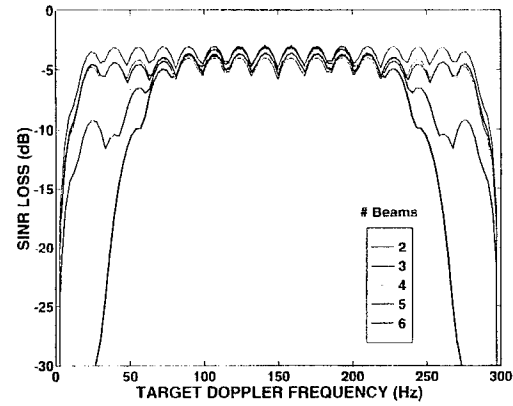
(a)



(b)



(c)



(d)

Figure 51. SINR loss performance for beamspace pre-Doppler STAP in a clutter-only scenario. (a) Displaced beams, untapered. (b) Adjacent beams, untapered. (c) Displaced beams, 30-dB Chebyshev taper. (d) Adjacent beams, 30-dB Chebyshev taper.

6.2.4 Jammer Cancellation: A Two-Step Nulling Approach

The beamspace architectures discussed thus far have not yet addressed jamming interference. For best performance, beamspace architectures typically require beams that are pointed toward the dominant interference sources. If jamming signals lie outside the region covered by the beam cluster for clutter cancellation, jamming will degrade performance. One way to deal with jamming is to form additional beams that are pointed toward any jamming signals. This is possible if a priori information on the approximate jammer locations are known. Without a priori knowledge, one may derive jammer information data-adaptively. Doing this, however, is subject to contamination by any clutter present unless clutter-free data exists or can be derived (through Doppler filtering, perhaps).

Assuming that additional jammer-cancellation beams have been allocated, the beamformer matrix may be partitioned as

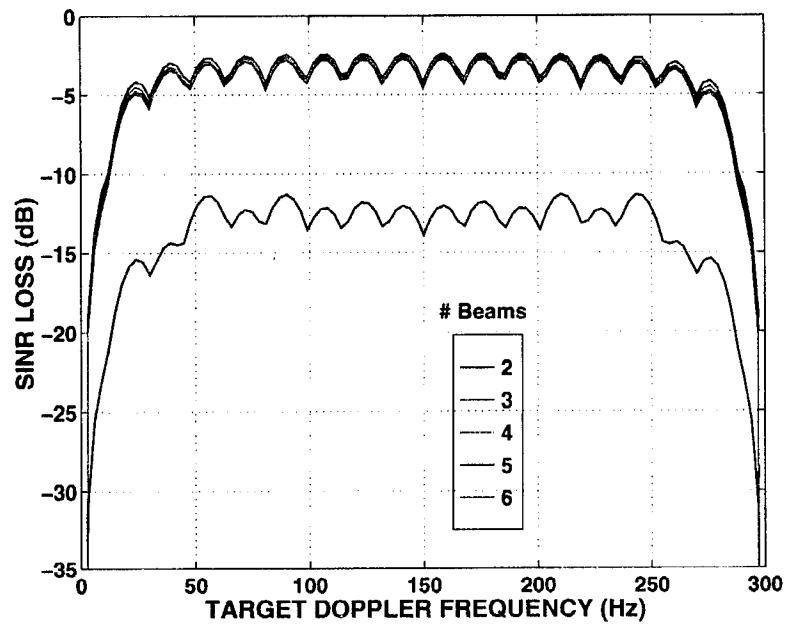
$$\tilde{\mathbf{G}} = [\tilde{\mathbf{G}}_c, \tilde{\mathbf{G}}_j] , \quad (261)$$

where $\tilde{\mathbf{G}}_j$ is an $N \times K_j$ matrix containing beamformers pointed in the approximate jammer directions. Here $\tilde{\mathbf{G}}_c$ is the $N \times K_c$ beamformer matrix used before. Each additional jammer beam adds K_t adaptive degrees of freedom, for a total increase in dimensionality of $K_j K_t$.

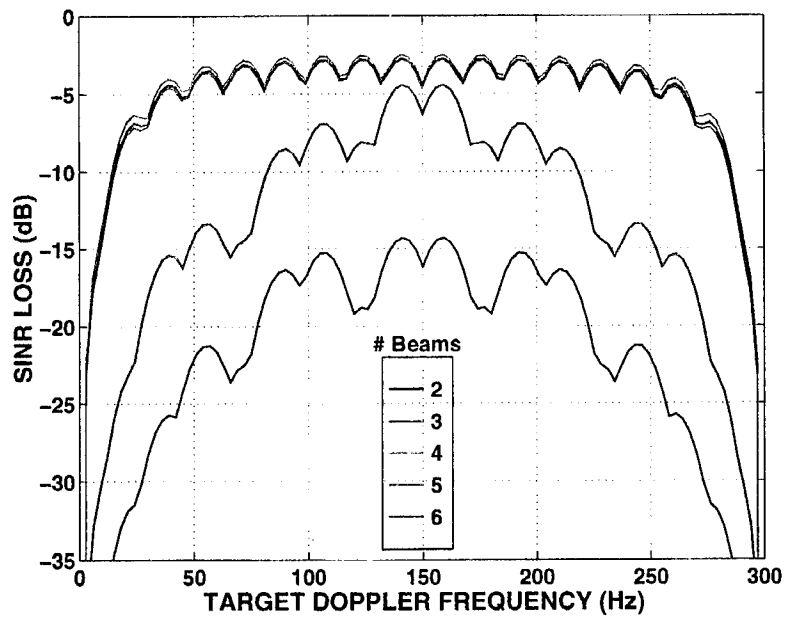
If appropriate clutter-free data can be obtained, another way of dealing with jamming is to implement a two-step nulling architecture as described in Section 4.4. This approach is illustrated in Figure 53. The clutter-free data is used to estimate the spatial jammer-plus-noise covariance matrix

$$\Phi_{jn} = \Phi_j + \sigma^2 \mathbf{I}_N . \quad (262)$$

Within each sub-CPI, Φ_{jn} is used to construct adaptive beamformers that null the jamming as part of the beamspace transformation. The resulting beam outputs, with the jamming removed, are then used for adaptive clutter suppression as before. Thus, interference cancellation occurs in two steps, the first being to suppress jamming and the second to remove the clutter.



(a)



(b)

Figure 52. Displaced-beam pre-Doppler STAP performance with ambiguous clutter and no jamming. A 30-dB Chebyshev taper is assumed. (a) $\beta = 2$. (b) $\beta = 3$.

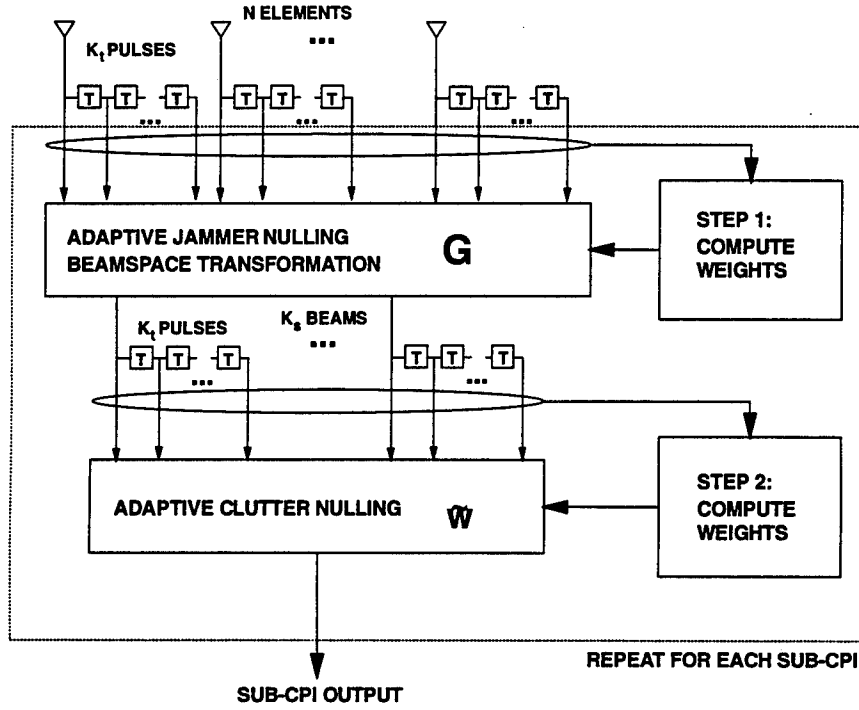


Figure 53. A two-step nulling approach to beamspace pre-Doppler STAP.

The exact form of the adaptive beamspace transform depends on whether a displaced-beam or adjacent-beam approach is used. With displaced-beam pre-Doppler, a separate jammer-nulling weight vector must be formed for each subaperture. Let \mathbf{J}_p be an $N \times N'$ selection matrix that chooses the elements for the p th subaperture, where $p = 0 : K_s - 1$. The adaptive beamformer $\mathbf{g}_p = [g_{p,0}; \dots; g_{p,N'-1}]$ is computed from

$$\mathbf{g}_p = \left(\mathbf{J}_p^T \Phi_{jn} \mathbf{J}_p \right)^{-1} \mathbf{g}_q, \quad (263)$$

where \mathbf{g}_q is the $N' \times 1$ subaperture desired response. Assume that \mathbf{g}_q is a tapered N' -element beamformer steered to the target direction,

$$\mathbf{g}_q = \mathbf{t} \odot \mathbf{a}(\vartheta_t; N'). \quad (264)$$

The resulting beamformer matrix is then

$$\tilde{\mathbf{G}} = \begin{bmatrix} g_{0,0} & & & \mathbf{0} \\ g_{0,1} & \ddots & & \\ \vdots & & g_{K_s-1,0} & \\ g_{0,N'-1} & & g_{K_s-1,1} & \\ & \ddots & \vdots & \\ \mathbf{0} & & g_{K_s-1,N'-1} \end{bmatrix} \quad (265)$$

The beamformer matrix for a two-step-nulling form of adjacent-beam pre-Doppler is found similarly. Let $\tilde{\mathbf{G}}_q$ be the desired beams in the absence of jamming. With jamming, the beamformer matrix is instead given by Within each sub-CPI, Φ_{jn} is used to construct adaptive beamformers that null the jamming as part of the beamspace transformation. The resulting beam outputs, with the jamming removed, are then used for adaptive clutter suppression as before. Thus, interference cancellation occurs in two steps, the first being to suppress jamming and the second to remove the clutter. Within each sub-CPI, Φ_{jn} is used to construct adaptive beamformers that null the jamming as part of the beamspace transformation. The resulting beam outputs, with the jamming removed, are then used for adaptive clutter suppression as before. Thus, interference cancellation occurs in two steps, the first being to suppress jamming and the second to remove the clutter.

$$\tilde{\mathbf{G}} = \Phi_{jn}^{-1} \tilde{\mathbf{G}}_q \quad (266)$$

Once the adaptive beamformer matrix $\tilde{\mathbf{G}}$ is found from Equations (265) or (266), the beamspace snapshots (with jamming removed) are formed. Clutter nulling and subsequent Doppler filtering proceed exactly as before. In practice, any jamming that is not suppressed in the first step, for whatever reason, may have a severe impact on the clutter nulling performance.

The two-step nulling architecture is attractive for several reasons. First, it does not increase the dimensionality of the clutter nulling step. Instead, it adds a jammer nulling problem that must be solved for every PRI. Typically, jammer nulling is a low-dimensional problem. However, two-step-nulling is predicated on the ability to obtain clutter-free jamming samples that are valid estimates of the jamming that occurs during the time interval when clutter is present. This in turn may depend on the radar system parameters and the characteristics, particularly the short-term stationarity of the expected jamming signals.

