

12

Space-Time Adaptive Processing for Radar

William L. Melvin

Georgia Institute of Technology, Atlanta, GA, USA

2.12.1 Introduction

Space-time adaptive processing (STAP) is a two-dimensional, adaptive filtering technique foundational to modern radar system design and implementation. Generally, spatial sampling is given by the outputs of an array of antenna elements (sometimes called “channels”); the Fourier transform of space yields angle. There are essentially two time measurements in radar: fast-time and slow-time. Fast-time samples correspond to outputs of the analog-to-digital converter abiding with the Nyquist rate for the reflected radar waveform; the Fourier transform of fast-time is radio-frequency (RF), which relates to radar range. The processor collects slow-time voltage samples at a given range realization from distinct, reflected, pulsed transmissions; the Fourier transform of slow-time yields Doppler frequency. Depending on the statistical properties of the interference source, either slow-time or fast-time is applicable. For example, the adaptive suppression of ground clutter received by airborne or spaceborne radar requires joint spatial and slow-time adaptivity since clutter returns appear correlated in the angle-Doppler domain, but are uncorrelated in fast-time. Fast-time adaptive degrees of freedom play in role in wideband and terrain-bounce interference mitigation in both aerospace and surface-based radar systems.

Suppose we consider adaptive mitigation of Doppler-spread ground clutter returns in further detail. This problem arises in airborne and spaceborne radar systems searching for moving targets. The clutter returns are generally much stronger than the target signal, in many cases by orders of magnitude, thereby rendering detection in certain regions of the range-Doppler map virtually impossible using conventional antenna beamforming and Doppler processing techniques. Spatial and Doppler beamforming represent one-dimensional matched filtering operations. The matched filter maximizes signal-to-noise ratio (SNR), thereby maximizing probability of detection for a fixed false alarm rate for uncorrelated, circular Gaussian disturbance [1, 2]. STAP is needed when clutter and radio frequency interference (RFI)—colored noise—are present; the STAP applies a space-time weighting incorporating estimated characteristics of the interference environment to asymptotically maximize signal-to-interference-plus-noise ratio (SINR), a sufficient condition to maximize the probability of detection for a fixed false alarm rate under the assumption the disturbance is circular Gaussian [3].

Fixed points on the Earth’s surface exhibit a Doppler frequency shift consistent with their angular displacement from the platform velocity vector, v_p . In turn, this angular displacement corresponds to a specific angle of arrival at the antenna array. Thus, given a particular azimuth and elevation angle in the antenna coordinate system, the clutter Doppler frequency is precisely specified, and vice versa.

The clutter angle-Doppler region of support is described as a ridge in airborne and spaceborne radar. This ridge is a line in sidelooking radar, opens up into an ellipse for varying degrees of platform yaw, and becomes a circle for the forward-looking array. Figure 12.1 provides example power spectral density (PSD) and minimum variance distortionless response (MVDR) super-resolution plots of the clutter angle-Doppler region of support for the side-looking array radar (SLAR), a forward-looking array radar (FLAR), and an array with forty-five (45°) of yaw. Typical radar parameters are used in this example and will be described in further detail in subsequent sections.

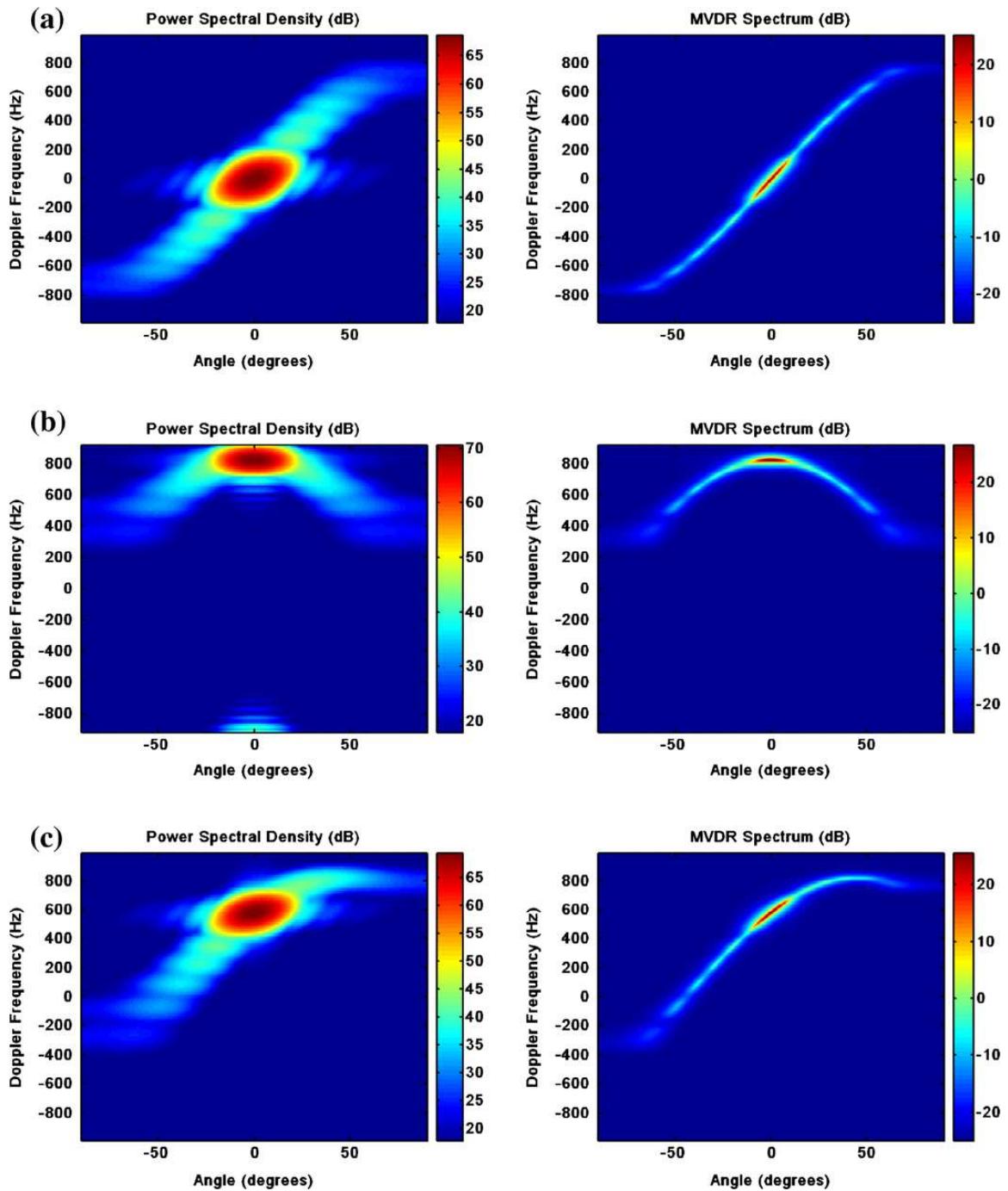
The PSD plots are a result of classical, two-dimensional, Fourier analysis. Resolution is limited in Fourier analysis by roughly the wavelength of the signal divided by the aperture length; to resolve more closely spaced signals, either the wavelength must be decreased for a fixed aperture size, or the physical aperture must be increased for a fixed wavelength. Super-resolution spectral estimators overcome the diffraction limitations of Fourier-based methods, oftentimes providing resolution enhancement by a factor of two to five, or even more. It turns out that STAP is intimately related to the MVDR spectrum, as discussed in Section 2.12.2.2. Hence, STAP likewise exhibits super-resolution properties, enabling it to detect slow moving targets buried in mainbeam clutter or targets closely spaced to other interference sources.

As mentioned, STAP coherently combines spatial and temporal voltage samples from a particular range realization using a data-dependent weighting to maximize SINR. The corresponding adaptive filter frequency response exhibits sharp nulls along the clutter region of support while maximizing the gain at another, specified angle-Doppler location where a potential target might exist. A target with a Doppler frequency outside of the mainlobe clutter Doppler spread (given as the highest power, circular region in the PSD) is called *exoclutter*, whilst targets within the mainlobe clutter spread are *endoclutter*. The STAP frequency response null is the inverse of the super-resolution contour shown in Figure 12.1. Hence, STAP itself exhibits super-resolution properties, providing detection performance within the diffraction limits of the traditional space-time aperture. This important characteristic makes STAP invaluable on airborne and spaceborne platforms, enabling drastic reduction in antenna aperture length for a desired minimum detectable velocity (MDV). The asymptotic STAP frequency responses for each of the PSDs in Figure 12.1 are given in Figure 12.2. (The asymptotic STAP response is optimal, since this is the weighting that precisely maximizes SINR.) The filter is tuned to broadside and 400 Hz Doppler.

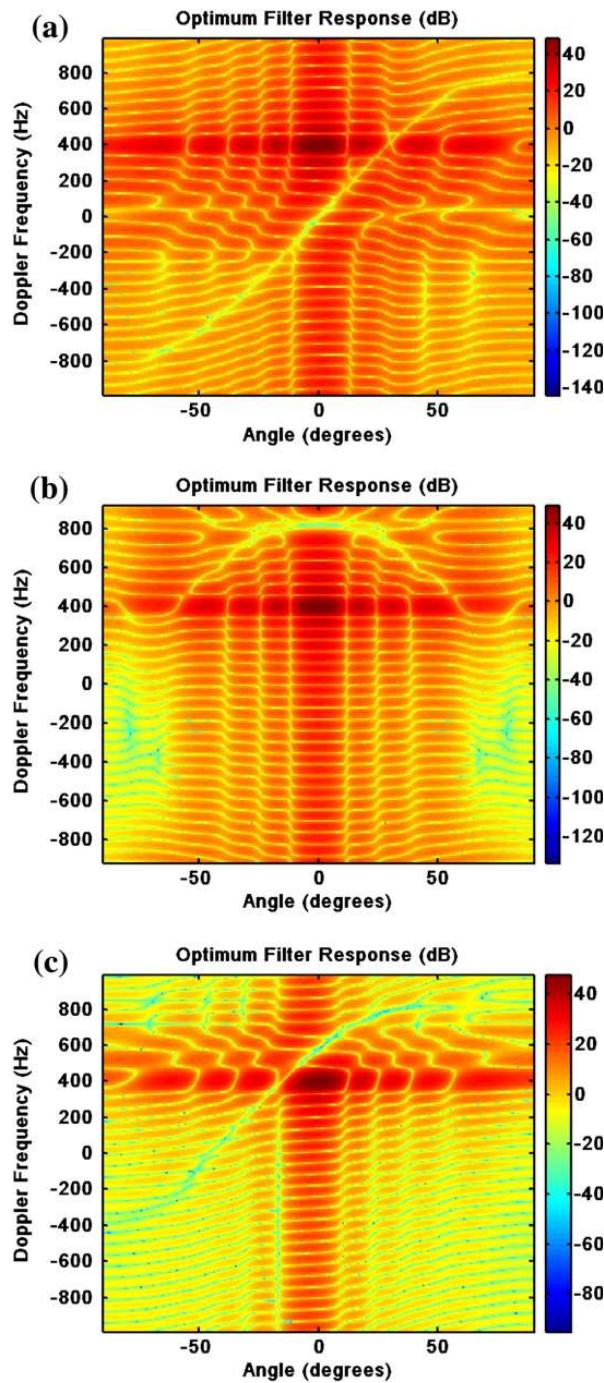
The MDV is the lowest target range-rate where sufficient SINR exists to permit a specified probability of detection for an acceptable probability of false alarm. Conversely, the MDV occurs at those points where the loss in SINR due to clutter is maximally acceptable. Figure 12.3 provides an example of the loss in SINR due to clutter for the radar array configuration of Figures 12.1 and 12.2. SINR loss will be described further in Section 2.12.2. Noise-limited performance is at 0 dB. The depth of the loss null is related to the clutter-to-noise ratio (CNR).