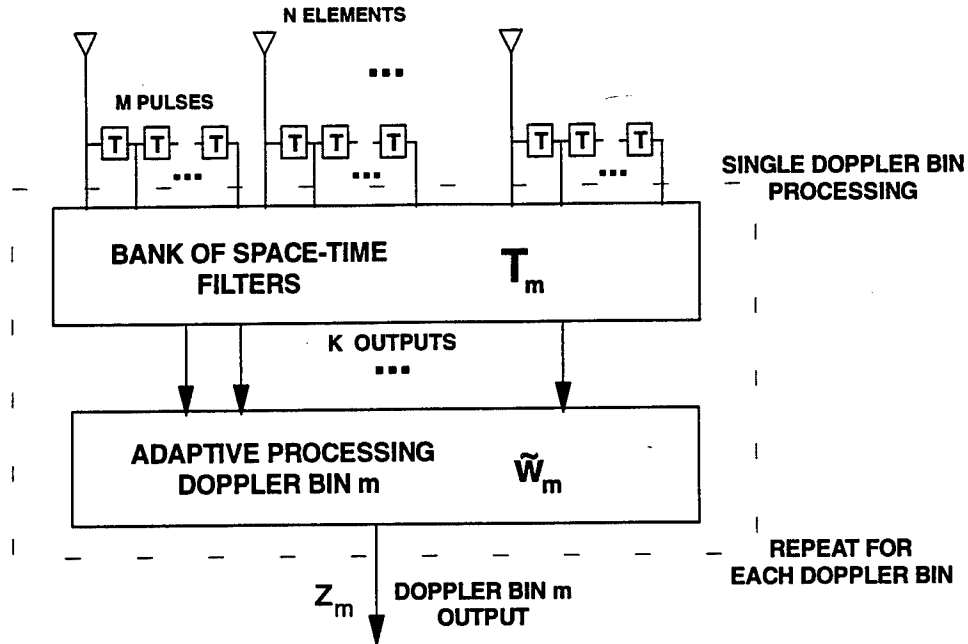


Figure 54. Generic block diagram for beamspace post-Doppler. Single Doppler bin processing is shown.

A novel two-step nulling approach that utilizes a sparse selection of elements and pulses for adaptation is described by Baranoski [36]. His approach, called sparse network processing, more closely resembles element-space pre-Doppler processing because of the way the selected signals are combined.

6.3 Beam-Space Post-Doppler STAP

This section discusses the fourth and final quadrant of the algorithm taxonomy that was presented in Figure 33. Preprocessors consisting of a bank of space-time filters are considered. These filters may be formed by cascading spatial beamformers on each pulse with Doppler filters on each beam; this class is therefore called *beamspace post-Doppler* STAP. A generic block diagram is shown in Figure 54. The filtered signals are then adaptively combined to produce the Doppler bin output. This process is then repeated for each Doppler bin. Combined beamforming and Doppler filtering can provide substantial suppression of portions of the interference, thereby localizing the interference prior to adaptation. Significant reductions in the number of adaptive degrees of freedom are possible with beamspace post-Doppler processing.

Beamspace post-Doppler techniques are the extensions of the beamspace pre-Doppler and element-space post-Doppler approaches to two-dimensional preprocessors. Both “displaced-filter” and “adjacent-filter” methods will be developed and other types of filter allocation discussed.

6.3.1 Basic Analysis

As with all post-Doppler approaches, beamspace post-Doppler requires solving a separate adaptive problem in each target Doppler bin. Let \mathbf{T}_m be the $MN \times K$ preprocessor for the m th bin adaptation, where K is the size of the reduced-dimension snapshot. Attention will be restricted to preprocessors of two types:

1. $\mathbf{T}_m = \tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}$ (267)

2. $\mathbf{T}_m = (\mathbf{F} \otimes \mathbf{G})\mathbf{J}_m$. (268)

In Equation (267), $\tilde{\mathbf{F}}_m$ is an $M \times K_t$ matrix of Doppler filters, $\tilde{\mathbf{G}}$ is $N \times K_s$ matrix of beamformers, and $K = K_t K_s$. This type of preprocessor is said to be *separable* because it may be implemented by cascading multiple beamformers on each pulse with multiple Doppler filters on each beam (or vice versa). Separability in this sense is more general than factored processing; factored processing is a special case when $K_t = 1$. The second type of preprocessor is formed by choosing a subset of the outputs of a separable processor. In Equation (268), \mathbf{F} is $M \times M$, \mathbf{G} is $N \times N$, and \mathbf{J}_m is an $MN \times K$ selection matrix that picks a subset of angle-Doppler filters. A transform of this second type is separable only if the selection matrix is itself separable. Both preprocessor types can be implemented with two-step nulling if appropriate clutter-free data exists to design the beamformer portion of the preprocessor adaptively.

The separable preprocessor, Equation (267), will be focused on first. Assume that both $\tilde{\mathbf{F}}_m$ and $\tilde{\mathbf{G}}$ are of full column rank. The transformed snapshot for the m th bin is given by

$$\tilde{\chi}_m = (\tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}})^H \chi. \quad (269)$$

The snapshot $\tilde{\chi}_m$ consists of an interference-plus-noise component $\tilde{\chi}_{um}$ and, if a target is present, the target component

$$\tilde{\chi}_{tm} = (\tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}})^H (\alpha_t \mathbf{v}_t) = \alpha_t \tilde{\mathbf{F}}_m^H \mathbf{b}_t \otimes \tilde{\mathbf{G}}^H \mathbf{a}_t. \quad (270)$$

The adaptive weight vector for the m th Doppler bin is computed from

$$\tilde{\mathbf{w}}_m = \tilde{\mathbf{R}}_{um}^{-1} \tilde{\mathbf{u}}_t, \quad (271)$$

where $\tilde{\mathbf{R}}_{um}$ is now the $K_s K_t \times K_s K_t$ interference-plus-noise covariance matrix and $\tilde{\mathbf{u}}_t$ is a $K_s K_t \times 1$ desired response. The desired response is chosen according to Equation (141),

$$\tilde{\mathbf{u}}_t = (\tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}})^H \mathbf{g}_t . \quad (272)$$

The final output of the m th Doppler bin is given by

$$z_m = \tilde{\mathbf{w}}_m^H \tilde{\chi}_m = \mathbf{w}_m^H \chi , \quad (273)$$

where the beamspace post-Doppler composite weight vector is given by

$$\mathbf{w}_m = (\tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}) \tilde{\mathbf{w}}_m . \quad (274)$$

The SINR performance can then be computed from Equations (185) and (186) in the usual ways.

Consider the space-time covariance matrix after preprocessing. The thermal noise component has covariance matrix

$$\tilde{\mathbf{R}}_{nm} = \sigma^2 \tilde{\mathbf{F}}_m^H \tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}^H \tilde{\mathbf{G}} . \quad (275)$$

The noise is still full rank but is now colored both spatially and temporally by the preprocessor. The jamming signals produce a covariance matrix

$$\tilde{\mathbf{R}}_{jm} = \tilde{\mathbf{F}}_m^H \tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}^H \Phi_j \tilde{\mathbf{G}} . \quad (276)$$

The jamming is colored temporally and filtered spatially by the beamformer matrix.

Finally, the clutter covariance matrix after preprocessing is found to be

$$\tilde{\mathbf{R}}_{cm} = \sum_{k=1}^{N_c} \sigma^2 \xi_k \tilde{\mathbf{F}}_m^H \mathbf{b}_k \mathbf{b}_k^H \tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}^H \mathbf{a}_k \mathbf{a}_k^H \tilde{\mathbf{G}} . \quad (277)$$

For given K_s and K_t , the rank of the clutter covariance depends upon both $\tilde{\mathbf{F}}_m$ and $\tilde{\mathbf{G}}$. A beamspace post-Doppler version of Brennan's rule provides conditions for the design of the preprocessor that result in minimum clutter rank.

Theorem 6 *If the assumptions underlying Brennan's rule hold and if the following two conditions*

1. *There exists a $K_s \times K_s$ nonsingular matrix \mathbf{Q}_s and a length $N' = N - K_s + 1$ vector $\mathbf{g} = [g_0; g_1; \dots; g_{N'-1}]$ such that*

$$\tilde{\mathbf{G}}\mathbf{Q}_s = \begin{bmatrix} g_0 & & & \mathbf{0} \\ g_1 & & \ddots & \\ \vdots & & & g_0 \\ g_{N'-1} & & & g_1 \\ & & \ddots & \vdots \\ \mathbf{0} & & & g_{N'-1} \end{bmatrix} \quad (278)$$

2. *There exists a $K_t \times K_t$ nonsingular matrix \mathbf{Q}_t and a length $M' = M - K_t + 1$ vector \mathbf{f} such that*

$$\tilde{\mathbf{F}}_m\mathbf{Q}_t = \begin{bmatrix} f_0 & & & \mathbf{0} \\ f_1 & & \ddots & \\ \vdots & & & f_0 \\ f_{M'-1} & & & f_1 \\ & & \ddots & \vdots \\ \mathbf{0} & & & f_{M'-1} \end{bmatrix} \quad (279)$$

are satisfied, then the clutter rank is minimum and equal to

$$\tilde{r}_c = \text{rank}(\tilde{\mathbf{R}}_{cm}) = K_s + (K_t - 1)\beta. \quad (280)$$

Thus, if the space-time preprocessor possesses a Toeplitz structure in space and time, the clutter rank is given by Equation (280). Theorem 6 is illustrated in Figure 55 for the case $N = 4$, $M = 3$, $K_s = K_t = 2$, and $\beta = 1$. Here $N' = 3$ and $M' = 2$. The element-pulse groupings for filters 2 and 3 are effectively from the same position; therefore, they produce the same clutter signal. As clutter observations are repeated by different filter outputs, the clutter rank is less than full. In this case it is three, which is as predicted by Equation (280). If Theorem 6 is satisfied, the transform $\tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}}$ is equivalent to a two-dimensional convolution of the $M \times N$ input snapshot with the $M' \times N'$ space-time filter

$$\mathbf{h} = \mathbf{f} \otimes \mathbf{g}, \quad (281)$$

where the $K_s K_t$ meaningful output samples form the signals to be adaptively combined.

The proof of Equation (277) follows as before. Only $\mathbf{Q}_s = \mathbf{I}_{K_s}$ and $\mathbf{Q}_t = \mathbf{I}_{K_t}$ need to be considered. If Equations (278) and (279) are satisfied, a target component after preprocessing has form

$$\tilde{\mathbf{F}}_m^H \mathbf{b}(\varpi) \otimes \tilde{\mathbf{G}}^H \mathbf{a}(\vartheta) = (\mathbf{f}^H \mathbf{b}(\varpi; M')) \tilde{\mathbf{b}}(\varpi) \otimes (\mathbf{g}^H \mathbf{a}(\vartheta; N')) \tilde{\mathbf{a}}(\vartheta), \quad (282)$$

where $\tilde{\mathbf{b}}(\varpi) = \mathbf{b}(\varpi; K_t)$ and $\tilde{\mathbf{a}}(\vartheta) = \mathbf{a}(\vartheta; K_s)$. The clutter covariance matrix is then

$$\tilde{\mathbf{R}}_{cm} = \sum_{k=1}^{N_c} \sigma^2 \xi_k |\mathbf{f}^H \mathbf{b}(\varpi_k; M')|^2 |\mathbf{g}^H \mathbf{a}(\vartheta_k; N')|^2 \tilde{\mathbf{b}}_k \tilde{\mathbf{b}}_k^H \otimes \tilde{\mathbf{a}}_k \tilde{\mathbf{a}}_k^H \quad (283)$$

$$= \sum_{k=1}^{N_c} \sigma^2 \tilde{\xi}_k \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^H, \quad (284)$$

where

$$\tilde{\xi}_k = \xi_k |\mathbf{f}^H \mathbf{b}(\varpi_k; M')|^2 |\mathbf{g}^H \mathbf{a}(\vartheta_k; N')|^2 \quad (285)$$

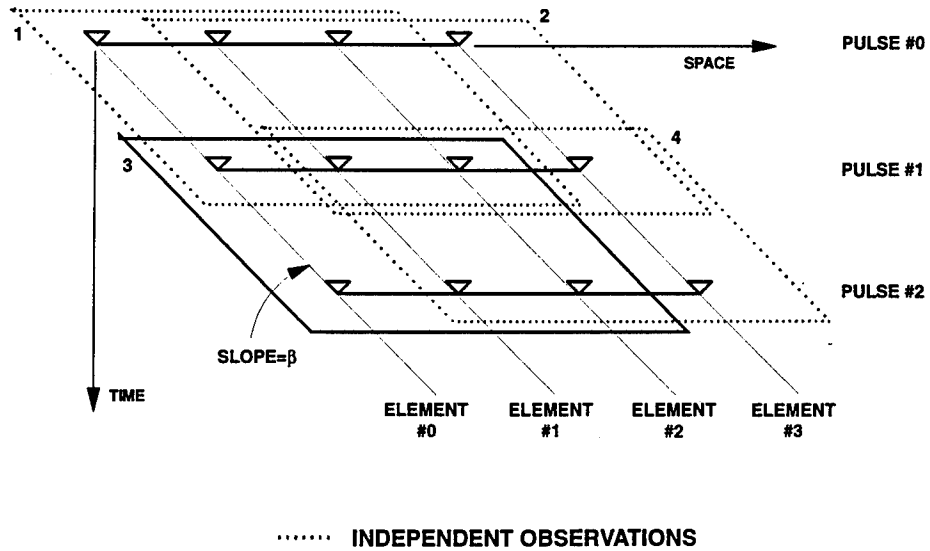


Figure 55. Beamspace post-Doppler variation on Brennan's Rule. $N = 4$, $M = 3$, $K_s = K_t = 2$, $\beta = 1$. If the conditions are satisfied, clutter observations are repeated by different outputs.

is the filtered CNR for the k th clutter patch and $\tilde{\mathbf{v}}_k = \tilde{\mathbf{b}}_k \otimes \tilde{\mathbf{a}}_k$ is a K_t -pulse, K_s -element space-time steering vector. Application of Brennan's rule to Equation (284) yields the desired result. The extensions to nonsingular \mathbf{Q}_s and \mathbf{Q}_t follow directly, as multiplication by a nonsingular matrix will not change the rank.

In practice, the numerical rank of the clutter covariance matrix may be even less than Equation (280), as parts of the clutter ridge may be suppressed to well below thermal noise by the combined angle and Doppler sidelobes of the preprocessor.

To differentiate between beamspace post-Doppler approaches, the special case $\mathbf{Q}_s = \mathbf{I}_{K_s}$, $\mathbf{Q}_t = \mathbf{I}_{K_t}$ will be referred to as *displaced-filter beamspace post-Doppler*. Another appropriate name, in view of Figure 55 and Equation (281), might be *space-time filter-then-adapt*.

6.3.2 Adjacent-filter beamspace post-Doppler

Now consider the case where the preprocessor is given by Equation (268),

$$\mathbf{T}_m = (\mathbf{F} \otimes \mathbf{G})\mathbf{J}_m. \quad (286)$$

This type of architecture has been studied by Cai and Wang [10], and Wicks [39], and is shown in Figure 56. In their algorithm, the input snapshot is passed through a two-dimensional DFT that may be represented as

$$\mathbf{T} = \mathbf{F} \otimes \mathbf{G}, \quad (287)$$

where

$$\mathbf{F} = \text{diag}(\mathbf{t}_f)\mathbf{U}_M^*, \quad \mathbf{G} = \text{diag}(\mathbf{t}_g)\mathbf{U}_N^*. \quad (288)$$

\mathbf{U}_M and \mathbf{U}_N are $M \times M$ and $N \times N$ DFT matrices, respectively, and $\mathbf{t}_f, \mathbf{t}_g$ are tapers in Doppler and angle. The output for the m th bin target filter is formed by adaptively combining the signals from a subset of K_m filters denoted by the $MN \times K_m$ selection matrix \mathbf{J}_m . The chosen subset must contain the target filter. The performance of this approach will depend on the preprocessor filter sidelobes and on the subset selection method.

The snapshot for the m th bin adaptation is given by

$$\begin{aligned}\tilde{\chi} &= ((\mathbf{F} \otimes \mathbf{G})\mathbf{J}_m)^H \chi \\ &= \mathbf{J}_m^T (\mathbf{F} \otimes \mathbf{G})^H \chi .\end{aligned}\tag{289}$$

The issue is now one of filter selection, i.e. the choice of \mathbf{J}_m for each Doppler bin.

One selection strategy is to choose a rectangular block of filters that is centered on and includes the target filter. In this strategy, *adjacent-filter beamspace post-Doppler*, $K_m = K_{sm}K_{tm}$, where K_{sm} is the number of angle beams and K_{tm} is the number of Doppler filters in the m th block. This selection method is illustrated in Figure 57. The 2D-DFT provides a grid of filters whose center frequencies (steering directions and Dopplers) are indicated by the circles. For the target filter at zero azimuth and 100-Hz Doppler (Doppler bin 6), the rectangle indicates the 3×3 block of $K_m = 9$ filters that are adaptively weighted. The block moves with the target filter as each Doppler (or angle) bin is processed. The adjacent filter selection strategy is the approach suggested by Cai and Wang [10].

The adjacent-filter selection matrix is separable,

$$\mathbf{J}_m = \mathbf{J}_{tm} \otimes \mathbf{J}_{sm} ,\tag{290}$$

where \mathbf{J}_{tm} and \mathbf{J}_{sm} are appropriate $M \times K_t$ and $N \times K_s$ selection matrices. This property leads to the separable preprocessor

$$\mathbf{T}_m = \tilde{\mathbf{F}}_m \otimes \tilde{\mathbf{G}},\tag{291}$$

where

$$\tilde{\mathbf{F}}_m = \mathbf{F}\mathbf{J}_{tm} \quad \text{and} \quad \tilde{\mathbf{G}} = \mathbf{G}\mathbf{J}_{sm} .\tag{292}$$

It can be shown that the preprocessor for adjacent-filter beamspace post-Doppler satisfies Theorem 6 when the 2D-DFT is not tapered.

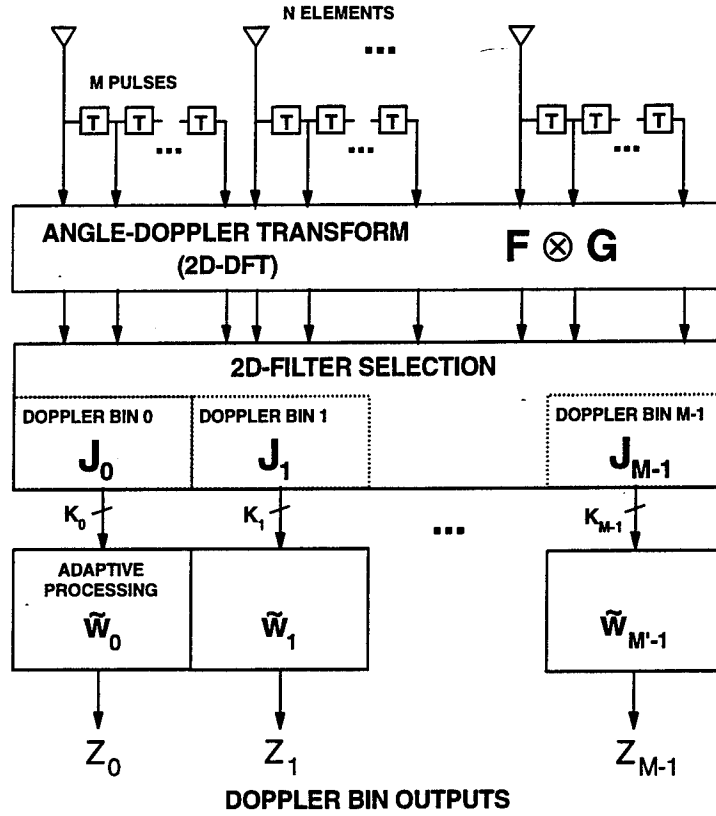


Figure 56. Block diagram for beamspace post-Doppler utilizing a single 2D-DFT, Cai, Wang, and Wicks[10,39].

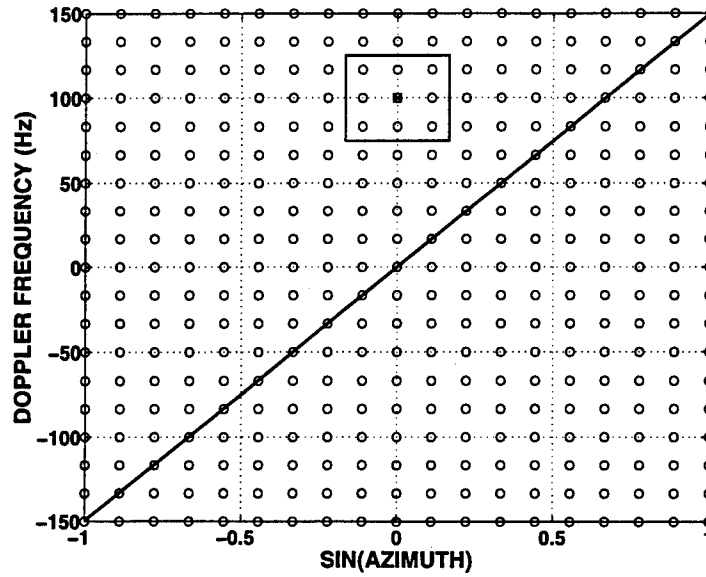


Figure 57. Filter selection for adjacent-filter beamspace post-Doppler. Key: 'o'-available filters, '⊗'-target filter. The box encircles the selected filters for $K_{sm} = 3$, $K_{tm} = 3$.

There are many other methods of filter selection. Given a target filter, one could choose the interference cancellation filters to be those filters with maximum interference power. Alternately, one could select the set of filters that have maximum cross-correlation with the target filter. Yet another method is to select those filters whose center frequencies are closest to the expected location of the clutter ridge. The post-Doppler beamspace architectures discussed thus far have utilized either displaced space-time filters or adjacent space-time filters. A set of filters from adjacent beams but displaced temporally, or vice versa, could also be considered. One of these methods may be appropriate for a specific application.

Jamming or combined clutter and jamming can be handled the same ways as for the beamspace pre-Doppler approaches. Simultaneous jammer and clutter suppression is best performed when some filters have high gain in the jammer directions. A filter-selection method should therefore have provisions to select filters specifically for jammer degrees of freedom. If a two-step nulling architecture is preferred, the spatial beamformer portion of the beamspace post-Doppler preprocessor will be adaptively computed to null the jamming prior to clutter suppression in each Doppler bin. Two-step nulling with adjacent and displaced beams was discussed in Section 6.2.4.

6.3.3 Performance Results

The performance of displaced and adjacent-filter beamspace post-Doppler will be compared first in an environment of clutter only. The example system and clutter scenario of Tables 2 and 4 is assumed. Figure 58 shows SINR loss plots for the case $K = 4$ degrees of freedom, with $K_{tm} = K_{sm} = 2$. Note that with uniform taper, both displaced filter and adjacent filter post-Doppler provide near optimum performance with four adaptive degrees of freedom. Results are also shown for tapered filters, assuming 30-dB Chebyshev in both angle and Doppler. While the displaced-filter approach, which satisfies Theorem 6 regardless of taper, suffers only a small taper loss over the whole Doppler space, the adjacent-filter approach suffers a significant loss in performance at Doppler frequencies close to mainlobe clutter. Figures 59 and 60 show similar plots as the dimensionality is increased to $K_{sm} = 3, K_{tm} = 3$, and then $K_{sm} = 4, K_{tm} = 4$. As the number of selected filters increases, the relative difference between displaced and adjacent filter approaches lessens, as there are plenty of degrees of freedom to suppress the clutter.

Now add the jamming scenario of Table 4. Figures 61 through 63 compare performance with and without two-step nulling (TSN), assuming the displaced-filter beamspace post-Doppler approach. In those curves with two-step nulling, the beamspace portion of the preprocessor is computed adaptively, assuming clutter-free data is available. Thus, the first jammer nulling step is an 18-element spatially adaptive solution. The plots are for different numbers of degrees of freedom in the beamspace post-Doppler nulling. Without TSN, jamming and clutter must be nulled simultaneously, which requires more degrees of freedom than nulling clutter alone. Having only $K_{tm} = K_{sm} = 2$ is not have enough and performance without TSN is poor. With TSN and the jamming removed prior to clutter nulling, performance is nearly equal to that of the clutter-only case. As more adaptive degrees of freedom are added, realizing that the clutter in this case obeys Theorem 6, more excess degrees of freedom are available to deal with the jamming. With $K = 16$ or $K_{tm} = K_{sm} = 4$, beamspace post-Doppler without TSN has enough degrees of freedom to handle the combined clutter and two-jammer environment. For a fixed number of degrees of freedom, it has also been observed that the performance without TSN may also vary with the tapering on the preprocessor. Heavy tapering can suppress much of the clutter nonadaptively, thereby allowing the adaptive weights to better suppress jamming.