2.12.1.1 Historical overview

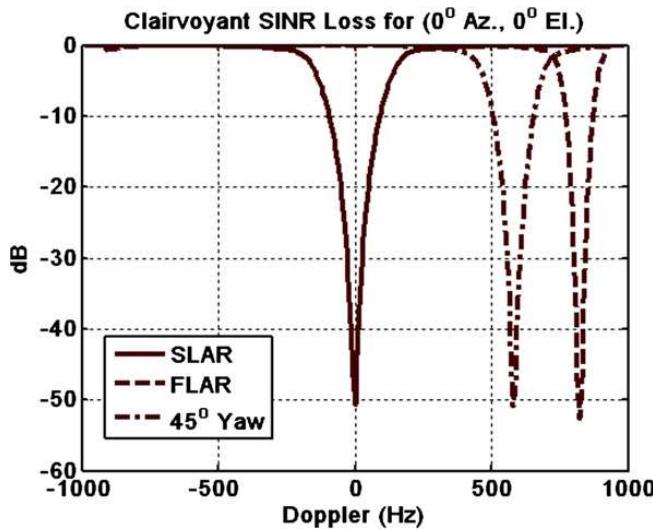
Historical overviews of STAP are given in [4,5]. Reed [4] describes the earlier history of adaptive arrays, whilst Klemm [5] covers more contemporary developments in STAP. Additionally, Entzminger et al. [6] discusses the history and future of Joint STARS and ground moving target indication (GMTI), identifying the central role of STAP in the success of both.

**FIGURE 12.1**

Clutter angle-Doppler region of support for three array configurations. (a) PSD (left) and super-resolution clutter image (right) for side-looking array, (b) PSD (left) and super-resolution clutter image (right) for forward-looking array, and (c) PSD (left) and super-resolution clutter image (right) for array with 45° of yaw.

**FIGURE 12.2**

Comparison of STAP frequency response for the three scenarios in Figure 12.1. (a) Side-looking array, (b) forward-looking array, and (c) array with 45° of yaw.

**FIGURE 12.3**

Clairvoyant SINR loss for example radar scenarios.

Table 12.1 is an historical accounting of STAP from the author’s perspective, as well as information derived from [4–6]. As such, it attempts to convey some key developments and general trends, but is necessarily incomplete. (Note: references to specific works are provided throughout the remainder of this text. Additionally, Klemm [5] provides an excellent, historical accounting and perspective, and is highly recommended to the interested reader.)

2.12.1.2 Organization

We organize the remaining discussion as follows:

- Section 2.12.2 describes basic concepts, include target detection; different space-time optimal filter formulations, including the maximum SINR filter, the minimum variance beamformer, and the generalized sidelobe canceler; the sample matrix inversion (SMI) approach to adaptive filtering; and, key performance metrics, such as SINR loss.
- Section 2.12.3 discusses clutter, RFI, target, and receiver space-time signal models.
- Section 2.12.4 discusses different adaptive filter topologies, including several reduced-rank methods, reduced-dimension STAP, and parametric adaptive matched filtering.
- Section 2.12.5 describes the utility of STAP to clutter suppression in airborne radar, as an example. This section also includes benchmark results for some of the algorithms discussed in Section 2.12.4.
- Much of current STAP research focuses on extending textbook discussion to real-world application. Section 2.12.6 describes various STAP challenges and research areas, including the impact of heterogeneous or nonstationary clutter on detection performance and mitigating techniques, application to airborne bistatic and conformal array radar, and knowledge-aided STAP. Each of these challenge areas predominantly focuses on issues surrounding effective estimation of the interference covariance matrix.

Table 12.1 Some Significant Events in the History of STAP

Year	Topic
1957	Paul Howells of General Electric, Syracuse (now Lockheed-Martin), develops technique to electronically scan antenna null in direction of a jammer
1959	Howells receives US Patent, "Intermediate frequency side-lobe canceller"
1962	Howells and Sidney Applebaum successfully test five-loop side-lobe canceller
1963	Howells and Applebaum, at Syracuse University Research Corp. (then Syracuse Research Corp., now SRC) investigate application of adaptive techniques to radar, including OTHR and BMD radar
1965 (1976)	Applebaum publishes "Adaptive Arrays," SURC TR-66-001, later available in IEEE Transactions on Antennas and Propagation in 1976
1969	Lloyd Griffiths publishes an adaptive algorithm for wideband antennas in Proceedings of the IEEE
1969	V. Anderson, H. Cox and N. Owsley separately publish works on adaptive arrays for sonar
1971	R.T. Compton describes the application of adaptive arrays for communications systems in Ohio State University Quarterly Rept. 3234-1, December 1971
1972	O.L. Frost publishes an adaptive algorithm for antenna arrays incorporating constraints
1972	Howells and Applebaum investigate adaptive radar for Airborne Early Warning (AEW) radar
1973	Larry Brennan and Irving Reed publish the seminal paper, "Theory of adaptive radar," in IEEE Trans. AES
1974	Reed, John Mallett and Brennan publish a paper describing the Sample Matrix Inverse (SMI) method for adaptive arrays in IEEE Transactions on Aerospace and Electronic Systems
1976	"Adaptive antenna systems," published by Bernard Widrow et al. in IEEE Proceedings
1980	R.A. Monzingo and T.W. Miller publish the book <i>Adaptive Arrays</i>
1983	Rule for calculating clutter subspace dimension proposed by Richard Klemm in IEE Proceedings Part F, Vol 130, February 1983
1991	Joint STARS prototypes deploy to Gulf War using adaptive clutter suppression methods (1978, Pave Mover was pre-cursor)
1992	Klemm proposes spatial transform techniques for dimensionality reduction in paper on antenna design for STAP
1992	Real-time STAP implementation by Alfonso Farina et al.
1992	Bob DiPietro presents post-Doppler STAP algorithm at Asilomar Conference on Signals, Systems and Computers
1994	Post-Doppler, beamspace method proposed by Wang and Cai in IEEE Transactions on Aerospace and Electronic Systems
1994	Jim Ward of MIT Lincoln Lab describes STAP techniques in ESC-TR-94-109, <i>Space-Time Adaptive Processing for Airborne Radar</i>
1995	Air Force Rome Laboratory (now Air Force Research Laboratory) and Westinghouse (now Northrop Grumman) collect twenty-four channel airborne radar data for STAP research and development under the Multichannel Airborne Radar Measurements (MCARM) program

Continued

Table 12.1 *Continued*

Year	Topic
1998	Richard Klemm publishes first STAP text book, <i>Space-Time Adaptive Processing: Principles and Applications</i>
1999	IEE Electronics and Communications Engineering Journal (ECEJ) Special Issue on STAP (Klemm, Ed.)
1999	STAP techniques for space-based radar, by Rabideau and Kogon, presented at IEEE Radar Conference
1999	3-D STAP for hot and cold clutter mitigation appears in IEEE Trans. AES by Techau, Guerci, Slocumb and Griffiths
2000	Bistatic STAP techniques appear in literature (numerous authors)
2000	IEEE Transactions on Aerospace and Electronic Systems Special Section on STAP (William Melvin, Ed.)
2002	DARPA's Knowledge-Aided Sensor Signal Processing & Expert Reasoning (KASSPER) commences under leadership of Joseph Guerci
2003	Guerci publishes the book <i>Space-Time Adaptive Processing for Radar</i>
2004	IEE publishers release the book <i>The Applications of Space-Time Processing</i> (Klemm, Ed.)
c. 2005	Multi-Input, Multi-Output (MIMO) adaptive systems (numerous authors)
2006	IEEE Transactions on Aerospace and Electronic Systems, Special Section on Knowledge-Aided Signal and Data Processing (Melvin, Guerci, Ed.)
2009	Cognitive Radar (Guerci)

- A description of some additional, practical issues in STAP implementation is given in Section 2.12.7, including maximum likelihood angle estimation and adaptive matched filter normalization.
- In Section 2.12.8 we briefly discuss some multichannel radar data collection efforts.
- Summary comments are given in Section 2.12.9, along with notation and acronym lists in the appendices.

2.12.1.3 Key points

STAP is a key component of modern radar. Some important points are given below:

- STAP is really a catch-all phrase for a variety of weight calculation schemes and adaptive architectures.
- Radar terminology usually includes two time measurements: fast-time, corresponding to a sample rate consistent with the analog-to-digital converter clock speed and effectively measuring range; and, slow-time, where the pulse repetition interval (PRI) defines the sample rate and separating targets and clutter in Doppler frequency is the primary motivation. Space and slow-time are used to mitigate ground clutter returns in aerospace radar, whereas space and fast-time are used to cope with wideband RFI or multipath signals in aerospace or surface radar systems.
- Herein, we focus on space and slow-time adaptive processing for clutter suppression. The ideas discussed are easily extended to other radar degrees of freedom (DoFs), include fast-time, polarization, and multiple passes.

- Implementation of the optimal processor requires precise knowledge of the null-hypothesis covariance matrix and the target space-time steering vector. The adaptive filter is an approximation to the optimal processor, since neither covariance matrix nor steering vector are known in practice. There are two primary factors leading to differences between adaptive and optimal filter outputs: covariance matrix estimation error and unknown array manifold due to amplitude and phase error as a function of frequency. Adaptive SINR loss describes the impact of both of these errors on detection performance potential.
- Current STAP research focuses on coupling basic STAP theory, physics-based models, and the realities of the real-world operating environment to construct practical detection architectures.
- STAP is applied at the coherent processing interval (CPI) level, often typically prior to noncoherent—or post-detection—integration. Maximizing SINR is always effective at improving detection performance, since the separation between the null and alternative hypothesis output probability density functions increases.
- In many respects, STAP development is in its early stages, at least from an implementation perspective. It has been only recently that embedded computing power has improved to the point where more flexible, real-time STAP implementation is feasible.

2.12.2 Basic concepts

In this section we discuss radar space-time detection and several approaches to calculate the optimal space-time weight vector: the maximum SINR filter, the minimum variance beamformer, the generalized sidelobe canceler, and use of the Rayleigh quotient and generalized eigen-analysis. We introduce the sample matrix inverse adaptive filter as the most common of the various adaptive implementations used in radar application. We also discuss several important STAP metrics, such as clairvoyant and adaptive SINR loss and improvement factor.

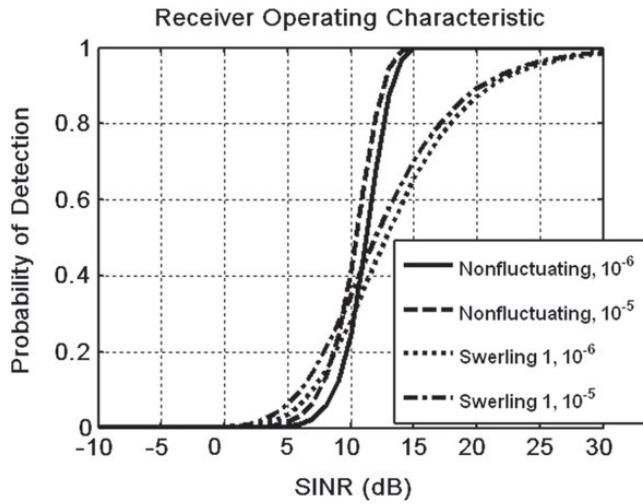
2.12.2.1 Detection

Suppose we collect M spatial samples across the antenna array for range realization k and each of the N received pulses. These voltages are organized in a vector as

$$\mathbf{x}_k = \left[\mathbf{x}_{s/k}^T(1) \quad \mathbf{x}_{s/k}^T(2) \quad \cdots \quad \mathbf{x}_{s/k}^T(N) \right]^T, \quad (12.1)$$

where $\mathbf{x}_k \in C^{NM \times 1}$ is the space-time snapshot and $\mathbf{x}_{s/k}(n) \in C^{M \times 1}$ is the spatial snapshot for the n th pulse. Two hypotheses correspond to (12.1): the null hypothesis, H_0 , where only additive interference and noise are present; and, the alternative hypothesis, H_1 , which includes target presence in addition to the components of the null hypothesis. Defining $\mathbf{c}_k \in C^{NM \times 1}$ as the clutter space-time snapshot, $\mathbf{j}_k \in C^{NM \times 1}$ as the RFI space-time snapshot, $\mathbf{n}_k \in C^{NM \times 1}$ as the uncorrelated noise space-time snapshot, and $\mathbf{t}_k \in C^{NM \times 1}$ the target space-time snapshot, we have

$$\begin{aligned} H_0 : \mathbf{x}_k &= \mathbf{c}_k + \mathbf{j}_k + \mathbf{n}_k, \\ H_1 : \mathbf{x}_k &= \mathbf{c}_k + \mathbf{j}_k + \mathbf{n}_k + \mathbf{t}_k = \mathbf{x}_{k/H_0} + \mathbf{t}_k. \end{aligned} \quad (12.2)$$

**FIGURE 12.4**

Receiver operating characteristic (ROC) curves for nonfluctuating and Swerling 1 targets (probability of false alarm given in legend).

The radar detection problem is to decide, given \mathbf{x}_k , which of the two hypotheses is most probable. It is shown in [3] that maximizing SINR is equivalent to maximizing probability of detection for a fixed false alarm rate when \mathbf{x}_k/H_0 is jointly complex normal with zero mean and covariance matrix, \mathbf{R}_k , viz. $\mathbf{x}_k/H_0 \sim CN(\mathbf{0}, \mathbf{R}_k)$. We wish to combine the elements of \mathbf{x}_k using a linear, finite impulse response (FIR) filter, and compare the scalar to a threshold, η_T , to decide which hypothesis is a best fit,

$$y_k = \mathbf{w}_k^H \mathbf{x}_k = \sum_{m=1}^{NM} [\mathbf{w}_k]^*_m [\mathbf{x}_k]^*_m \begin{matrix} H_1 \\ > \\ H_0 \end{matrix} \quad (12.3)$$

where \mathbf{w}_k is the space-time weight vector.

Figure 12.4 shows probability of detection, P_d , as a function of output SINR for varying probability of false alarm, P_{fa} . The curves correspond to nonfluctuating and Swerling 1 target types for $P_{fa} = 10^{-5}$ and $P_{fa} = 10^{-6}$. As seen, P_d increases monotonically with SINR. Also, it is seen that for reasonably high P_d we generally need higher SINR to detect a Swerling 1 target at the same rate as the nonfluctuating target due to fading; the SINR difference between Swerling 1 and nonfluctuating detection curves is called fluctuation loss.

2.12.2.2 Space-time filter formulations

2.12.2.2.1 Maximum SINR filter

A primary objective in STAP is to approximate as closely as possible the optimal weighting that maximizes SINR. By attempting to maximize output SINR, the STAP maximizes the probability of

detection for a fixed false alarm rate, as discussed in the prior section. We now formulate the optimal weight vector.

The target snapshot is of the form, $\mathbf{t}_k = \alpha_T \mathbf{s}_{s-t}(\phi, \theta, f_d)$, where α_T is a complex gain term proportional to the square root of the target cross section, $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the space-time steering vector of dimension NM -by-one, ϕ is azimuth, θ is elevation, and f_d is Doppler frequency. In the nonfluctuating target case, α_T is a constant, whereas in the Swerling 1 case, $\alpha_T \sim \mathcal{CN}(0, \sigma_T^2)$, where σ_T^2 is the target signal variance [1, 7, 8]. The space-time steering vector is the Kronecker product of the spatial steering vector, $\mathbf{s}_s(\phi, \theta) \in \mathbb{C}^{M \times 1}$, and the temporal steering vector, $\mathbf{s}_t(f_d) \in \mathbb{C}^{N \times 1}$. Generally, the spatial steering vector describes the phase variation among the spatial antenna channels (of common size and gain) to a signal with a direction of arrival of (ϕ, θ) , whereas the temporal steering vector describes the phase variation over the slow-time aperture to a signal with a particular range-rate [9–11]. If the aperture in either space or time is uniformly sampled, the corresponding steering vector will be Vandermonde.

The output SINR of the FIR filter with weight vector, \mathbf{w}_k , is the ratio of target power, P_T , to interference-plus-noise power, P_{I+N} , given as

$$\text{SINR} = \frac{P_T}{P_{I+N}} = \frac{E[y_T y_T^*]}{E[y_k/H_0 y_k^*/H_0]} = \frac{E[\mathbf{w}_k^H \mathbf{t}_k \mathbf{t}_k^H \mathbf{w}_k]}{E[\mathbf{w}_k^H \mathbf{x}_k/H_0 \mathbf{x}_k^H/H_0 \mathbf{w}_k]} = \frac{\mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{R}_{k/H_0} \mathbf{w}_k}. \quad (12.4)$$

Using the definition for \mathbf{t}_k gives, $\mathbf{R}_T = E[\mathbf{t}_k \mathbf{t}_k^H] = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, and the Schwarz Inequality lets us write (12.4) as

$$\text{SINR} = \sigma_T^2 \frac{|\tilde{\mathbf{w}}^H \tilde{\mathbf{s}}|^2}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}} \leq \sigma_T^2 \frac{(\tilde{\mathbf{w}}^H \tilde{\mathbf{w}})(\tilde{\mathbf{s}}^H \tilde{\mathbf{s}})}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}}. \quad (12.5)$$

In this case, $\tilde{\mathbf{w}} = \mathbf{R}_{k/H_0}^{-1/2} \mathbf{w}_k$ and $\tilde{\mathbf{s}} = \mathbf{R}_{k/H_0}^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$. By inspection, we find that the right side of the inequality achieves the upper bound when $\tilde{\mathbf{w}} = \mu \tilde{\mathbf{s}}$, thereby providing the solution for the optimal weight vector. Through substitution, it follows that

$$\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.6)$$

for arbitrary scalar, μ , is the weighting that yields maximal output SINR. When the disturbance is uncorrelated noise, $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$, where σ_n^2 is the noise power; then, (12.6) becomes the space-time matched filter, viz. $\mathbf{w}_k = \mu \mathbf{s}_{s-t}(\phi, \theta, f_d)$.

Inserting the optimal weight vector, $\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, into (12.4) yields the maximum output SINR,

$$\text{SINR}_{\text{Max}} = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.7)$$

For simplicity, and without loss in meaning, in subsequent sections we use $\mathbf{R}_k = \mathbf{R}_{k/H_0}$.

The maximum SINR weight vector has the following interpretation. If $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$, then $\tilde{\mathbf{x}}_k = \mathbf{R}_k^{-1/2} \mathbf{x}_k$ is the whitened data vector, where $\tilde{\mathbf{x}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{NM})$ and \mathbf{I}_{NM} is the NM -by- NM identity matrix. This results because

$$E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^H] = \mathbf{R}_k^{-1/2} E[\mathbf{x}_k \mathbf{x}_k^H] \mathbf{R}_k^{-1/2} = \mathbf{R}_k^{-1/2} \mathbf{R}_k \mathbf{R}_k^{-1/2} = \mathbf{I}_{NM}. \quad (12.8)$$

Then, the filter employing the maximum SINR weight vector is seen as a cascade of whitening and warped matched filtering operations,

$$y_k = \mathbf{w}_k^H \mathbf{x}_k = \left(\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \right) \left(\mathbf{R}_k^{-1/2} \mathbf{x}_k \right) = \tilde{\mathbf{s}}_{s-t}^H \tilde{\mathbf{x}}_k, \quad (12.9)$$

where $\tilde{\mathbf{s}}_{s-t} = \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the warped matched filter accounting for the linear transformation of the target signal through the whitening process and $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the original matched filter weight vector.

2.12.2.2 Minimum variance beamformer

The minimum variance (MV) space-time beam former is another common formulation leading to the adaptive processor [12, 13]. The MV beamformer employs a weight vector yielding minimal output power subject to a linear constraint on the desired target spatial and temporal response:

$$\min_{\mathbf{w}_k} E[y_k y_k^*] \text{ subject to } \mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = g, \quad (12.10)$$

where g is a complex scalar. We can determine the weight vector satisfying the problem statement in (12.10) using the method of Lagrange. Thus, define the following cost function

$$J = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k + \operatorname{Re} \left\{ \tilde{\lambda}^* \left(\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) - g \right) \right\}, \quad (12.11)$$

where $\tilde{\lambda}^*$ is a complex Lagrangian multiplier and $E[y_k y_k^*] = E[\mathbf{w}_k^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{w}_k] = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k$. Taking the gradient of (12.11) with respect to the conjugate of the weights and setting to zero gives

$$\nabla J = \mathbf{R}_k \mathbf{w}_k + \tilde{\lambda}^* \mathbf{s}_{s-t}(\phi, \theta, f_d) = 0. \quad (12.12)$$

Solving (12.12) for the weight vector yields

$$\mathbf{w}_k = -\tilde{\lambda}^* \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.13)$$

Next, we find by applying the constraint in (12.10) that

$$\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = -\tilde{\lambda} \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = g. \quad (12.14)$$

Solving (12.14) for $\tilde{\lambda}$ then gives

$$\tilde{\lambda} = \frac{-g}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.15)$$

The minimum variance weight vector follows from (12.13) and (12.15) as

$$\mathbf{w}_k = \frac{g^* \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.16)$$

Equation (12.16) abides by the form $\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, where

$$\mu = \frac{g^*}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.17)$$

Hence, (12.16) likewise yields maximal SINR.

Setting $g = 1$ is known as the *distortionless response*, and the corresponding weight vector yields the minimum variance distortionless response (MVDR) beamformer. The output power of the MVDR beamformer is given as

$$P_{\text{MVDR}} = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k = \frac{1}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.18)$$

The MVDR spectrum, as given in Figure 12.1, follows from (12.18) by scanning the space-time steering vector in the denominator over the angle and Doppler locations of interest.

2.12.2.2.3 Generalized sidelobe canceler

The GSLC is a formulation that conveniently converts the minimum variance constrained beamformer described in the preceding section into an unconstrained form [12,14]. Many prefer the GSLC interpretation of STAP over the whitening filter-warped matched filter interpretation of Section 2.12.2.2.1; since the GSLC implements the MV beamformer for the single linear constraint [14], this structure likewise can be shown to maximize output SINR.

Figure 12.5 provides a block diagram of the GSLC. The top leg of the GSLC generates a quiescent response by forming a beam at the angle and Doppler of interest. A blocking matrix, $\mathbf{B}_G = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$, prevents the target signal from entering the lower leg; essentially, the blocking matrix forms a notch filter tuned to (ϕ, θ, f_d) . With the desired signal absent from the lower leg, the processor tunes the weight vector to provide a minimal mean square error (MMSE) estimate of the interference in the top leg. In the final step, the processor differences the desired signal, $d_{o/k}$, with the estimate from the lower leg, $\hat{d}_{o/k}$, to form the filter output, $y_k = d_{o/k} - \hat{d}_{o/k}$. Ideally, any large residual at the output corresponds to an actual target.

The desired signal is given as

$$d_{o/k} = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{x}_k. \quad (12.19)$$

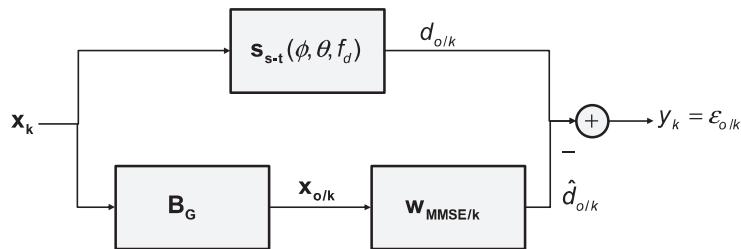


FIGURE 12.5

GSLC block diagram.