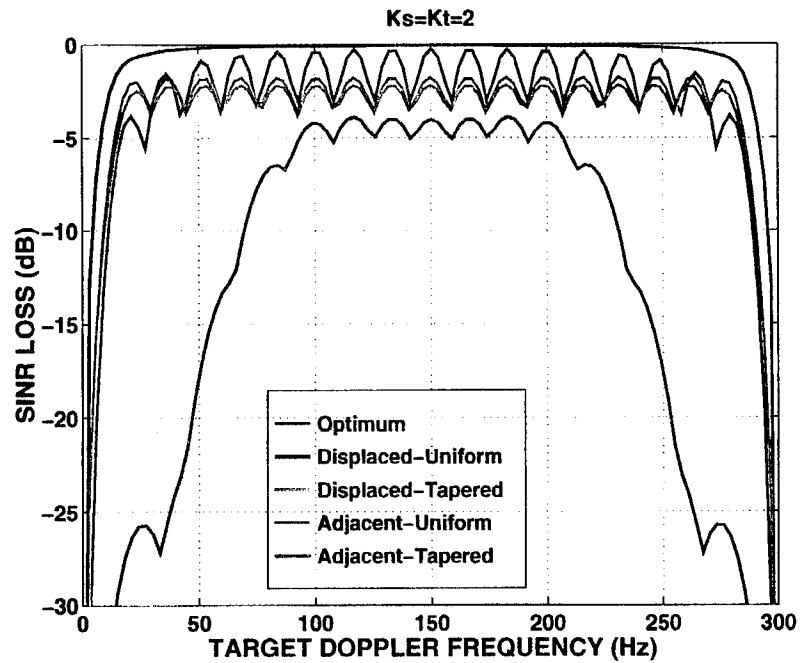


Figure 58. Beamspace post-Doppler in a clutter-only environment. $K = 4$, $K_{sm} = K_{tm} = 2$.

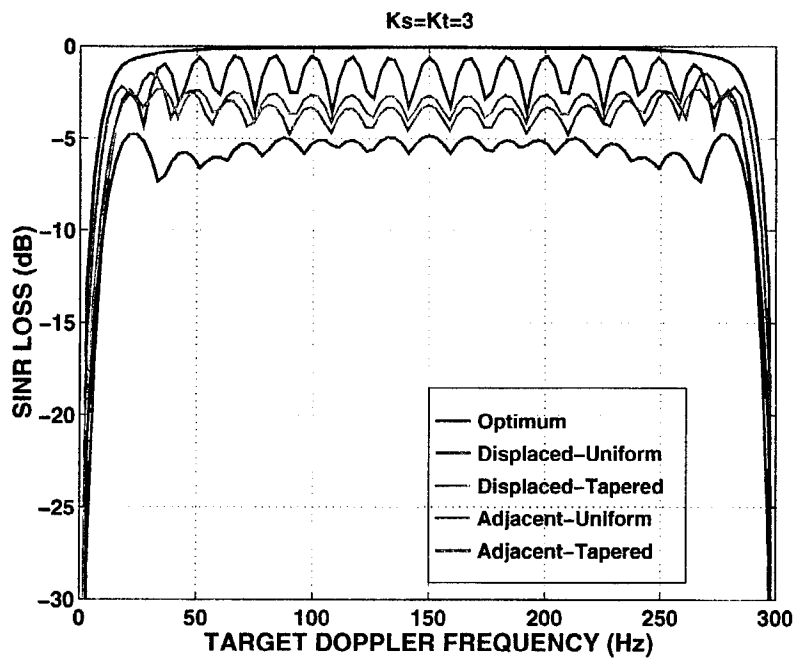


Figure 59. Beamspace post-Doppler in a clutter-only environment. $K = 9$, $K_{sm} = K_{tm} = 3$.

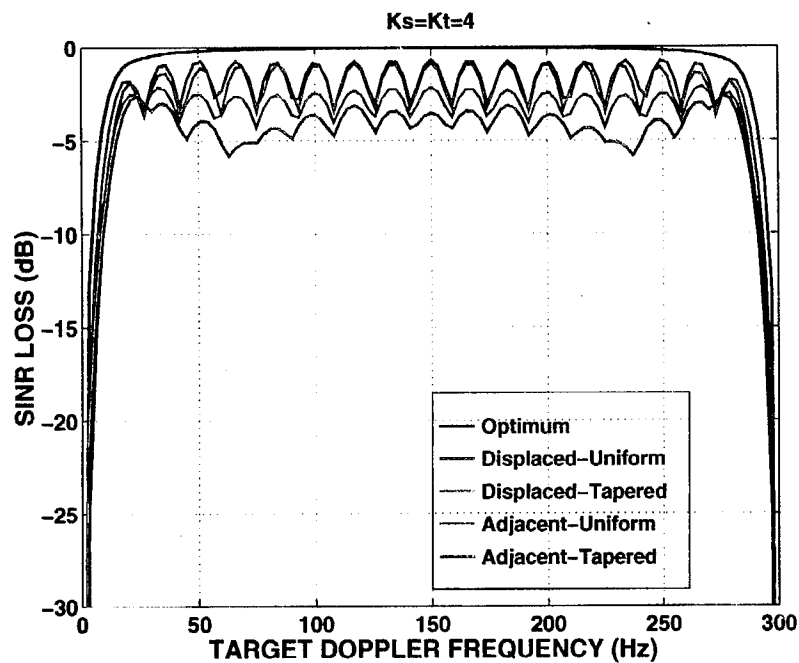


Figure 60. Beamspace post-Doppler in a clutter-only environment. $K = 16$, $K_{sm} = K_{tm} = 4$.

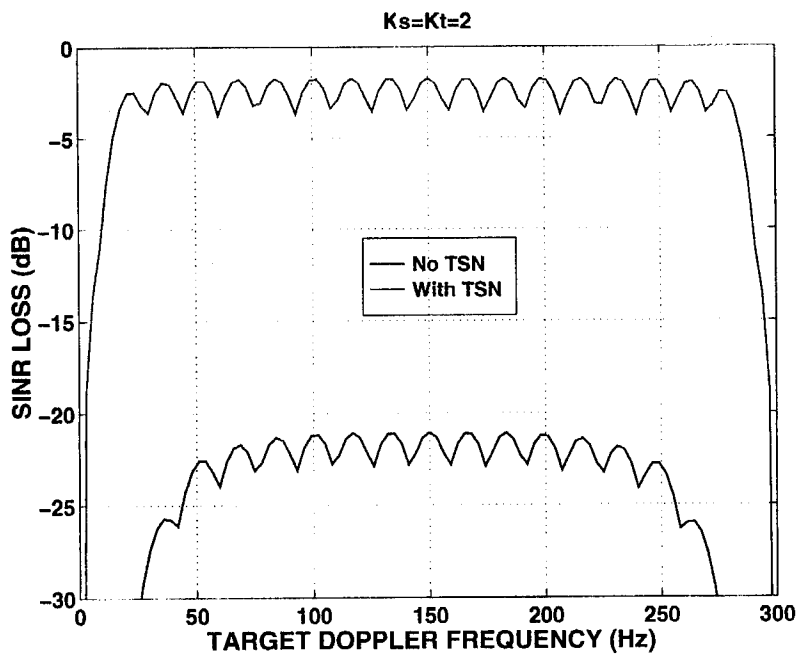


Figure 61. Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 4$, $K_{sm} = K_{tm} = 2$.

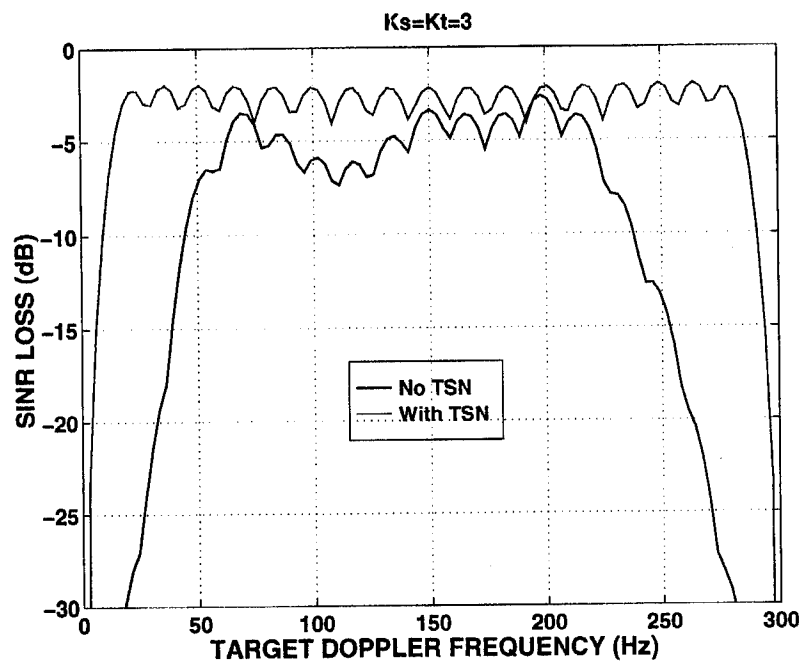


Figure 62. Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 9$, $K_{sm} = K_{tm} = 3$.

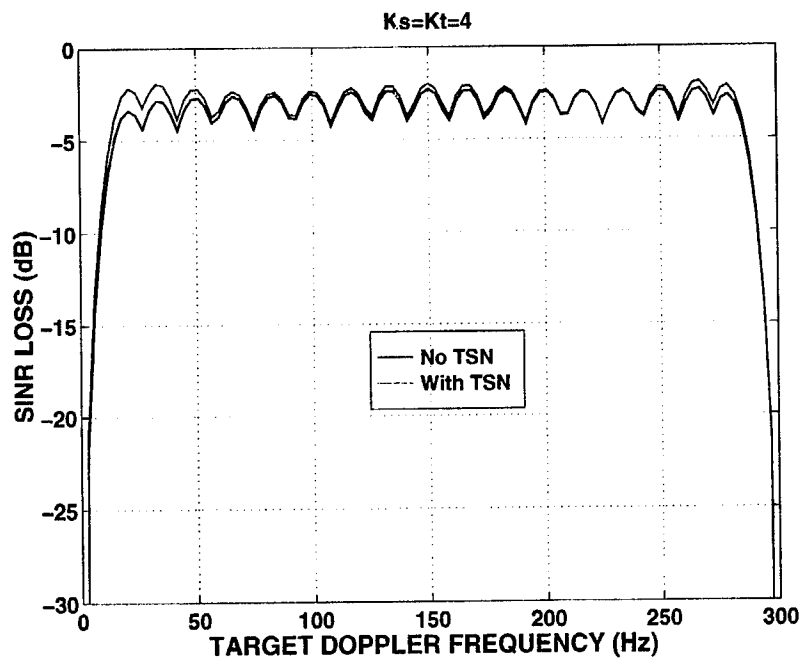


Figure 63. Beamspace post-Doppler in a clutter-plus-jamming environment. $K = 16$, $K_{sm} = K_{tm} = 4$.

7. ADDITIONAL PERFORMANCE RESULTS

The previous chapters have been devoted to detailed discussions of particular space-time processing approaches. This chapter presents simulation results that are chosen to illustrate relative performance for some other typical scenarios. This chapter is not meant to be a comprehensive evaluation leading to a "best" approach. Rather, it is intended to highlight some of the important issues that should be considered when selecting an approach for a specific radar system platform.

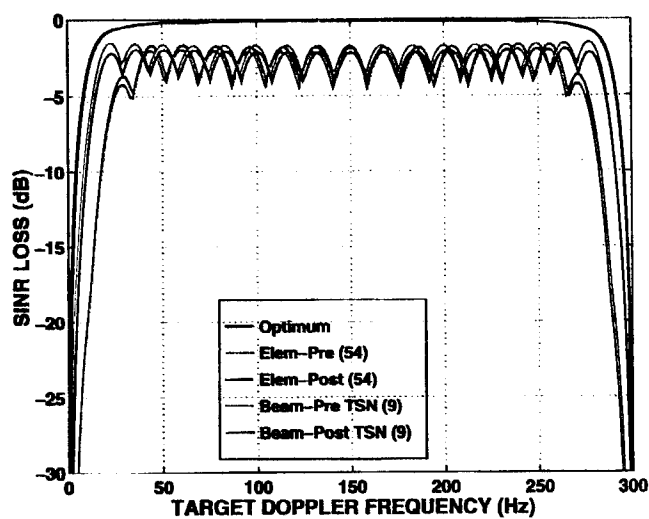
Although partially adaptive STAP architectures have been categorized into four major classes, a large number of variations, depending on bin or beam selection, the inclusion of two-step nulling, etc., have been discussed. This discussion will consider the optimum fully adaptive approach and one algorithm from each of the four classes. The algorithms to be examined are:

1. Optimum fully adaptive
2. Element-space pre-Doppler
3. Element-space post-Doppler: PRI-staggered post-Doppler
4. Beamspace pre-Doppler: Displaced beams with two-step nulling
5. Beamspace post-Doppler: Displaced filters with two-step nulling.

All steering vectors, beamformers, and Doppler filters are designed with a 30-dB Chebyshev taper. Some approaches are more dependent on the specific taper than others. For the beamspace approaches, assume that appropriate clutter-free data exists so that ideal two-step nulling can be implemented. By definition, the beamspace approaches are lower dimension than element-space approaches with the same number of temporal degrees of freedom. For the same level of performance, the beamspace approaches have the advantage of less computational complexity and training requirements.

The baseline system and scenario of Tables 2 and 4 will be used again. Recall that the major attributes of this example system are low PRF, UHF, with a short CPI (18 pulses or 50 ms) and a moderate-size aperture of 18 elements. The baseline scenario consists of heavy clutter, 47 dB CNR per element, and two jamming sources that combine for a 38-dB JNR per element. In addition to SINR loss, minimum detectable velocity (MDV) and usable Doppler space fractions (UDSF), which were defined in Chapter 3, will be tabulated. These quantities are computed at 5-dB and 12-dB cutoff points, representing values of SINR loss that would cause a loss in range coverage of 25% and 50%, respectively.

For the baseline scenario, the performance of the five algorithms is plotted in Figure 64. In this case all four partially adaptive approaches provide near-optimum performance. The post-Doppler approaches have slightly better MDV, resulting in better UDSF. In this case the beamspace approaches with only nine degrees of freedom perform as well as their element-space counterparts.



(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	2.74	1.14	0.94	0.97
Elem-Pre	54	7.81	5.11	0.84	0.89
Elem-Post	54	4.42	2.44	0.91	0.95
Beam-Pre	9	8.26	5.26	0.81	0.89
Beam-Post	9	4.74	2.58	0.90	0.94

(b)

Figure 64. Performance for example system, $\beta = 1$, no misalignment, no intrinsic clutter motion. (a) SINR loss. (b) MDV and UDSF performance.

For the second scenario, the platform velocity is increased so that $\beta = 2.6$; all other parameters are held fixed. The results are plotted in Figure 65. As the increased slope of the clutter ridge causes mainlobe clutter to spread over a larger region in Doppler, the resultant nulls in the SINR loss curves are wider, and a little more of the Doppler space is lost. Both element-space techniques perform well, as expected. Beam-space approaches with $K_s = 3$ and $K_s = 5$ are shown. With only three beams, the Doppler-ambiguous clutter is nearly full rank. Performance suffers mainly around the mainlobe clutter notch, as evidenced by the poorer MDV and UDSF. With $K_s = 5$ beams, the beam-space performance is restored to nearly that of the element-space approaches.

Now add intrinsic clutter motion to the scenario of Figure 65. Choose a velocity standard deviation $\sigma_v = 0.2$ m/s that could represent antenna scan modulation, for example. The intrinsic motion broadens the clutter ridge in Doppler, which also results in wider clutter notches and a loss in Doppler space coverage. The results are shown in Figure 66.

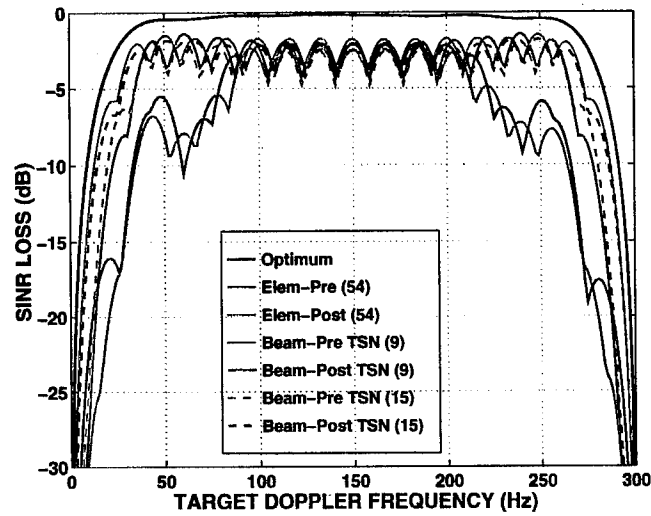
The combination of ICM and Doppler-ambiguous clutter causes a significant loss in the Doppler space to all algorithms. Again, observe that post-Doppler techniques generally provide slightly better UDSF, particularly at small values of SINR loss.

To complicate the scenario more, add velocity misalignment. Misalignment increases the clutter rank, so one can expect low-dimension algorithms catered to DPCA conditions to suffer. The degradation is particularly severe if strong backlobe clutter is present. The baseline system assumes a one-way backlobe level of -30 dB, which produces a backlobe clutter-to-noise ratio of -13 dB per element per pulse. Subsequent spatial and temporal integration with 18 elements and 18 pulses is sufficient to increase the backlobe clutter level above thermal noise.

First assume the misalignment angle to be 10° ; the clutter ridge for this scenario was given in Figure 14. Performance results are given in Figure 67. The strong backlobe clutter results in a second notch at the backlobe clutter Doppler frequency of 230 Hz. The need to suppress (even slightly) the backlobe clutter notch further reduces the amount of Doppler space with desirable performance. Pre-Doppler algorithms are degraded more than post-Doppler approaches; with both ICM and backlobe clutter, three temporal degrees of freedom are insufficient to place broad nulls at two Doppler frequencies. A significant loss in both MDV and UDSF is observed. When close to mainlobe clutter, post-Doppler algorithms benefit from the suppression of backlobe clutter achieved through the Doppler filter sidelobes. As a result, the difference in MDV and UDSF between pre- and post-Doppler algorithms is more pronounced.

When misalignment is present, a small improvement in backlobe level can improve performance significantly. Figure 68 considers the same scenario as Figure 67 but with a backlobe reduced to -40 dB. The result is a 20-dB reduction in backlobe clutter level, which is sufficient to effectively eliminate the problem. The SINR loss shows only a very slight dip at the backlobe clutter Doppler; performance over this part of the Doppler space is restored. Finally, consider the system of Figure 68, with the -40 -dB backlobe but at a misalignment angle of 90° ; this corresponds to a forward-looking scenario. The results are shown in Figure 69. With the low backlobe there is again only one notch in the SINR curves. The MDV and UDSF are noticeably better than in Figure 69, because the Doppler spread of the mainlobe clutter depends on the misalignment angle. In the forward-looking direction, the slope of the clutter ridge at the look direction is actually zero, as shown in the clutter ridge of Figure 13(d). The resulting Doppler spread over the angular extent of the mainlobe is minimum. Therefore, the clutter can be suppressed with a much narrower null in Doppler than for the sidelooking cases of Figures 67 and 68.

These results were intended to provide a few comparative results for a single set of radar system parameters. They were chosen to illustrate some of the major issues. There are certainly other areas for more detailed comparative analysis. This report concludes with a summary and some suggestions for further study.

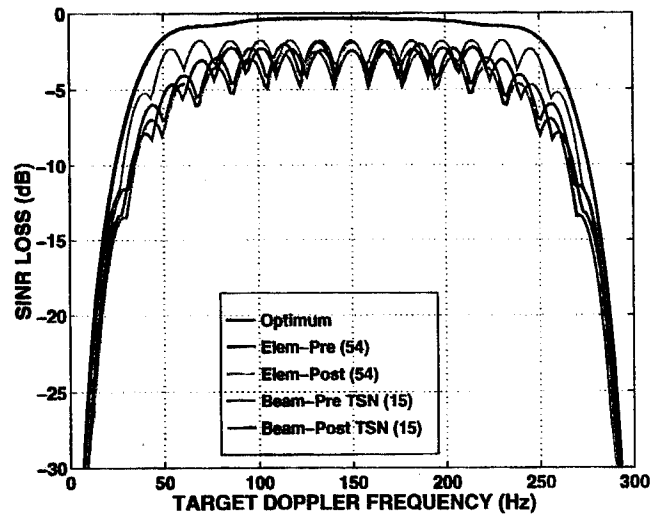


(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	5.55	2.32	0.88	0.95
Elem-Pre	54	11.48	6.48	0.76	0.86
Elem-Post	54	8.88	3.63	0.82	0.92
Beam-Pre	9	27.11	10.74	0.43	0.78
Beam-Post	9	26.20	10.39	0.47	0.79
Beam-Pre	15	11.52	5.85	0.76	0.88
Beam-Post	15	9.22	3.94	0.81	0.92

(b)

Figure 65. Performance for example system, $\beta = 2.6$, no misalignment, no intrinsic clutter motion. (a) SINR loss. (b) MDV and UDSF.

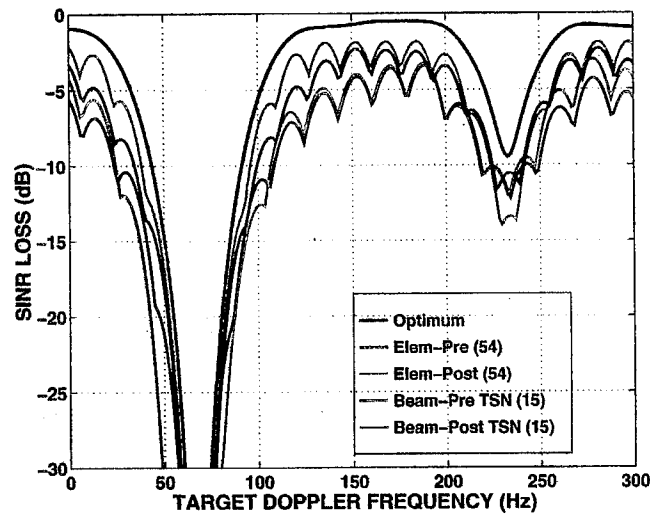


(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	11.59	7.01	0.76	0.85
Elem-Pre	54	18.40	8.50	0.61	0.82
Elem-Post	54	15.06	8.35	0.69	0.82
Beam-Pre	15	19.12	10.63	0.53	0.78
Beam-Post	15	17.14	9.71	0.62	0.80

(b)

Figure 66. Performance for example system, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, no velocity misalignment. (a) SINR Loss (b) MDV and UDSF.

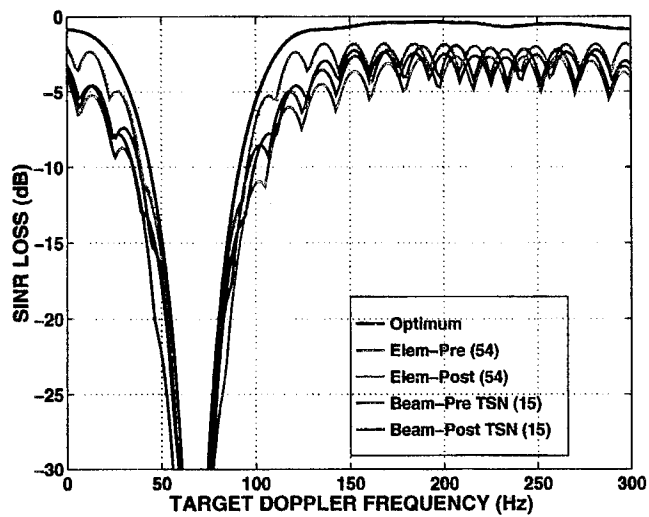


(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	11.47	6.93	0.69	0.86
Elem-Pre	54	27.23	10.10	0.16	0.79
Elem-Post	54	15.04	8.73	0.54	0.78
Beam-Pre	15	21.80	12.45	0.24	0.74
Beam-Post	15	17.18	9.71	0.44	0.81

(b)

Figure 67. Performance for example system, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 10° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.

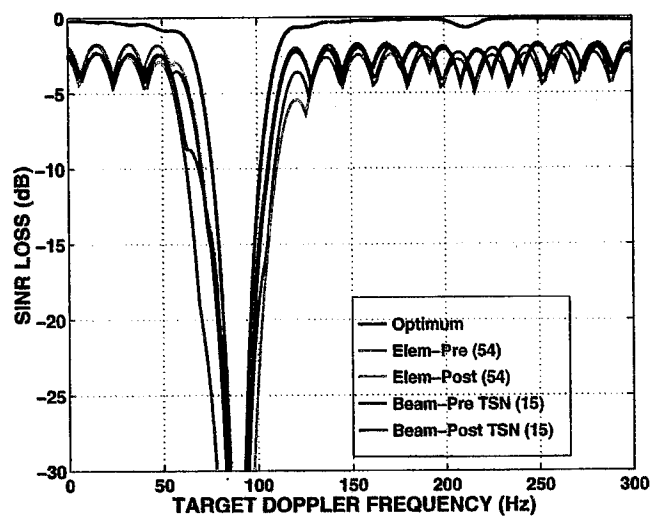


(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	11.45	6.93	0.77	0.86
Elem-Pre	54	16.73	8.63	0.60	0.83
Elem-Post	54	14.08	8.20	0.70	0.84
Beam-Pre	15	21.33	9.66	0.51	0.81
Beam-Post	15	16.96	9.67	0.64	0.81

(b)

Figure 68. Performance for example system with -40 -dB backlobe, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 10° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.