The signal vector in the lower leg is

$$\mathbf{x}_{0/k} = \mathbf{B}_G \mathbf{x}_k; \quad \mathbf{B}_G \in C^{(NM-1) \times NM}; \quad \mathbf{x}_{0/k} \in C^{(NM-1) \times 1}. \quad (12.20)$$

Forming the quiescent response of (12.19) uses a single degree of freedom (DoF), resulting in the reduced dimensionality of $\mathbf{x}_{0/k}$. By weighting the signal vector in the lower leg, the GSLC forms a MMSE estimate of $d_{o/k}$ as

$$\hat{d}_{o/k} = \mathbf{w}_{\text{MMSE}/k}^H \mathbf{x}_{0/k}, \quad (12.21)$$

where the MMSE weight vector follows from the well-known Wiener-Hopf equation [12, 13],

$$\mathbf{w}_{\text{MMSE}/k} = \mathbf{R}_{x_{0/k}}^{-1} \mathbf{r}_{x_{0/k} d}. \quad (12.22)$$

The lower leg covariance matrix is

$$\mathbf{R}_{x_{0/k}} = E \left[\mathbf{x}_{0/k} \mathbf{x}_{0/k}^H \right] = \mathbf{B}_G \mathbf{R}_k \mathbf{B}_G^H \quad (12.23)$$

whilst the cross-correlation between lower and upper legs is

$$\mathbf{r}_{x_{0/k} d} = E \left[\mathbf{x}_{0/k} d_{o/k}^* \right] = \mathbf{B}_G \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.24)$$

The GSLC filter output is then

$$\begin{aligned} y_k &= \varepsilon_{o/k} = d_{o/k} - \hat{d}_{o/k} = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{x}_k - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{x}_{0/k} \\ &= \left(\mathbf{s}_{s-t}^H(\phi, \theta, f_d) - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \right) \mathbf{x}_k. \end{aligned} \quad (12.25)$$

Comparing (12.25) and (12.3), we identify

$$\mathbf{w}_{\text{GSLC}/k} = \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k}. \quad (12.26)$$

We compute the output SINR as the ratio of output signal power to interference-plus-noise power, as in (12.4). The signal-only output power of the GSLC is

$$\begin{aligned} P_T &= E \left[\mathbf{w}_{\text{GSLC}/k}^H \mathbf{t}_k \mathbf{t}_k^H \mathbf{w}_{\text{GSLC}/k} \right] = E \left[\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{t}_k \mathbf{t}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right] \\ &= \sigma_T^2 \left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2 = \sigma_T^2 (NM)^2, \end{aligned} \quad (12.27)$$

and the output interference-plus-noise power is

$$P_{I+N} = E \left[\mathbf{w}_{\text{GSLC}/k}^H \mathbf{x}_k / \mathbf{H}_0 \mathbf{x}_k^H / \mathbf{H}_0 \mathbf{w}_{\text{GSLC}/k} \right] = \mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k}. \quad (12.28)$$

The ratio of (12.27) and (12.28) is

$$\text{SINR}_{\text{GSLC}} = \frac{\sigma_T^2 N^2 M^2}{\mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k}} = \frac{\sigma_T^2 N^2 M^2}{\sigma_d^2 - \mathbf{r}_{x_{0/k} d}^H \mathbf{R}_{x_{0/k}}^{-1} \mathbf{r}_{x_{0/k} d}}, \quad (12.29)$$

where $\sigma_d^2 = E[|d_{o/k}|^2] = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)$. We arrive at the denominator in (12.29) by using (12.22)–(12.24),

$$\begin{aligned} \mathbf{w}_{\text{GSLC}/k}^H \mathbf{R}_k \mathbf{w}_{\text{GSLC}/k} &= \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k} \\ &\quad - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d) + \mathbf{w}_{\text{MMSE}/k}^H \mathbf{B}_G \mathbf{R}_k \mathbf{B}_G^H \mathbf{w}_{\text{MMSE}/k} \\ &= \sigma_d^2 - \mathbf{r}_{x_0/k}^H \mathbf{w}_{\text{MMSE}/k} - \mathbf{w}_{\text{MMSE}/k}^H \mathbf{r}_{x_0/k} + \mathbf{w}_{\text{MMSE}/k}^H \mathbf{R}_{x_0/k} \mathbf{w}_{\text{MMSE}/k} \\ &= \sigma_d^2 - \mathbf{r}_{x_0/k}^H \mathbf{R}_{x_0/k}^{-1} \mathbf{r}_{x_0/k}. \end{aligned} \quad (12.30)$$

2.12.2.2.4 Rayleigh quotient

The optimal weight vector in Section 2.12.2.2.1 maximizes SINR. As seen from (12.4), for weight vector \mathbf{w}_k , SINR is given as the ratio of quadratics, $\text{SINR} = \mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k / \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k$. Thus, our problem is to find the weight vector that maximizes this function. In accord with [15], and similar to our discussion in Section 2.12.2.2.1, SINR is written as

$$\text{SINR} = \frac{\tilde{\mathbf{w}}^H \left(\left(\mathbf{R}_k^{-1/2} \right)^H \mathbf{R}_T \mathbf{R}_k^{-1/2} \right) \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}}, \quad (12.31)$$

where $\tilde{\mathbf{w}} = \mathbf{R}_k^{1/2} \mathbf{w}_k$. The weight vector, \mathbf{w}^0 , that maximizes this Rayleigh quotient form is known to be proportional to the eigenvector \mathbf{v}_{\max} associated with the largest eigenvalue λ_{\max} of the matrix in the numerator [13],

$$\tilde{\mathbf{w}} \propto \mathbf{v}_{\max} = \max_{\text{Eigenvector}} \left(\left(\mathbf{R}_k^{-1/2} \right)^H \mathbf{R}_T \mathbf{R}_k^{-1/2} \right) = \max_{\text{Eigenvector}} \left(\mathbf{R}_k^{-1/2} \mathbf{R}_T \mathbf{R}_k^{-1/2} \right), \quad (12.32)$$

and so

$$\mathbf{w}^0 \propto \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \max_{\text{Eigenvector}} \left(\mathbf{R}_k^{-1/2} \mathbf{R}_T \mathbf{R}_k^{-1/2} \right). \quad (12.33)$$

Take the case where $\mathbf{R}_T = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, as previously introduced in Section 2.12.2.2.1. The maximum eigenvector in (12.32) must abide by

$$\sigma_T^2 \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \lambda_{\max} \mathbf{v}_{\max}. \quad (12.34)$$

The solution for \mathbf{v}_{\max} in (12.34) is thus seen to be

$$\mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.35)$$

and so, as expected,

$$\mathbf{w}^0 \propto \mathbf{R}_k^{-1/2} \mathbf{v}_{\max} = \mathbf{R}_k^{-1/2} \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.36)$$

(Note: inserting (12.35) in (12.34) gives $\mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d)$ times a scalar on each side of the equation, thereby satisfying the eigen-relation.)

By plugging $\mathbf{w}_k = \mathbf{w}^0$ found in (12.33) into (12.4), it is straightforward to show that the maximum SINR is equal to λ_{\max} . In this vein, using (12.35) in (12.34) gives

$$\lambda_{\max} = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1/2} \mathbf{R}_k^{-1/2} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \sigma_T^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.37)$$

which is the desired result and matches the expression given in (12.7).

2.12.2.2.5 Generalized eigen-analysis

The second approach, described in [13], begins by noting that the solution to maximizing SINR is equivalent to finding the maximum eigenvalue and associated eigenvector for the generalized Eigen-problem

$$\mathbf{R}_T \mathbf{w}_k = \lambda \mathbf{R}_k \mathbf{w}_k. \quad (12.38)$$

This can be solved as an ordinary eigen-equation by writing it as

$$(\mathbf{R}_k^{-1} \mathbf{R}_T) \mathbf{w}_k = \lambda \mathbf{w}_k. \quad (12.39)$$

Then, the weight vector that maximizes (12.4) is proportional to the eigenvector \mathbf{v}_{\max} associated with the largest eigenvalue λ_{\max} of $\mathbf{R}_k^{-1} \mathbf{R}_T$,

$$\mathbf{w}^0 \propto \mathbf{v}_{\max} = \max_{\text{Eigenvector}} (\mathbf{R}_k^{-1} \mathbf{R}_T). \quad (12.40)$$

The maximum eigenvalue of $\mathbf{R}_k^{-1} \mathbf{R}_T$ is the maximum achievable SINR in (12.4).

If we again consider the case where $\mathbf{R}_T = \sigma_T^2 \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)$, the generalized eigen-analysis equation of (12.39) becomes

$$(\sigma_T^2 \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{s}_{s-t}^H(\phi, \theta, f_d)) \mathbf{w}_k = \lambda \mathbf{w}_k, \quad (12.41)$$

which, we find, is satisfied when $\mathbf{w}_k = \mathbf{w}^0 = \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$. Substituting this eigenvector, \mathbf{w}_k , into (12.41) gives,

$$(\sigma_T^2 \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \lambda_{\max} \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.42)$$

from which it is seen the maximum eigenvalue, λ_{\max} , is identical to the result in (12.37) and also the same as the maximum SINR expression found in Section 2.12.2.2.1.

2.12.2.6 Summary

Table 12.2 summarizes the weight vector formulation given in the prior sections.

2.12.2.3 Sample matrix inversion

All of the previously described approaches to filter design presume the availability of the interference-plus-noise covariance matrix for the k th realization, \mathbf{R}_k , and perfectly known target space-time steering

Table 12.2 Summary of Space-Time Filter Formulations

Formulation	Objective	Weight Vector
Maximum SINR filter	Maximize output SINR	$\mathbf{w}_k = \mu \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$
Minimum variance (MV) beamformer	Minimize output power subject to linear constraint maintaining gain in the look direction	$\mathbf{w}_k = \frac{\mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}$
Generalized sidelobe canceler	Convert constrained beamformer formulation to unconstrained form	$\mathbf{w}_k = \mathbf{w}_{GSLC/k} = \mathbf{s}_{s-t}(\phi, \theta, f_d) - \mathbf{B}_G^H \mathbf{w}_{MMSE/k}$
Rayleigh quotient	Maximize output SINR	$\mathbf{w}_k \propto \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$
Generalized eigen-analysis	Maximize output SINR	$\mathbf{w}_k \propto \mathbf{R}_{k/H_0}^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$

vector, $\mathbf{s}_{s-t}(\phi, \theta, f_d)$. In practice, neither \mathbf{R}_k nor $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ is known. The disturbance covariance matrix, \mathbf{R}_k , changes as look-angle, platform location, and platform attitude change. Errors in $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ are generally a result of straddling the precise target spatial or temporal frequency when searching over angle and Doppler, as well as system hardware errors that most greatly affect the spatial component, $\mathbf{s}_s(\phi, \theta)$. These system errors generally manifest as spatially varying, complex gain errors due to factors such as spatial variation in element gain patterns, varying line lengths, radome and airframe interactions, etc.

A number of techniques are available to adapt the weight vector, including least means square (LMS) and recursive least squares (RLS) formulations [12]. However, the sample matrix inversion (SMI) method is by far the most popular choice for radar application for two primary reasons: (1) convergence of the SMI approach depends only on the number of samples used to estimate the unknown interference-plus-noise covariance matrix, regardless of environmental conditions as long as the data are independent and identically distributed (IID), for a system of specified DoFs; and, (2) prior to processing, radar systems generally collect their data in blocks, called coherent processing intervals (CPIs), and so a batch processing strategy using SMI is fully acceptable. In the SMI approach, the optimal weight vector,

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.43)$$

is simply replaced with the adaptive weight vector,

$$\hat{\mathbf{w}}_k = \hat{\mu} \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d), \quad (12.44)$$

where $\hat{\mu}$ is a scalar that often depends on estimated quantities, $\hat{\mathbf{R}}_k$ is the interference-plus-noise covariance matrix estimate, and $\mathbf{v}_{s-t}(\phi, \theta, f_d)$ is the hypothesized space-time steering vector. In the absence of straddle error, system errors dominate the error between $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{v}_{s-t}(\phi, \theta, f_d)$. Generally, array errors are explicitly modeled as

$$\mathbf{s}_{s-t}(\phi, \theta, f_d) = \mathbf{s}_t(f_d) \otimes (\mathbf{e}_s(\phi, \theta) \odot \mathbf{s}_s(\phi, \theta)), \quad (12.45)$$

where $\boldsymbol{\varepsilon}_s(\phi, \theta)$ is the spatially-varying error between the ideal and actual array manifold, and \otimes and \odot are Kronecker (tensor) and Hadamard (element-wise) products, respectively.

In practice, $\boldsymbol{\varepsilon}_s(\phi, \theta)$ is “calibrated out” of the system using a variety of techniques, including array tuning on an antenna range or *in situ* using the clutter background and navigation data to estimate differences between ideal and actual multi-channel antenna responses [16]. Correspondingly, the hypothesized space-time steering vector is then

$$\mathbf{v}_{s,t}(\phi, \theta, f_d) = \mathbf{v}_t(f_d) \otimes (\hat{\boldsymbol{\varepsilon}}_s(\phi, \theta) \odot \mathbf{v}_s(\phi, \theta)), \quad (12.46)$$

where $\hat{\boldsymbol{\varepsilon}}_s(\phi, \theta) \in C^{M \times 1}$ is an estimate of the array error vector, and $\mathbf{v}_t(f_d)$ and $\mathbf{v}_s(\phi, \theta, f_d)$ are otherwise the hypothesized temporal and spatial steering vectors accounting for potential straddle. Temporal errors due to system non-ideality are commonly very small for the typical STAP CPI, and thus usually are not considered further.

The complex baseband, pulse compressed, space-time snapshots comprise the voltage vectors for the p th CPI,

$$\mathbf{X}_{\text{CPI}}(p) = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_L] \in C^{NM \times L}. \quad (12.47)$$

A particular realization, \mathbf{x}_k , chosen from amongst the space-time vectors in (12.47) for filtering and detection thresholding is called the cell-under-test (CUT); multiple CUTs form the primary data set. The remaining vectors in (12.47) are available to estimate the unknown interference-plus-noise covariance matrix and are called training or secondary data. The primary data set can be as small as a single CUT with several adjacent realizations serving as guard cells; the purpose of the guard cell region is to prevent any target energy from leaking into the training interval. Let L_g be the number of guard cells on either side of the primary data region. Then, define the training set as

$$\mathbf{X}_{\text{Train}}(p) = \{\mathbf{x}_m\}_{m \in 1:\kappa; m \neq k-L_g:k+L_g}. \quad (12.48)$$

It is shown in [17] that if the data vectors comprising the training interval—the columns of (12.48)—are IID with respect to the null-hypothesis of the CUT, a maximum likelihood estimate for \mathbf{R}_k is

$$\hat{\mathbf{R}}_k = \frac{1}{\kappa} \mathbf{X}_{\text{Train}}(p) \mathbf{X}_{\text{Train}}^H(p) = \frac{1}{\kappa} \sum_{m=1}^{\kappa} \mathbf{x}_m \mathbf{x}_m^H. \quad (12.49)$$

A fundamental question centers on how many training samples, κ , lead to an acceptable covariance matrix estimate. This matter was addressed in [17] and considered in further detail in Section 2.12.2.4.2.

2.12.2.4 Metrics

In radar, performance is generally measured according to the specific goals of the collection and processing mode. In moving target indication (MTI) radar, probability of detection, P_d , and false alarm rate (FAR)—or probability of false alarm, P_{fa} —are the primary measures of performance. As discussed in Section 2.12.2.1, the probability of detection is a monotonic function of output SINR for a fixed false alarm rate. Thus, measures of SINR are key to understanding achievable system performance. Among these measures, SINR loss metrics are preferable, since they characterize performance over certain independent variables—usually range rate or Doppler frequency, as shown in Figure 12.3—and are not

tied to a specific radar cross section, thus folding into traditional link budget analysis. Specifically, given SNR calculated from the radar range equation [1, 7, 8] or measured in the field, the output SINR is

$$\text{SINR}(\phi, \theta, f_d) = \text{SNR}(\phi, \theta) \cdot \prod_{m=1}^Q L_{s,m}(\phi, \theta, f_d), \quad (12.50)$$

where $L_{s,m}(\phi, \theta, f_d)$ is the m th SINR loss term and $0 \leq L_{s,m}(\phi, \theta, f_d) \leq 1$. (N.b., the term “SINR loss” is widely used, even though such losses are negative valued on a decibel scale. This oftentimes contradicts typical radar systems engineering usage. However, the losses are numerator terms in the full link calculation.) In general, SINR losses vary with angle and Doppler, since the interference PSD likewise varies.

In imaging radar, terrain-to-noise ratio (TNR) is a primary metric and is a function of multiplicative noise ratio and SINR loss. We focus our attention specifically on MTI radar, but the basic metrics are adaptable.

2.12.2.4.1 *Clairvoyant SINR loss*

Clairvoyant SINR loss is the ratio of output SINR for a filter implementation, \mathbf{w}_k , where all parameters are known precisely, to the maximum SINR. Clairvoyant SINR is given as

$$L_{s,1} = \frac{\text{SINR}}{\text{SNR}} = \frac{\left(\frac{\sigma_T^2 |\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right)}{\left(\frac{\sigma_T^2}{\sigma_n^2} NM \right)} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}. \quad (12.51)$$

A few cases are worth considering:

Case 1—Noise-Limited Condition: In this case, $\mathbf{R}_k = \sigma_n^2 \mathbf{I}_{NM}$, where σ_n^2 is the noise power. Then, the maximum SINR weight vector,

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \rightarrow \mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.52)$$

and so

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\sigma_n^2 \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)} = 1. \quad (12.53)$$

As expected, the optimal weighting defaults to the matched filter and there is no loss relative to the bound defined by the uncorrelated noise.

Case 2—Color-Limited Condition, Matched Filter Weights: Clairvoyant loss characterizes the impact of interference on detection performance. Consider the case where the weight vector is set to the space-time matched filter, $\mathbf{w}_k = \mathbf{s}_{s-t}(\phi, \theta, f_d)$. The clairvoyant SINR loss follows from (12.51) as

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{|\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{s}_{s-t}(\phi, \theta, f_d)|^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)} = \frac{\sigma_n^2 NM}{\text{PSD}(\phi, \theta, f_d)} \leq 1, \quad (12.54)$$

where $\text{PSD}(\phi, \theta, f_d) = \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the power spectral density (PSD), equal to the two-dimensional Fourier transform of the space-time covariance matrix [12]; Figure 12.1 provides

example clutter-plus-noise PSD plots. Equation (12.54) shows the impact of interference on performance relative to the noise-limited case. Since the PSD is diffraction-limited (the mainlobe is determined by the size of the space-time aperture), (12.54) characterizes the performance impacts of using the deterministic matched filter. At those angles and Doppler frequencies away from the clutter angle-Doppler region of support, the clairvoyant SINR loss approaches the noise-limited case.

Case 3—Color-Limited Condition, Optimal Weights: Using the maximum SINR weight vector, $\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)$, in (12.51), gives

$$L_{s,1} = \frac{\sigma_n^2}{NM} \frac{\left| \mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) \right|^2}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)} = \frac{\sigma_n^2}{NM} P_{MVDR}^{-1}(\phi, \theta, f_d), \quad (12.55)$$

where $P_{MVDR}(\phi, \theta, f_d)$ is a sample of the MVDR spectrum given in (12.18). Considering the MVDR plots in Figure 12.1, (12.55) suggests regions of loss confined to the sharp, super-resolution contours of the clutter MVDR response (note the inverse in (12.55)). For this reason, STAP is able to detect targets in close proximity to the center of the clutter angle-Doppler region of support. In contrast, (12.54) suggests that SINR loss extends to the full width of the diffraction-limited spectrum when using nonadaptive weights. Figure 12.25, given in Section 2.12.5.3, compares SINR loss for adaptive and nonadaptive solutions, confirming these observations.

2.12.2.4.2 Adaptive SINR loss

Adaptive SINR loss is the ratio of the output SINR for the filter using estimated quantities, $\hat{\mathbf{w}}_k$, to the optimal filter output, \mathbf{w}_k , where all parameters are clairvoyantly known, viz.

$$L_{s,2} = \left(\frac{\hat{\mathbf{w}}_k^H \mathbf{R}_T \hat{\mathbf{w}}_k}{\hat{\mathbf{w}}_k^H \mathbf{R}_k \hat{\mathbf{w}}_k} \right) \Bigg/ \left(\frac{\mathbf{w}_k^H \mathbf{R}_T \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k} \right). \quad (12.56)$$

Observe that when $\hat{\mathbf{w}}_k = \mathbf{w}_k$, $L_{s,2} = 1$. Moreover, $0 \leq L_{s,2} \leq 1$, since \mathbf{w}_k yields the maximum output SINR and so the denominator in (12.56) will always be greater than or equal to the numerator.

Substituting (12.43) and (12.44) into (12.56) yields

$$L_{s,2}(\hat{\mathbf{R}}_k) = (\text{SINR}|\hat{\mathbf{w}}_k(\hat{\mathbf{R}}_k)) / (\text{SINR}_{\text{Max}}), \quad (12.57)$$

where SINR_{Max} is given by (12.7). Assuming $\hat{\mathbf{w}}_k$ uses (12.49) in its calculation, as (12.57) suggests, then determining κ is critical. This important matter was addressed in [17], where it is shown that (12.57) is Beta distributed, with mean value

$$E[L_{s,2}(\hat{\mathbf{R}}_k)] = (\kappa + 2 - NM) / (\kappa + 1). \quad (12.58)$$

Setting (12.58) equal to 0.5 (or, 3 dB of loss) and solving for κ yields

$$\kappa = 2NM - 3. \quad (12.59)$$

It is popular to refer to the result in (12.59), where setting κ roughly equal to twice the processor's DoFs yields 3 dB of loss, as the Reed-Mallett-Brennan (RMB) rule after its originators. In practice, 3 dB of loss is substantial, and so choosing κ to be at least five times the processor's DoF is desirable.

The IID assumption inherent in the calculation of (12.57) sets the bound on performance for the adaptive processor given κ homogeneous training samples. We address the impact of non-IID clutter conditions in further detail in Section 2.12.5.

2.12.2.4.3 Improvement factor

Improvement factor (IF) is given as the ratio of the output SINR to the input (element-level) SINR [18, 19]. IF is given as

$$\text{IF} = \frac{\left(\frac{\sigma_I^2 |w_k^H s_{s-t}(\phi, \theta, f_d)|^2}{w_k^H R_k w_k} \right)}{\left(\frac{\sigma_I^2}{\sigma_I^2 + \sigma_n^2} \right)} = \left| w_k^H s_{s-t}(\phi, \theta, f_d) \right|^2 \frac{(\sigma_I^2 + \sigma_n^2)}{w_k^H R_k w_k}, \quad (12.60)$$

where σ_I^2 is the input (element-level) interference power.

Case 1—Noise-Limited Condition: In the noise-limited case, $R_k = \sigma_n^2 I_{NM}$ and $\sigma_I^2 = 0$, and considering the matched filter, $w_k = s_{s-t}(\phi, \theta, f_d)$, then

$$\text{IF} = \frac{\sigma_n^2 |s_{s-t}^H(\phi, \theta, f_d) s_{s-t}(\phi, \theta, f_d)|^2}{\sigma_n^2 s_{s-t}^H(\phi, \theta, f_d) s_{s-t}(\phi, \theta, f_d)} = NM. \quad (12.61)$$

As expected, the improvement factor equals the space-time integration gain in this instance.

Case 2—Clutter-Limited Condition, Matched Filter Weights: In the clutter-limited case, with $w_k = s_{s-t}(\phi, \theta, f_d)$, the improvement factor is

$$\text{IF} = \left| s_{s-t}^H(\phi, \theta, f_d) s_{s-t}(\phi, \theta, f_d) \right|^2 \frac{\sigma_I^2 + \sigma_n^2}{s_{s-t}^H(\phi, \theta, f_d) R_k s_{s-t}(\phi, \theta, f_d)} = (NM)^2 \frac{\sigma_I^2 + \sigma_n^2}{\text{PSD}(\phi, \theta, f_d)} \leq NM. \quad (12.62)$$

The presence of interference in the PSD degrades the improvement factor. If we let $R_k = \sigma_n^2 I_{NM}$ and $\sigma_I^2 = 0$ in (12.62), we arrive at (12.61).

Case 3—Clutter-Limited Condition, Optimal Weights: Using the maximum SINR weight vector, $w_k = \mu R_k^{-1} s_{s-t}(\phi, \theta, f_d)$, in (12.60), gives

$$\text{IF} = (\sigma_I^2 + \sigma_n^2) \frac{\left| s_{s-t}^H(\phi, \theta, f_d) R_k^{-1} s_{s-t}(\phi, \theta, f_d) \right|^2}{s_{s-t}^H(\phi, \theta, f_d) R_k^{-1} s_{s-t}(\phi, \theta, f_d)} = (\sigma_I^2 + \sigma_n^2) P_{\text{MVDR}}^{-1}(\phi, \theta, f_d). \quad (12.63)$$

This expression indicates reduced improvement for those angle-Doppler locations aligning with the clutter, and otherwise good performance in proximity to the clutter response owing to the super-resolution characteristic of $P_{\text{MVDR}}(\phi, \theta, f_d)$ (see Figure 12.1).