(a)

Algorithm	# DOF	MDV (m/s)		UDSF	
		-5 dB	-12 dB	-5 dB	-12 dB
Optimum	324	5.96	3.81	0.88	0.93
Elem-Pre	54	9.70	5.91	0.80	0.88
Elem-Post	54	7.81	5.20	0.84	0.90
Beam-Pre	15	12.04	7.31	0.76	0.85
Beam-Post	15	8.06	5.34	0.84	0.89

(b)

Figure 69. Performance for example system with -40 -dB backlobe, $\beta = 2.6$, $\sigma_v = 0.2$ m/s intrinsic clutter motion, 90° velocity misalignment. (a) SINR loss. (b) MDV and UDSF.

8. SUMMARY AND FUTURE WORK

This report has been devoted to space-time adaptive processing for airborne radar systems. From a generic radar system description, a model was developed for the signals received by the elements of an airborne radar antenna array. A space-time adaptive processor was defined to be an adaptive filter that simultaneously combines the signals from multiple array elements and multiple pulses of a coherent processing waveform. The issues of computational complexity and sample support for covariance estimation motivated the search for reduced-dimension, or partially adaptive, STAP algorithms that are both computation- and data-efficient. Adaptive array processing and partially adaptive array theory were utilized to develop a taxonomy of reduced-dimension STAP architectures. Approaches were classified by the type of preprocessor transform or, equivalently, by the data domain in which adaptive weighting occurs. Different combinations of spatial and/or Doppler filtering prior to adaptation comprise the four classes of architectures. Within each class, a number of variations were studied that involve different preprocessor designs or different approaches to cancelling jamming. The clutter seen by an airborne radar has a special structure induced by the platform motion; this structure can in many cases be exploited to design preprocessors that result in efficient STAP architectures.

Generally speaking, beamspace architectures (pre- or post-Doppler) result in lower-degree-of-freedom processors than element-space techniques. However, beamspace architectures must incorporate a robust method for handling jamming and combined clutter-and-jamming scenarios. Two-step, or sequential, nulling was discussed as a potential solution for those systems where appropriate clutter-free data can be obtained. Element-space approaches, provided with sufficient data for training, are inherently more flexible because full spatial adaptivity is retained. However, they require more adaptive degrees of freedom and therefore are likely only practical for small or moderately sized antenna arrays. Pre-Doppler approaches provide a more rapid adaptation capability than post-Doppler techniques. Post-Doppler algorithms provide slightly better Doppler space coverage and are more robust in scenarios with velocity misalignment and backlobe clutter.

Several issues relevant to a STAP implementation were not discussed in this report. One issue is the optimization of training strategies, considering the nonhomogeneous nature of radar clutter, and the constraints imposed by a target digital processor. The whole topic of mapping a STAP algorithm onto a digital processor was not discussed. Seemingly similar algorithms may impose dramatically different requirements on buffer memory, interprocessor communication bandwidth, latency, etc. The analysis herein focused on reducing the dimensionality for a single adaptive weight computation. The total computational complexity of a STAP algorithm depends not only on the number of adaptive weights but also, through the training method and the radar parameters, on the number of adaptive problems that must be solved for each CPI. The choice of an algorithm in practice should certainly be made with the target digital processor architecture in mind.

While STAP is but one portion of the airborne radar signal and data processing, an adaptive capability impacts several other functions. Achieving large levels of interference suppression in practice requires channel equalization to mitigate degradation due to channel mismatch. As discussed in Chapter 3, the weight computation can be done in conjunction with CFAR detection. This area merits further consideration, including post-detection data processing to edit clutter discretized and sidelobe targets. Finally, STAP will impact estimation of the angle and Doppler of target detections. Some work has been reported on maximum likelihood estimation for use with STAP [42].

The STAP architectures discussed in this report have considered adaptation in space and time on a PRI scale. In a wideband system seeking wideband jammer cancellation, or if multipath jamming must be cancelled [40, 42], additional adaptation in time on the scale of range gates may be necessary. The resulting architecture would have three dimensions of adaptivity: space, time (PRI), and time (range), or angle, Doppler, and instantaneous frequency. The investigation of efficient adaptive architectures for these scenarios is a subject for future work.

The motivation for this report has been the airborne early warning radar problem. Although not discussed herein, it is believed that the adaptive algorithms developed here have application to a wide variety of radar platforms. In fact, the inclusion of an adaptive capability may have an impact on basic aspects of radar system design. The ability to adaptively suppress clutter and jamming is a powerful capability that may impact basic aspects of radar system design. In the past, the need for ultra-low sidelobes or high-PRF waveforms has been driven largely by the need to suppress interference. The ability to do this adaptively may allow the rethinking of conventional wisdom in some areas. For example, medium- or low-PRF waveforms (with better range information) may be permissible in systems that had been primarily high PRF. A STAP capability also offers the potential for low-velocity target detection to some systems that currently are only able to operate in clutter-free zones. Investigations of the applicability of STAP as defined herein to other classes of radar systems is another area for future investigation.

APPENDIX A

PROOF OF BRENNAN'S RULE

In Chapter 2 an approximate rule for the rank of the clutter covariance matrix was provided. This rule, originally put forth by Brennan and Staudaher [18], was based on extensive analysis of simulation data. This appendix proves a special case of Brennan's rule, for β equal to an integer.

Theorem 1 (Brennan's rule): *If the interpulse motion per half-interelement spacing $\beta = 2v_a T_r / d$ is an integer less than or equal to N , then the rank of the clutter covariance matrix is*

$$r_c = N + (M - 1)\beta .$$

The assumptions are that the platform velocity vector is aligned with the linear array axis and that the element patterns are identical. The proof begins with the clutter covariance matrix from Equation (60):

$$\mathbf{R}_c = \mathbf{V}_c \mathbf{\Xi}_c \mathbf{V}_c^H , \quad (\text{A.1})$$

where \mathbf{V}_c is an $MN \times N_c$ matrix of the space-time steering vectors to each clutter patch, $\mathbf{\Xi}_c$ is a diagonal matrix of the power from each patch, and N_c is the number of clutter patches. Assume that a number of clutter patches is large $N_c > MN$. Because $\mathbf{\Xi}_c$ is positive definite,

$$r_c = \text{rank}(\mathbf{R}_c) = \text{rank}(\mathbf{V}_c) \quad (\text{A.2})$$

and only \mathbf{V}_c needs to be considered.

The linear relationship, Equation (70), between clutter Doppler and spatial frequency means that when β is an integer, all elements of the space-time steering vector $\mathbf{v}(\vartheta, \varpi(\vartheta))$ are integer multiples of the same base quantity $z = \exp(j2\pi\vartheta)$. Let $z_k = \exp(j2\pi\vartheta_k)$ correspond to the k th scatterer. Assume that the patches are distributed such that the $\{z_k\}$ are distinct.¹³

¹³This is guaranteed if the scatterers are distributed over $[-90^\circ, 90^\circ]$ and the interelement spacing is a half-wavelength.

The clutter steering vector matrix V_c is then given by

$$V_c = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & & z_{N_c} \\ \vdots & \vdots & & \vdots \\ z_1^{N-1} & z_2^{N-1} & & z_{N_c}^{N-1} \\ \hline z_1^\beta & z_2^\beta & & z_{N_c}^\beta \\ z_1^{\beta+1} & z_2^{\beta+1} & & z_{N_c}^{\beta+1} \\ \vdots & \vdots & & \vdots \\ z_1^{\beta+N-1} & z_2^{\beta+N-1} & & z_{N_c}^{\beta+N-1} \\ \hline \vdots & \vdots & & \vdots \\ \hline z_1^{(M-1)\beta} & z_2^{(M-1)\beta} & & z_{N_c}^{(M-1)\beta} \\ z_1^{(M-1)\beta+1} & z_2^{(M-1)\beta+1} & & z_{N_c}^{(M-1)\beta+1} \\ \vdots & \vdots & & \vdots \\ z_1^{(M-1)\beta+N-1} & z_2^{(M-1)\beta+N-1} & \dots & z_{N_c}^{(M-1)\beta+N-1} \end{bmatrix} \quad (A.3)$$

Of the MN elements in the k th column, only the 0th through the $(N-1 + (M-1)\beta)$ powers of z_k are represented. The rank of V_c is invariant under a permutation of its rows; rearrange the rows so that the distinct elements of each column occupy the top rows and the replicated entries occupy the bottom rows. The result is the new matrix \tilde{V}_c :

$$\tilde{V}_c = \begin{bmatrix} 1 & \dots & 1 \\ z_1 & & z_{N_c} \\ z_1^2 & & z_{N_c}^2 \\ \vdots & & \vdots \\ z_1^{(M-1)\beta+N-1} & \dots & z_{N_c}^{(M-1)\beta+N-1} \\ \hline \text{repeated rows} \end{bmatrix} = \begin{bmatrix} U_c \\ \text{repeated rows} \end{bmatrix} \quad (A.4)$$

The upper partition of $\tilde{\mathbf{V}}_c$, denoted by \mathbf{U}_c , is a Vandermonde matrix of dimension $N + (M - 1)\beta \times N_c$. The columns of \mathbf{U}_c are linearly independent because the z_k are distinct; therefore \mathbf{U}_c is of full rank [20]. The repeated rows in \mathbf{V}_c do not contribute to its rank. This yields the results

$$r_c = \text{rank}(\mathbf{U}_c) = \min(N + (M - 1)\beta, N_c) = N + (M - 1)\beta. \quad (\text{A.5})$$

Equation (A.5) is the desired result, and the proof is complete.

One can also prove that the clutter covariance matrix is of rank less than MN for certain rational values of β by using a Vandermonde argument similar to the proof above [41]. Even for these special cases, however, the eigenspectra exhibit a knee that is accurately predicted by the general form of Brennan's rule, and eigenvalues beyond this knee are very small and likely insignificant. For this reason these special cases are not of much practical interest, and the additional details are omitted.

APPENDIX B

PROOF OF THEOREM 2

Chapter 4 considered a covariance matrix consisting of two components,¹⁴

$$\mathbf{R}_u = \mathbf{R}_1 + \mathbf{R}_2 , \quad (\text{B.1})$$

with \mathbf{R}_1 being full rank and \mathbf{R}_2 of low rank $r_2 < MN$ with spectral decomposition

$$\mathbf{R}_2 = \mathbf{E}_2 \mathbf{\Lambda}_2 \mathbf{E}_2^H . \quad (\text{B.2})$$

A partially adaptive processor with preprocessor $\mathbf{T} = \mathbf{R}_1^{-1} \mathbf{B}$ was suggested, where \mathbf{B} is an $MN \times D$ beam matrix with $D < MN$. The beam matrix is also assumed to be of full column rank. The following theorem provided conditions for the design of the matrix \mathbf{B} under which the performance of the partially adaptive processor, with known covariance, will be identical to that of a fully adaptive processor.

Theorem 2 *If the matrix \mathbf{B} is designed such that*

$$\text{span} \{[\mathbf{g}_t, \mathbf{E}_2]\} \subset \text{span} \{\mathbf{B}\} , \quad (\text{B.3})$$

then the partially adaptive processor of Figure 31(a) achieves performance equal to that of the fully adaptive processor.

Another way of stating the condition Equation (B.3) is to require

$$\mathbf{P}_B [\mathbf{g}_t, \mathbf{E}_2] = [\mathbf{g}_t, \mathbf{E}_2] , \quad (\text{B.4})$$

where $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$ is a projection matrix.

To prove this theorem, the composite weight vector $\mathbf{w}_{pa} = \mathbf{T} \tilde{\mathbf{w}}$ will be shown to equal the fully adaptive weight vector if Equation (B.3) holds. The output signals and therefore the performance will then be the same for both processors.

¹⁴The proof in this section follows a path first suggested by D. Marshall.