2.12.2.4.4 Optimal and adaptive filter patterns

The filter gain pattern follows directly by evaluating the filter response over the angles (or spatial frequencies) and Doppler frequencies of interest. Define the space-time steering matrix,

$$S = [s_{s-t}(\phi(1), \theta(1), f_d(1)) \ s_{s-t}(\phi(2), \theta(1), f_d(1)) \ \cdots \ s_{s-t}(\phi(P), \theta(U), f_d(V))], \quad (12.64)$$

for all $\{\phi(1), \phi(2), \dots, \phi(P)\}$, $\{\theta(1), \theta(2), \dots, \theta(U)\}$, and $\{f_d(1), f_d(2), \dots, f_d(V)\}$. The optimal gain pattern is

$$\mathbf{g}_o = 20 \log_{10} \left| \mathbf{w}_k^H \mathbf{S} \right|, \quad (12.65)$$

where \mathbf{w}_k is the optimal weighting given by (12.43). The adaptive gain pattern follows similarly as

$$\mathbf{g}_a = 20 \log_{10} \left| \hat{\mathbf{w}}_k^H \mathbf{S} \right| \quad (12.66)$$

with (12.44) yielding $\hat{\mathbf{w}}_k$.

Example gain patterns corresponding to the clutter environments shown in Figure 12.1 are given in Figure 12.2. This gain pattern corresponds to (12.65).

2.12.3 Signal models

In this section we consider basic space-time signal and covariance models.

2.12.3.1 Clutter

A plethora of reflected signals from the Earth's surface comprise what is known as radar ground clutter. Ground clutter is a primary impediment to the detection of moving targets. The clutter signal is predominantly made up of signal reflections from distributed objects, such as returns from soil, forests, fields, etc. Discrete clutter sources are less frequently occurring and spatially distributed with a particular density as a function of RCS, leading to strong, point-like returns. Since clutter discrete returns occur relatively infrequently, their suppression is more challenging and they tend to drive false alarm rate. Many discrete returns result from manmade objects, like buildings, water towers, utility poles, etc. Fence lines, train tracks, and power lines are examples of extended sources of discrete clutter.

We discuss basic distributed and discrete clutter models.

2.12.3.1.1 Distributed clutter

Distributed clutter returns result from the integrated response of all scatterers within the range resolution cell. A model for distributed clutter involves discretizing the range resolution cell into fine angle bins called clutter patches. The location of the angle bin relative to the platform velocity vector determines the clutter patch Doppler frequency. The clutter snapshot then follows as the sum of the returns from each of the individual patches.

Figure 12.6 depicts the discretized clutter patch model for distributed clutter. The antenna gain varies around the range resolution annulus on the Earth's surface in accordance with the array normal and steering direction. It is common to model the clutter complex voltage as a Gaussian random variable due to the constructive and destructive sum of the returns from the sub-scatterers comprising each resolution cell. The patch clutter-to-noise ratio (CNR) at the channel-level (assuming matched channels) is

$$\text{CNR}_k(p) = \left(\frac{P_t G_t(\phi_p, \theta_p)}{4\pi r^2} \right) \left(\frac{\sigma_{k,p}}{4\pi r^2} \right) \left(\frac{g_c(\phi_p, \theta_p) \lambda^2}{4\pi k_B T_o B_r F_n L_{rf}} \right) G_{sp}, \quad (12.67)$$

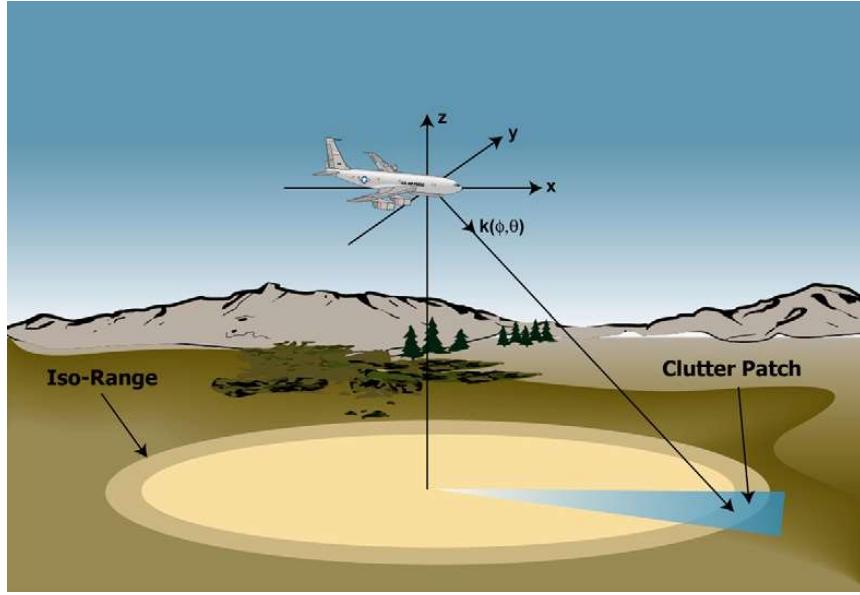
**FIGURE 12.6**

Illustration of space-time clutter patch calculation (after [9], © 2003 IEEE).

where P_t is the transmit power, $G_t(\phi_p, \theta_p)$ is the transmit gain in the p th direction of interest, r is the slant range, $\sigma_{k,p}$ is the p th clutter patch radar cross section, $g_c(\phi_p, \theta_p)$ is the receive channel gain, λ is the center wavelength, k_B is Boltzman's constant, T_o is standard operating temperature, B_r is the receiver bandwidth, F_n is noise figure, L_{rf} is RF system loss, and G_{sp} is signal processing gain (the pulse compression ratio, in this case) [1, 7, 8]. The clutter RCS is a function of the patch reflectivity, $\sigma_{k,p}^0$, and area, A_p , where the area is determined to be a fraction of the resolution cell,

$$\sigma_{k,p} = \sigma_{k,p}^0 A_p. \quad (12.68)$$

The constant gamma model is a common choice used to model reflectivity and is given as $\sigma_{k,p}^0 = \gamma_{k,p} \sin(\psi_p)$, where $\gamma_{k,p}$ is a normalized reflectivity term dependent on the terrain type and ψ_p is the grazing angle to the clutter patch [20]. The clutter patch power then follows by multiplying (12.67) by the receiver noise, σ_n^2 . The clutter voltage for the k th range bin and p th patch voltage is then

$$v_k(p) = \varsigma_{k,p} \sigma_n \sqrt{\text{CNR}_k(p)}, \quad (12.69)$$

where $\varsigma \sim \text{CN}(0, 1)$ [9–11, 18, 19]. (If the response is non-fluctuating, as in the case of a clutter discrete subsequently discussed, then ς is a complex scalar with unity magnitude and uniformly distributed phase.)

The patch voltage in (12.69) varies over the space-time aperture: there is a phase change from channel-to-channel due to direction of arrival, described by the spatial steering vector; and, a phase change from pulse-to-pulse due to the change in range-rate between the clutter patch and the radar platform, described by the temporal steering vector. The clutter space-time snapshot then follows by summing the voltages over the P patches for each of Q range ambiguities, where $q = 0$ is the first

(unambiguous) range return

$$\mathbf{c}_k = \sum_{q=0}^Q \sum_{p=1}^P v_k(p, q) \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)). \quad (12.70)$$

It is common to assume each patch is statistically independent, in which case the clutter covariance matrix corresponding to (12.70) is

$$\mathbf{R}_{c/k} = \sum_{q=0}^Q \sum_{p=1}^P \sigma_{c,k}^2(p, q) \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)) \mathbf{s}_{s-t}^H(\phi_p(q), \theta_p(q), f_{d,p}(q)), \quad (12.71)$$

where $\sigma_{c,k}^2(p, q) = E[|v_k(p, q)|^2]$ is the clutter patch power and follows from (12.69). Equations (12.70) and (12.71) form the basic ground clutter models.

The clutter voltage decorrelates, mainly in time due to intrinsic clutter motion (ICM). Windswept vegetation and moving water are two cases where ICM results. The two basic models describing clutter temporal decorrelation include the Billingsley model [21] and the Gaussian model [7, 10]. The Billingsley model is most common for overland surveillance; it allows a certain amount of the clutter power to decorrelate, say due to leaves fluctuating in the breeze, whilst some of the clutter power is persistent, thus modeling the tree trunk, for example. The Billingsley model leads to an exponential autocorrelation function. In contrast, the Gaussian power spectrum of [7] leads to a Gaussian autocorrelation; the Gaussian model leads to complete decorrelation over a specified time interval and is most appropriate in marine or riverine environments. Dispersion (group delay) across the array is the common source of spatial decorrelation; a sinc autocorrelation model is commonly chosen to model this effect, since it is the inverse transform of a rectangular function in the frequency domain [22, 23].

Given the aforementioned discussion, define the length- NM space-time correlation taper as $\mathbf{a}_{s-t} = \mathbf{a}_t \otimes \mathbf{a}_s$, with correlation matrix $\mathbf{A}_{s-t} = E[\mathbf{a}_{s-t} \mathbf{a}_{s-t}^H] = \mathbf{A}_t \otimes \mathbf{A}_s$. \mathbf{A}_t is the temporal correlation matrix and \mathbf{A}_s is the spatial correlation matrix. The resulting clutter covariance matrix is

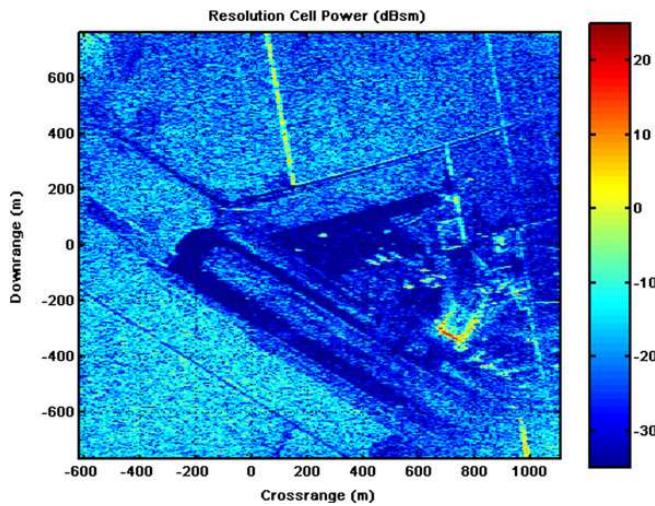
$$\mathbf{R}_{c/k} = \sum_{q=0}^Q \sum_{p=1}^P \sigma_{c,k}^2(p, q) \mathbf{A}_{s-t,k}(p, q) \odot \mathbf{s}_{s-t}(\phi_p(q), \theta_p(q), f_{d,p}(q)) \mathbf{s}_{s-t}^H(\phi_p(q), \theta_p(q), f_{d,p}(q)). \quad (12.72)$$

Generating space-time clutter snapshots exhibiting the correlation described by \mathbf{A}_{s-t} typically involves shaping white noise by multiplying the matrix square root of \mathbf{A}_t and \mathbf{A}_s by random vectors of length N or M , where the elements of each vector are uncorrelated, zero mean, unity variance Gaussian variates.

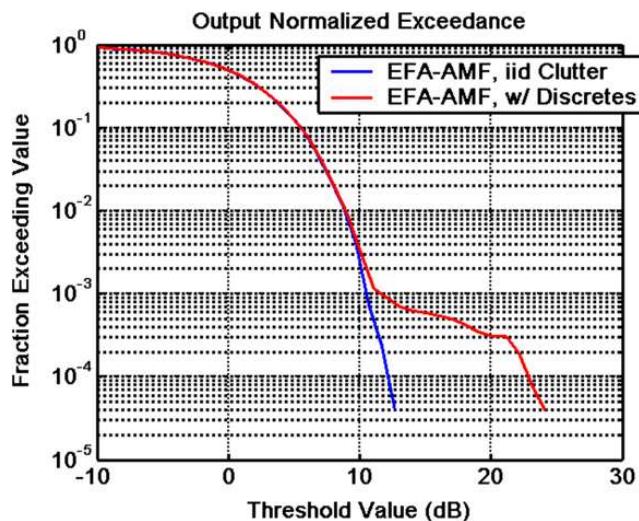
2.12.3.1.2 Discrete clutter

Clutter discretes are strong returns that occur relatively infrequently and correspond to stationary, point-like objects, such as parked vehicles, utility poles, water towers, etc. Generally, the RCS of the clutter discrete is large relative to the RCS of a typical, resolved clutter patch. In some cases, the discrete clutter can appear extended, such as in the case of a fence line or train track.

Figure 12.7 shows an example of discrete-like returns in a spotlight synthetic aperture radar (SAR) image of the Mojave Desert Airport. The hangar edges and fence lines, as well as some aircraft on the tarmac, are evident in the figure. Figure 12.8 shows an exceedance plot characterizing the typical

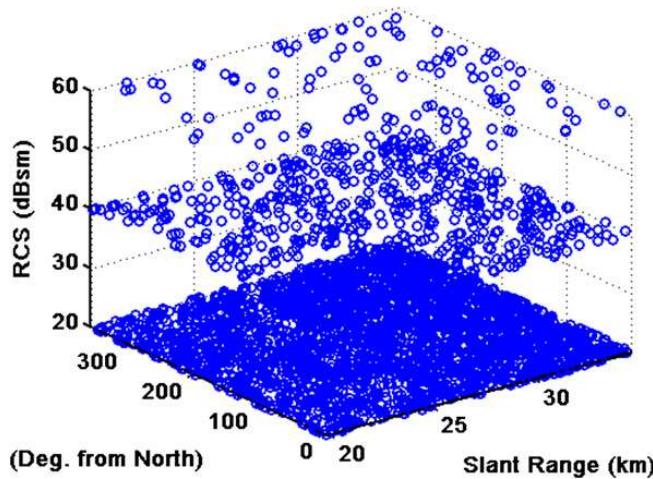
**FIGURE 12.7**

Spotlight SAR image of Mojave Desert Airport at 1 m resolution ([24], © 2004 IEEE).

**FIGURE 12.8**

Example exceedance plot showing the impact of discrete clutter on STAP performance.

impact of clutter discretes on detection performance. (EFA is a post-Doppler STAP method described in Section 2.12.4 and AMF refers to a STAP normalization discussed in Section 2.12.7.) In a homogeneous environment, the output of the STAP is well-behaved and the cumulative distribution follows the expected exponentially-shaped exceedance; in contrast, with discretes present, the tails of the exceedance plot are extended, indicating a requirement to raise the detection threshold to maintain a constant false alarm rate.

**FIGURE 12.9**

Example of clutter discrete seeding.

A plausible discrete model involves laying out the spatial locations of discrete clutter of varying densities. One way to do this is to specify the density per km^2 for a particular discrete RCS and employ a Poisson distribution to identify the random clutter discrete locations. An example of a range-angle seeding of discrete clutter of varying RCS from 20 dBsm to 60 dBsm is shown in Figure 12.9.

Since discrete clutter is stationary, the angular position uniquely specifies Doppler frequency. In other words, clutter discretes reside along the angle-Doppler region of support corresponding to stationary objects. As in the case of distributed clutter, the discrete-to-noise ratio (DNR) follows in a form similar to (12.67) for the l th discrete at range, r_l , and angle (ϕ_l, θ_l) , as

$$\text{DNR}(l) = \left(\frac{P_t G_t(\phi_l, \theta_l)}{4\pi r_l^2} \right) \left(\frac{\sigma_{d,l}}{4\pi r_l^2} \right) \left(\frac{g_c(\phi_l, \theta_l)\lambda^2}{4\pi k_B T_o B_r F_n L_{rf}} \right) G_{\text{sp}}. \quad (12.73)$$

As the discrete is considered point-like, the RCS is taken as nonfluctuating, whereas the phase is uniformly distributed. The corresponding discrete space-time snapshot is

$$\mathbf{c}_d(l; k(l)) = e^{j2\pi u} \sigma_n \sqrt{\text{DNR}(l)} \mathbf{s}_{s-t}(\phi_l, \theta_l, f_{d,l}), \quad (12.74)$$

where u is a uniformly-distributed random variable between $[0, 1]$ and $k(l)$ is the range realization for the l th clutter discrete. The discrete clutter is additive. The corresponding covariance matrix for each discrete is

$$\mathbf{R}_{cd}(l; k(l)) = \sigma_n^2 \text{DNR}(l) \mathbf{s}_{s-t}(\phi_l, \theta_l, f_{d,l}) \mathbf{s}_{s-t}^H(\phi_l, \theta_l, f_{d,l}). \quad (12.75)$$

2.12.3.2 Radio frequency interference

Typically, RFI appears as a noise-like, in-band signal source. The RFI can be intentional or not. In the narrowband case, the k th spatial snapshot for the i th RFI source and n th pulse is

$$\mathbf{j}_{s/k}(n; i) = \omega_k(nT) \mathbf{s}_s(\phi_i, \theta_i), \quad (12.76)$$

where $\omega_k(nT)$ is an uncorrelated source and T is the PRI. The waveform, $\omega_k(nT)$, is generally uncorrelated at nonzero lags, viz.,

$$E[\omega_k(nT)\omega_k^*(mT)] = \sigma_{J,i}^2 \delta((n-m)T) \quad (12.77)$$

but otherwise fully correlated for very short time lags corresponding to propagation across the multi-channel array, as (12.76) indicates (thus, $E[\omega_k(nT)\omega_j^*(nT)] = 0$, for $k \neq j$, in the narrowband case). The corresponding spatial covariance matrix is

$$\mathbf{R}_{\text{s/RFI}}(i) = \sigma_{J,i}^2 \mathbf{s}_s(\phi_i, \theta_i) \mathbf{s}_s^H(\phi_i, \theta_i) \in \mathbb{C}^{M \times M}. \quad (12.78)$$

As a result of (12.77), the space-time covariance matrix is

$$\mathbf{R}_{\text{RFI}}(i) = E \left[\begin{bmatrix} \mathbf{j}_{s/k}(0; i) \\ \vdots \\ \mathbf{j}_{s/k}((N-1); i) \end{bmatrix} \begin{bmatrix} \mathbf{j}_{s/k}^H(0; i) & \cdots & \mathbf{j}_{s/k}^H((N-1); i) \end{bmatrix} \right] = \mathbf{I}_N \otimes \mathbf{R}_{\text{s/RFI}}(i) \in \mathbb{C}^{NM \times NM}. \quad (12.79)$$

Thus, a spatial null is sufficient to remove the narrow-band RFI, as no temporal correlation is present.

Each RFI source is considered statistically independent, so that for J sources, $\mathbf{j}_k = \sum_{i=1}^J \mathbf{j}_k(i)$, where $\mathbf{j}_k \in \mathbb{C}^{NM \times 1}$ is the space-time snapshot and $\mathbf{R}_{\text{RFI}} = \sum_{i=1}^J \mathbf{R}_{\text{RFI}}(i)$.

As the fractional bandwidth—the ratio of waveform bandwidth to center frequency—of the receive signal increases, dispersion occurs. Dispersion leads to decorrelation of the RFI over the receive array; in this case, a single RFI signal appears as multiple, closely spaced sources in angle [23]. In general, the wideband RFI suppression case is handled similar to the narrowband case through the use of subband filtering or the use of fast-time taps. In the former case, the subbanding is commonly implemented using polyphase filtering, allowing the processor to implement the narrowband canceler in each subband prior to recombining [25].

2.12.3.3 Receiver noise

Uncorrelated noise sources—such as receiver noise or sky noise—are modeled as a white Gaussian noise (WGN) disturbance, $\mathbf{n}_k = [\beta_k(m, nT)]_{m=1:M, n=0:N}$, where m and n are channel and pulse indices, respectively, and $\beta_k(m, nT) \sim \text{CN}(0, \sigma_n^2)$. The waveform samples are assumed uncorrelated such that

$$\mathbf{R}_n = \sigma_n^2 \mathbf{I}_{NM}, \quad (12.80)$$

where σ_n^2 is the channel noise power. This noise source is also uncorrelated, independent, and identically distributed over the range dimension.

2.12.3.4 Target

The target snapshot is generally modeled using both Swerling 1 and Swerling 2 statistics [1, 7, 8]. Swerling 1 and Swerling 2 targets each exhibit voltages corresponding to a circular Gaussian distribution; the Gaussian nature of the voltage distribution models target fading effects. The target voltage is assumed

perfectly correlated within the CPI, in accord with the Swerling 1 model, and uncorrelated from CPI-to-CPI according to the Swerling 2 target model. Frequency hopping from CPI-to-CPI is a common reason for target voltage decorrelation and is used to minimize the impact of target fading on detection performance. STAP is applied on a CPI basis, as it is a coherent signal processing technique, with noncoherent addition (NCA) applied to the STAP output—from CPI-to-CPI—to mitigate target fading effects.

Using the Swerling 1 target model, the target snapshot at the k th range, Doppler frequency, $f_{d,T}$, and angle, (ϕ_T, θ_T) , is

$$\mathbf{t}_k = \varsigma_T \sigma_n \sqrt{\text{SNR}_{s-t}} (f_{d,T}, \phi_T, \theta_T), \quad (12.81)$$

where $\varsigma_T \sim \text{CN}(0, 1)$ and SNR is the single channel, single pulse signal-to-noise ratio. The random variable, ς_T , models targeting fading resulting from subscatterers adding in and out of phase. As indicated, to handle fading, it is common to frequency hop from CPI-to-CPI, in which case target voltages appear decorrelated (Swerling 2).

Nonfluctuating, point-target analysis employs (12.81) with $\varsigma_T \rightarrow e^{j2\pi u}$, where u is uniformly distributed between $[0, 1]$. Sometimes the nonfluctuating target model is used in analysis. However, the Swerling 2 model is the preferred choice, with the Swerling 1 model applicable at the fixed frequency, CPI-level.

2.12.3.5 The space-time snapshot

The space-time snapshot for realization, k , is given by (12.2) with the addition of clutter discrete, where the possibility of the single target case is

$$\begin{aligned} H_0 : \mathbf{x}_k &= \mathbf{c}_k + \sum_{l=1}^D \mathbf{c}_{dl}(l; k(l)) + \mathbf{j}_k + \mathbf{n}_k, \\ H_1 : \mathbf{x}_k &= \mathbf{x}_{k/H_0} + \mathbf{t}_k, \end{aligned} \quad (12.82)$$

where D is the total number of discrete scatterers and the snapshot terms, $\mathbf{c}_{dl}(l; k(l))$, are only added when $k = k(l)$. The null-hypothesis covariance matrix corresponding to (12.82) is

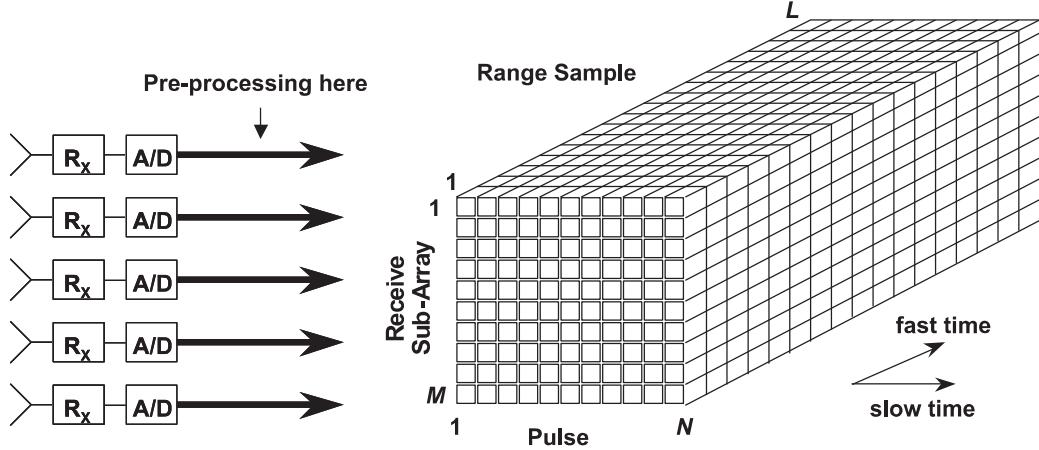
$$\mathbf{R}_k = \mathbf{R}_{c,k} + \sum_{l=1}^D \mathbf{R}_{cd}(l; k(l)) + \sum_{i=1}^J \mathbf{R}_j(i) + \mathbf{R}_n, \quad (12.83)$$

where, as expected, $\mathbf{R}_{cd}(l; k(l))$ is included only for those terms where $k = k(l)$.