The fully adaptive weight vector was shown in Equation (147) to be

$$\mathbf{w}_{fa} = \mathbf{R}_1^{-1} \mathbf{g}_t - \mathbf{R}_1^{-1} \mathbf{E}_2 \left(\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \Lambda_2^{-1} \right)^{-1} \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{g}_t . \quad (\text{B.5})$$

The partially adaptive weight vector is given by

$$\tilde{\mathbf{w}} = \tilde{\mathbf{R}}_u^{-1} \tilde{\mathbf{g}}_t , \quad (\text{B.6})$$

where

$$\tilde{\mathbf{R}}_u = \mathbf{T}^H \mathbf{R}_u \mathbf{T} = \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{R}_u \mathbf{R}_1^{-1} \mathbf{B} \quad (\text{B.7})$$

and $\tilde{\mathbf{g}}_t = \mathbf{T}^H \mathbf{g}_t = \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t$. The composite partially adaptive weight vector is then

$$\mathbf{w}_{pa} = \mathbf{T} \tilde{\mathbf{w}} = \mathbf{R}_1^{-1} \mathbf{B} \tilde{\mathbf{R}}_u^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t . \quad (\text{B.8})$$

$\tilde{\mathbf{R}}_u^{-1}$ is first found by substituting Equations (B.1) and (B.2) into Equation (B.7) and applying the matrix inversion lemma:

$$\begin{aligned} \tilde{\mathbf{R}}_u^{-1} &= \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} - \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 \times \\ &\quad \left[\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \Lambda_2^{-1} \right]^{-1} \times \\ &\quad \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} . \end{aligned} \quad (\text{B.9})$$

Equation (B.8) is then used to obtain the following cumbersome expression for the composite weight vector:

$$\begin{aligned} \mathbf{w}_{pa} &= \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t - \\ &\quad \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 \times \\ &\quad \left[\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \Lambda_2^{-1} \right]^{-1} \times \\ &\quad \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{B} \left(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t . \end{aligned} \quad (\text{B.10})$$

This last result applies for an arbitrary \mathbf{B} . To satisfy the condition of Equation (B.3) in Theorem 2 requires first that $D \geq r_2 + 1$ and that \mathbf{B} has form

$$\mathbf{B} = \begin{bmatrix} \mathbf{g}_t & \mathbf{E}_2 & \mathbf{F} \end{bmatrix} \mathbf{C}, \quad (\text{B.11})$$

where \mathbf{C} is a $D \times D$ invertible matrix. Also, \mathbf{F} is an $MN \times (D - r_2 - 1)$ matrix of additional degrees of freedom that turn out to be irrelevant here. It is seen from Equation (B.10) that because \mathbf{C} is invertible, it cancels itself out in all places; therefore, $\mathbf{C} = \mathbf{I}$ is taken without loss of generality.

The rest of the proof is easily seen once the following equality is observed:

$$\begin{aligned} \mathbf{I}_D &= (\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B} \\ &= (\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \begin{bmatrix} \mathbf{g}_t & \mathbf{E}_2 & \mathbf{F} \end{bmatrix}. \end{aligned} \quad (\text{B.12})$$

Equating the two sides in Equation (B.12) reveals that

$$(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t = \begin{bmatrix} 1 \\ \mathbf{0}_{(D-1) \times 1} \end{bmatrix} \quad (\text{B.13})$$

$$(\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 = \begin{bmatrix} \mathbf{0}_{1 \times r_2} \\ \mathbf{I}_{r_2} \\ \mathbf{0}_{(D-r_1-1) \times r_2} \end{bmatrix} \quad (\text{B.14})$$

so that

$$\mathbf{R}_1^{-1} \mathbf{B} (\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{g}_t = \mathbf{R}_1^{-1} \mathbf{g}_t \quad (\text{B.15})$$

$$\mathbf{R}_1^{-1} \mathbf{B} (\mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_1^{-1} \mathbf{E}_2 = \mathbf{R}_1^{-1} \mathbf{E}_2. \quad (\text{B.16})$$

This last result, when substituted into Equation (B.10), yields

$$\mathbf{w}_{pa} = \mathbf{R}_1^{-1} \mathbf{g}_t - \mathbf{R}_1^{-1} \mathbf{E}_2 \left(\mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{E}_2 + \mathbf{\Lambda}_2^{-1} \right)^{-1} \mathbf{E}_2^H \mathbf{R}_1^{-1} \mathbf{g}_t. \quad (\text{B.17})$$

From Equation (B.5), $\mathbf{w}_{pa} = \mathbf{w}_{fa}$ and the proof is complete.

APPENDIX C

DISPLACED PHASE CENTER ANTENNA (DPCA) PROCESSING

For a number of years, DPCA processing has referred to a class of techniques for suppressing the ground clutter seen by an airborne radar. This appendix reviews the basic DPCA concept. Many different variations on the DPCA theme have been considered for airborne and space-based radars; this discussion will focus on the type of DPCA described by Staudaher [2]. A DPCA processor is a fixed space-time processor, and the conditions for DPCA clutter cancellation are closely related to Brennan's rule for the clutter rank. Further analysis of the conditions under which STAP and DPCA are equivalent is given by Richardson [43].

Assume an array aligned with the platform velocity vector. Furthermore, assume that the platform velocity is matched to the radar PRF to satisfy the DPCA condition

$$v_a = \frac{df_r}{2} \quad \text{or} \quad \beta = 1. \quad (\text{C.1})$$

The DPCA condition above states that effectively, the array moves one interelement spacing per pulse repetition interval (see Figure 8). The basic DPCA idea is to displace the receive beam phase center on a pulse-by-pulse basis to compensate for the platform motion. Outputs from different beams with the same effective phase center are subtracted to suppress the clutter much like an MTI canceler. The key to effective clutter cancellation is that the receive beams from different subapertures of the array have exactly the same pattern so that cancellation is achieved at all angles.

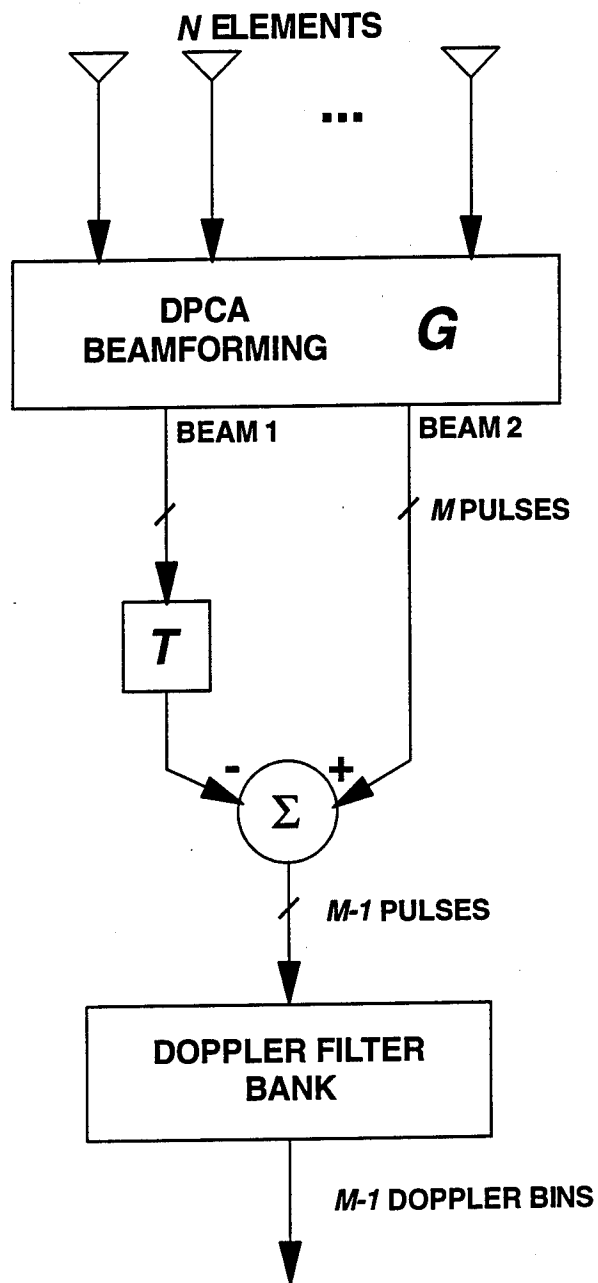


Figure C-1. Block diagram for DPCA processing.

A block diagram of a DPCA processor is shown in Figure C-1. Two receive beams are formed by applying a beamformer matrix

$$\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2] \quad (\text{C.2})$$

to each pulse of data. Typically, the receive phase center of \mathbf{g}_1 is shifted to the left of the array physical center, while that of \mathbf{g}_2 is shifted to the right. Beam 1 is then delayed by one pulse-repetition interval and subtracted from beam 2. This operation, called DPCA cancellation, is a two-beam, two-pulse space-time filter. This process is repeated for each two-pulse sub-CPI, and the resulting $M' = M - 1$ outputs are then coherently integrated in the length M' doppler filter bank. Because the DPCA canceler occurs after initial beamforming of the data and doppler processing occurs after DPCA cancellation, DPCA is a beamspace pre-doppler space-time (fixed, nonadaptive) processor.

To examine the angle doppler response of a DPCA processor, focus on a single two-pulse sub-CPI. Let χ denote the $2N \times 1$ sub-CPI space-time snapshot, from which

$$\tilde{\chi} = (\mathbf{I}_2 \otimes \mathbf{G})^H \chi \quad (\text{C.3})$$

is the 2×2 beamspace sub-CPI snapshot, as in Section 6.2. The DPCA canceler may be represented by the fixed weight vector

$$\tilde{\mathbf{w}} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}. \quad (\text{C.4})$$

The DPCA canceller output is then

$$z = \tilde{\mathbf{w}}^H \tilde{\chi} = [(\mathbf{I}_2 \otimes \mathbf{G}) \tilde{\mathbf{w}}]^H \chi = \mathbf{w}^H \chi, \quad (\text{C.5})$$

where the composite sub-CPI weight vector is

$$\mathbf{w} = (\mathbf{I}_2 \otimes \mathbf{G}) \tilde{\mathbf{w}} = \begin{bmatrix} \mathbf{g}_2 \\ -\mathbf{g}_1 \end{bmatrix}. \quad (\text{C.6})$$

The angle-Doppler response of the DPCA processor, through the canceler, may then be written as

$$\begin{aligned} P(\vartheta, \varpi) &= \mathbf{w}^H \mathbf{v}(\vartheta, \varpi) = \mathbf{g}_2^H \mathbf{a}(\vartheta) - e^{j2\pi\varpi} \mathbf{g}_1^H \mathbf{a}(\vartheta) \\ &= F_2(\vartheta) - e^{j2\pi\varpi} F_1(\vartheta), \end{aligned} \quad (\text{C.7})$$

where $F_1(\vartheta) = \mathbf{g}_1^H \mathbf{a}(\vartheta)$ and $F_2(\vartheta) = \mathbf{g}_2^H \mathbf{a}(\vartheta)$ are the patterns of the two beamformers. Ideally, the DPCA beamformers should satisfy

$$\mathbf{g}_1 = \begin{bmatrix} \mathbf{g} \\ 0 \end{bmatrix}, \quad \mathbf{g}_2 = \begin{bmatrix} 0 \\ \mathbf{g} \end{bmatrix}, \quad (\text{C.8})$$

where \mathbf{g} is an $(N - 1)$ -element beamformer. In this case the two beamformers are equivalent to a single beamformer that processes two overlapping subapertures that are displaced by one interelement spacing. Therefore,

$$F_2(\vartheta) = e^{j2\pi\vartheta} F_1(\vartheta). \quad (\text{C.9})$$

The DPCA response may then be found from Equation (C.7) to be

$$P(\vartheta, \varpi) = F_1(\vartheta) (e^{j2\pi\vartheta} - e^{j2\pi\varpi}). \quad (\text{C.10})$$

This is the response of a two-pulse processor; at any angle a single null in doppler frequency exists. The locus of this null is where the bracketed quantity in Equation (C.10) equals zero. From this the DPCA canceler has a null at

$$\varpi_{\text{null}}(\vartheta) = \vartheta. \quad (\text{C.11})$$

This is exactly the equation for the clutter ridge for $\beta = 1$ and no misalignment. Thus, DPCA nulls the clutter ridge completely.

Another way to view DPCA is to note that the beamformer Equation (C.8), satisfies Theorem 5; the resulting clutter subspace has dimension $2 + (2 - 1) = 3$. The DPCA weight vector, Equation (C.4), lies in the subspace orthogonal¹⁵ to the clutter subspace [43]. This is another way of saying that DPCA forms a space-time filter that completely nulls the entire clutter ridge. It can be shown that for two-pulse DPCA, the full CPI composite weight vector is a projection of the space-time steering vector into the subspace orthogonal to the clutter.

One way to derive beamformers that satisfy Equation (C.8) is by a proper combination of sum and difference beams. Let \mathbf{g}_s and \mathbf{g}_d denote sum and difference beamformers, and let their corresponding patterns be denoted $F_s(\vartheta)$ and $F_d(\vartheta)$. If the sum and difference patterns satisfy

$$\frac{F_d(\vartheta)}{F_s(\vartheta)} = \tan \pi \vartheta , \quad (\text{C.12})$$

then a set of DPCA beamformers can be formed by applying the 2×2 matrix transform

$$\mathbf{T} = \begin{bmatrix} 1 & 1 \\ -j & j \end{bmatrix} \quad (\text{C.13})$$

as follows:

$$\mathbf{G} = [\mathbf{g}_s, \mathbf{g}_d] \mathbf{T} = [\mathbf{g}_s - j\mathbf{g}_d, \mathbf{g}_s + j\mathbf{g}_d] . \quad (\text{C.14})$$

To see this, consider the corresponding patterns

$$\begin{aligned} F_1(\vartheta) &= F_s(\vartheta) - jF_d(\vartheta) = F_s(\vartheta)(1 - j \tan \pi \vartheta) \\ F_2(\vartheta) &= F_s(\vartheta) + jF_d(\vartheta) = F_s(\vartheta)(1 + j \tan \pi \vartheta) . \end{aligned} \quad (\text{C.15})$$

¹⁵In this case, the dimension of the beamspace sub-CPI snapshot is 4, the clutter subspace is of dimension 3, so its orthogonal subspace has dimension 1.

Equations (C.15) and (C.12) are combined to yield

$$F_2(\vartheta) = F_1(\vartheta) \left(\frac{1 + j \tan \pi \vartheta}{1 - j \tan \pi \vartheta} \right) = F_1(\vartheta) e^{j2\pi \vartheta} , \quad (\text{C.16})$$

which is the desired result. Because a sum pattern utilizes a symmetric taper and the difference pattern is antisymmetric, the matrix combiner, Equation (C.13), results in two beamformers whose phase centers are displaced to the left and right of the array physical center.

The sum and difference-beam form of DPCA may also be modified to provide some means for maintaining mainlobe clutter cancellation in mismatched conditions where the DPCA condition does not hold or velocity misalignment exists. Redefine the transform \mathbf{T} above as

$$\mathbf{T}(\kappa) = \begin{bmatrix} 1 & 1 \\ -j\kappa & j\kappa \end{bmatrix} , \quad (\text{C.17})$$

where κ is a gain control that adjusts the displacement between the two DPCA beamformers. With TACCAR adjustment of the clutter doppler and linear control of κ with velocity and the cosine of the misalignment angle, it can be shown that the DPCA filter can be made to match the slope of the clutter ridge at the mainbeam look direction. Therefore, even in mismatched conditions, the two-beam, two-pulse DPCA can provide some level of clutter cancellation.

A full-CPI analysis of DPCA follows in the same way as for beam-space pre-doppler STAP but with the sub-CPI weight vectors given by Equation (C.4) rather than by Equation (234).

The DPCA processor, being a fixed space-time filter, requires precise antenna and system calibration to provide high levels of clutter cancellation. In practice, several factors cause the performance of DPCA to degrade. Among these are imperfect array calibration, element pattern mismatches, velocity-PRF combinations that do not satisfy the DPCA condition, and velocity misalignment. Even with appropriate control as suggested above, DPCA performance in practice can fall far short of the ideal case. Moreover, DPCA has no inherent capability to suppress jamming beyond what the nominal sidelobe level provides. Whereas DPCA seeks to provide complete clutter cancellation, STAP optimizes SINR. Because it is data-adaptive, STAP does not *require* the above-mentioned conditions to achieve clutter cancellation. As a result, STAP is both more flexible and more robust against clutter and can suppress both jamming and clutter simultaneously.

REFERENCES

1. F.R. Dickey, Jr., M. Labitt, and F.M. Staudaher, "Development of airborne moving target radar for long range surveillance," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 27, November 1991, pp. 959-971.
2. F.M. Staudaher, "Airborne MTI," Chapter 16 of *Radar Handbook*, editor M.I. Skolnik, McGraw-Hill, 1990.
3. L.E. Brennan and I.S. Reed, "Theory of adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-9, March 1973, pp. 237-252.
4. L.E. Brennan, J.D. Mallett, and I.S. Reed, "Adaptive arrays in airborne MTI radar," *IEEE Transactions on Antennas and Propagation*, Vol. AP-24, Sept. 1976, pp. 607-615.
5. R. Klemm, "Adaptive clutter suppression for airborne Phased Array Radars," *IEE Proceedings*, Vol. 130, Pts. F and H, No. 1, February, 1983, pp. 125-131.
6. R. Klemm, "Adaptive airborne MTI: an auxiliary channel approach," *IEE Proceedings*, Vol. 134, Pt. F, No. 3, June 1987, pp. 269-276.
7. J. Ender and R. Klemm, "Airborne MTI via digital filtering," *IEE Proceedings*, Vol. 136, Pt. F, No. 1, February 1989, pp. 22-29.
8. R. Klemm, "Adaptive airborne MTI with two-dimensional motion compensation," *IEEE Proceedings-F*, Vol. 138, No. 6, December 1991, pp. 551-558.
9. E.C. Barile, R.L. Fante, and J.A. Torres, "Some limitations on the effectiveness of airborne adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-28, October 1992, pp. 1015-1032.
10. H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 30, no. 3, July 1994, pp. 660-669.
11. H. Wang, H.R. Park, and M.C. Wicks, "Recent results in space-time processing," *Proceedings of the 1994 National Radar Conference*, Atlanta, GA, pp. 104-109.
12. R. DiPietro, "Extended factored space-time processing for airborne radar systems," *Proceedings of the 26th Asilomar Conference on Signals, Systems, and Computing*, Pacific Grove, CA, October, 1992, pp 425-430.
13. R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.
14. J.V. DiFranco and W.L. Rubin, *Radar Detection*, Dedham, MA: Artech House, 1980.
15. W. Rihaczek, *Principles of High-Resolution Radar*, New York: McGraw-Hill Book Company, 1969.

REFERENCES

(Continued)

16. L.V. Blake, *Radar Range-Performance Analysis*, D.C. Heath and Company, Lexington, MA, 1980.
17. F.E. Nathanson, *Radar Design Principles, Second Edition*, McGraw-Hill, Inc., 1991.
18. L.E. Brennan and F.M. Staudaher, "Subclutter Visibility Demonstration," Technical Report RL-TR-92-21, Adaptive Sensors Incorporated, March, 1992.
19. T.A. Barton and D.R. Fuhrmann, "Covariance structures for multidimensional data," *Proceedings of the 35th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November, 1991.
20. F.A. Graybill, *Matrices with Applications in Statistics*, Belmont, CA: Wadworth, Inc., 1983.
21. E.W. Swokowski, *Calculus with Analytic Geometry*, 2nd Edition, Boston, MA: Prindle, Weber, and Schmidt, 1979.
22. R.A. Monzingo and T.W. Miller, *Introduction to Adaptive Arrays*, New York, NY: John Wiley and Sons, 1980.
23. R.T. Compton, Jr., *Adaptive Antennas — Concepts and Performance*, Englewood Cliffs, NJ: Prentice Hall, 1980.
24. I.S. Reed, J.D. Mallett, and L.E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, AES-10, no. 6, November, 1974.
25. S.Z. Kalson, "An adaptive array detector with mismatched signal rejection," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 28, No. 1, January 1992, pp. 195-207.
26. L. Cai and H. Wang, "Performance comparisons of modified SMI and GLR algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 27, No. 3, May 1991, pp. 487-491.
27. L.W. Brooks and I.S. Reed, "Equivalence of the likelihood ratio processor, the maximum signal-to-noise ratio filter, and the Wiener filter," *IEEE Transactions on Aerospace and Electronic Systems*, September, 1972, pp. 690-692.
28. F.C. Robey, D.R. Fuhrmann, E.J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, AES-28, no. 1, pp. 208-216.
29. E.J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, AES-22, no. 1, March 1986, pp. 115-127.
30. E.J. Kelly, E.J., "Performance of an adaptive detection algorithm; rejection of unwanted signals," *IEEE Transactions on Aerospace and Electronic Systems*, AES-25, no. 2, March 1989, pp. 122-133.

REFERENCES

(Continued)

31. D.M. Boroson, "Sample size considerations in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, AES-16, No. 4, July 1980, pp. 446-451.
32. D.D. Feldman and L.J. Griffiths, "A Projection approach for robust adaptive beamforming," *IEEE Transactions on Signal Processing*, Vol. 42, No. 4, April, 1994, pp. 867-876.
33. B.D. Carlson, L.M. Goodman, et al., "An ultralow sidelobe adaptive array antenna," *The Lincoln Laboratory Journal*, Vol. 3, No. 2, Summer 1990, pp. 291-310.
34. S.I. Krich, Private Communication.
35. A. Steinhardt, "Foundations of adaptive detection," *Proceedings of the Adaptive Sensor Array Processing Workshop*, Lincoln Laboratory, Massachusetts Institute of Technology, March 11-12, 1993, pp. 807-831.
36. E.J. Baranoski, "Sparse network array processing," *Proceedings of the Seventh SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, June 26-29, 1994.
37. D.F. Marshall, "A two step adaptive interference nulling algorithm for use with airborne sensor arrays," *Proceedings of the Seventh SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, June 26-29, 1994.
38. J. Ward, "Multiwindow post-Doppler space-time adaptive processing," *Proceedings of the Seventh SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, June 26-29, 1994.
39. M. Wicks and H. Wang, "Adaptive array technology for clutter rejection in airborne radar," *Proceedings of the 1993 National Radar Conference*, March 1993.
40. R.L. Fante, "Cancellation of specular and diffuse jammer multipath using a hybrid adaptive array," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 27, No. 5, September 1991, pp. 823-837.
41. S.T. Smith, private communication, February 1, 1994.
42. L.E. Brennan and J.D. Mallett, "Angle estimation and scattered jamming cancellation in STAP radars," Technical Report J114F, Adaptive Sensors, Inc., August 8, 1992.
43. P. Richardson, "Relationships between DPCA and adaptive space-time processing techniques for clutter suppression," *Proceedings of the International Conference on Radar*, Paris, 1994, pp.295-300.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 13 December 1994	3. REPORT TYPE AND DATES COVERED Technical Report	
4. TITLE AND SUBTITLE Space-Time Adaptive Processing for Airborne Radar			5. FUNDING NUMBERS C — F19628-95-C-0002 PR — 681 PE — 63226E	
6. AUTHOR(S) James Ward				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lincoln Laboratory, MIT P.O. Box 73 Lexington, MA 02173-9108			8. PERFORMING ORGANIZATION REPORT NUMBER TR-1015	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) ARPA/CEXEC 4001 No. Fairfax Dr. Arlington, VA 22203-1714			10. SPONSORING/MONITORING AGENCY REPORT NUMBER ESC-TR-94-109	
11. SUPPLEMENTARY NOTES None				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Future airborne radars will be required to detect targets in an interference background comprised of clutter and jamming. Space-time adaptive processing (STAP) refers to multidimensional adaptive filtering algorithms that simultaneously combine the signals from the elements of an array antenna and the multiple pulses of a coherent radar waveform, to suppress interference and provide target detection. STAP can improve detection of low-velocity targets obscured by mainlobe clutter, detection of targets masked by sidelobe clutter, and detection in combined clutter and jamming environments. This report analyzes a variety of approaches to STAP problem. Optimum, or fully adaptive processing is reviewed. Computational complexity and the need to estimate the interference from a limited amount of available data make fully adaptive STAP impractical. As a result, partially adaptive space-time processors are required. A taxonomy of reduced-dimension STAP algorithms is presented where algorithms are classified based on the type of preprocessor employed. For example, beamspace algorithms use spatial preprocessing, while post-Doppler approaches perform temporal (Doppler) filtering before adaptive processing. In some cases, the special structure of the clutter can be exploited to design preprocessors yielding minimum clutter tank. For each class, either sample-matrix-inversion (SMO) or subspace-based weight computation may be employed. Simulation results are presented to illustrate various performance metrics, including SINR, adapted patterns, minimum detectable velocity, and required degrees of freedom.				
14. SUBJECT TERMS Airborne Radar Space-time filtering Adaptive Array Processing Clutter-interface suppression			15. NUMBER OF PAGES 222	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Same as Report	