It is common to envision the collection of space-time snapshots organized as space-time data matrices,

$$\mathbf{X}_k = [\mathbf{x}_{s,k}(1) \ \mathbf{x}_{s,k}(2) \ \cdots \ \mathbf{x}_{s,k}(N)] \in C^{M \times N}. \quad (12.84)$$

The spatial snapshot for pulse, n , is denoted as $\mathbf{x}_{s,k}(n) \in C^{M \times 1}$; it results by removing N length- M segments from \mathbf{x}_k and stacking them side-by-side. The pictorial of (12.84) is given in Figure 12.10 and is known as the radar datacube. Generally, the STAP operates on each space-time data matrix, or slice, of the cube in Figure 12.10, while using slices at other ranges (or realizations) for training. We subsequently describe various adaptive implementations. (Note: space/fast-time processing uses slices along the pulse/fast-time domain and can train over the pulse dimension. As previously mentioned, we focus on space/slow-time adaptivity, but the basic formulations of the next section generally apply to any appropriately formatted data with the corresponding restrictions of each method.)

**FIGURE 12.10**

Radar datacube (after [9], © 2003 IEEE).

2.12.4 Adaptive filter implementations

Section 2.12.2 describes space-time filter formulations. As seen from the discussion, all solutions are similar, involving space-time weightings of the form of a covariance inverse and a steering vector to implement the matched filter.

We now discuss three basic paradigms to implement the weighting strategies discussed in Section 2.12.2: reduced-rank STAP (RR-STAP), reduced-dimension STAP (RD-STAP), and parametric STAP.

2.12.4.1 Reduced-rank STAP

As seen from (12.78), each narrowband RFI source is rank-one. Distributed clutter is oftentimes of rank significantly less than the dimensionality of \mathbf{x}_k , viz. $r_\lambda \ll \dim(\mathbf{x}_k)$. Through empirical analysis, the clutter rank for a sidelooking radar with minimal yaw is approximated as

$$r_\lambda \approx \left\lfloor N + (M - 1) \left(\frac{v_p T}{d/2} \right) \right\rfloor, \quad (12.85)$$

where v_p is the platform along-track velocity and d is the separation between channels of a uniform linear array (ULA) [10]. The expression in (12.85) is known as Brennan's Rule; a related rule is given by Klemm [26]. Brennan's rule is closely related to the number of independent antenna channel positions during the collection interval; redundancy in the measurements lowers the rank, and in fact leads to coloration of the clutter return. The idea behind RR-STAP is to essentially project the interference subspace—those eigenspaces corresponding to larger eigenvalues above the noise floor—out of the space-time snapshot.

The eigendecomposition of the space-time covariance matrix yields

$$\mathbf{R}_k = \sum_{m=1}^{NM} \lambda_{k/m} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H = \mathbf{Q}_{k/I} \Lambda_{k/I} \mathbf{Q}_{k/I}^H + \mathbf{Q}_{k/N} \Lambda_{k/N} \mathbf{Q}_{k/N}^H, \quad (12.86)$$

where $\lambda_{k/m}$ is the m th eigenvalue corresponding to eigenvector $\mathbf{q}_{\mathbf{k}/\mathbf{m}} \in C^{NM \times 1}$, $\Lambda_{\mathbf{k}/\mathbf{I}} = \text{diag}(\lambda_{k/m})_{m=1}^P$ represents the interference subspace eigenvalues, and $\Lambda_{\mathbf{k}/\mathbf{N}} = \text{diag}(\lambda_{k/m})_{m=P+1}^{NM}$ characterizes the noise subspace eigenvalues. The interference eigenvectors have the special property that each span the collection of interference steering vectors. Moreover, $\mathbf{Q} = [\mathbf{Q}_{\mathbf{k}/\mathbf{I}} \ \mathbf{Q}_{\mathbf{k}/\mathbf{N}}]$ is unitary, so

$$\mathbf{Q}_{\mathbf{k}/\mathbf{I}} \mathbf{Q}_{\mathbf{k}/\mathbf{I}}^H + \mathbf{Q}_{\mathbf{k}/\mathbf{N}} \mathbf{Q}_{\mathbf{k}/\mathbf{N}}^H = \mathbf{I}_{NM}. \quad (12.87)$$

We can write the distributed clutter-plus-noise part of (12.83) in a generic, simplified form

$$\mathbf{R}_{\mathbf{k}} = \sum_{p=1}^{\tilde{N}} \sigma_{s/p}^2 \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi_p, \theta_p, f_{d/p}) + \sigma_n^2 \mathbf{I}_{NM}, \quad (12.88)$$

where \tilde{N} is the number of signal sources and $\sigma_{s/p}^2$ is the power for the p th signal source. While this simplified covariance form suggests a model for clutter-plus-noise, the subsequent results are extensible to other correlated sources. Then, from (12.88),

$$\mathbf{R}_{\mathbf{k}} \mathbf{q}_{\mathbf{k}/\mathbf{m}} = \lambda_{k/m} \mathbf{q}_{\mathbf{k}/\mathbf{m}} = \sum_{p=1}^{\tilde{N}} \sigma_{s/p}^2 \gamma_{k/m,p} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) + \sigma_n^2 \mathbf{q}_{\mathbf{k}/\mathbf{m}}, \quad (12.89)$$

where

$$\gamma_{k/m,p} = \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi_p, \theta_p, f_{d/p}) \mathbf{q}_{\mathbf{k}/\mathbf{m}}. \quad (12.90)$$

Solving (12.89) for the m th eigenvector, $\mathbf{q}_{\mathbf{k}/\mathbf{m}}$,

$$\mathbf{q}_{\mathbf{k}/\mathbf{m}} = \sum_{p=1}^{\tilde{N}} \left(\frac{\sigma_{s/p}^2 \gamma_{k/m,p}}{\lambda_{k/m} - \sigma_n^2} \right) \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}) = \sum_{p=1}^{\tilde{N}} \alpha_{k/m,p} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p}), \quad (12.91)$$

where in this case

$$\alpha_{k/m,p} = \frac{\sigma_{s/p}^2 \gamma_{k/m,p}}{\lambda_{k/m} - \sigma_n^2}. \quad (12.92)$$

Naturally, these equations are only valid for the dominant subspace, $\lambda_{k/m} - \sigma_n^2 > 0$. Equation (12.91) shows $\mathbf{q}_{\mathbf{k}/\mathbf{m}} \in \text{span}(\{\mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi_p, \theta_p, f_{d/p})\}_{p=1}^{\tilde{N}})$.

Also, it is known that any wide-sense stationary (WSS) process with zero mean and covariance matrix, $\mathbf{R}_{\mathbf{k}}$, can be written as a linear combination of the eigenvectors of $\mathbf{R}_{\mathbf{k}}$, via what is known as the Karhunen-Loëve Transform (KLT),

$$\mathbf{x}_{\mathbf{k}} = \sum_{m=1}^{NM} a_{k/m} \mathbf{q}_{\mathbf{k}/\mathbf{m}}, \quad (12.93)$$

where $a_{k/m} = \mathbf{q}_{\mathbf{k}/\mathbf{m}}^H \mathbf{x}_{\mathbf{k}}$ are the Karhunen-Loëve (KL) coefficients [12].

The prior expressions, (12.86)–(12.93), provide necessary mathematical background for our subsequent discussion on RR-STAP methods.

Motivation for RR-STAP includes the following:

- Clutter and jamming tend to be of low numerical rank and the processor explicitly removes only those signal subspaces corresponding to interference.
- The eigendecomposition maximally compresses the interference into the fewest basis vectors [12].
- The RMB rule applies to the RR-STAP case, with rank replacing DoFs in the formulation, viz. training over twice the rank yields roughly 3 dB loss on average [27].

In effect, RR-STAP is a weight calculation strategy. There are challenges implementing RR-STAP, including high computational burden and difficulties in rank determination. Nevertheless, RR-STAP methods hold meaningful insight and, in some cases, are useful in weight vector determination.

2.12.4.1.1 Adaptive beamformer pattern synthesis

It is shown in [28] that (12.43) can be written

$$\mathbf{w}_k = \mu \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d) = \frac{\mu}{\lambda_0} \left[\mathbf{s}_{s-t}(\phi, \theta, f_d) - \sum_{m=1}^{NM} \frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \vartheta_{k/m} \mathbf{q}_{k/m} \right], \quad (12.94)$$

where $\lambda_0 = \min(\{\lambda_{k/m}\}_{\forall m})$ is the noise-floor eigenvalue level and $\vartheta_{k/m} = \mathbf{q}_{k/m}^H \mathbf{s}_{s-t}(\phi, \theta, f_d)$ is the projection of the m th eigenvector onto the quiescent pattern. As seen from (12.94), the STAP response appears as a notching of the space-time beampattern given by $\mathbf{s}_{s-t}(\phi, \theta, f_d)$ by the weighted, interference eigenvectors. When $\lambda_{k/m} = \lambda_0$, no subtraction occurs, since the corresponding eigenvector lies in the noise subspace.

Naturally, (12.94) applies to the SMI formulation of (12.44), and can provide robustness in the presence of training sample support limitations if the interference rank is known (which is usually only practical for strong interferers). Specifically, running the sum in (12.94) only over $r_\lambda \ll NM$ subspaces leads to adaptive pattern robustness, since low sample support tends to predominantly perturb the noise subspace; the perturbed noise subspace estimate leads to poor adaptive sidelobe performance when the sum in (12.94) is run over all values of the index, m .

2.12.4.1.2 Principle components inverse

The idea behind principle components inverse (PCI) is to apply an orthogonal projection to \mathbf{x}_k , and then apply a matched filter to the remaining transformed data [29]. The PCI formulation starts by writing (12.94) applied to the data as

$$y_k = \frac{\mu \mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \left[\mathbf{I}_{NM} - \sum_{m=1}^{NM} \frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k. \quad (12.95)$$

Then, for the $P \ll NM$ eigenvalues where $\lambda_{m/k} \gg \lambda_0$,

$$y_k \approx \frac{\mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \left[\mathbf{I}_{NM} - \sum_{m=1}^{P \ll NM} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k = \frac{\mathbf{s}_{s-t}^H(\phi, \theta, f_d)}{\lambda_0} \tilde{\mathbf{x}}_k. \quad (12.96)$$

We see that

$$\tilde{\mathbf{x}}_k = \left[\mathbf{I}_{NM} - \sum_{m=1}^{P \ll NM} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right] \mathbf{x}_k \stackrel{\text{e.g.}}{=} \mathbf{x}_k - a_{k/1} \mathbf{q}_{k/1} - a_{k/2} \mathbf{q}_{k/2}, \quad (12.97)$$

where, for the case of $P = 2$ as an example, (12.97) shows the coherent removal of $\mathbf{q}_{k/1}$ and $\mathbf{q}_{k/2}$. The term in brackets in (12.97) is called an orthogonal projection for this reason.

2.12.4.1.3 Eigencanceler

Haimovich [30] gives two different developments; one of these essentially leads to (12.96) via an alternate route.

The minimum power eigencanceler (MPE) weighting results from

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k \text{ such that } \mathbf{Q}_{k/I}^H \mathbf{w}_k = \mathbf{0} \text{ and } \mathbf{C}^H \mathbf{w}_k = \mathbf{f}, \quad (12.98)$$

where \mathbf{C} is a constraint matrix and \mathbf{f} is a desired response vector. The solution to (12.98) is shown in [30] to be

$$\mathbf{w}_{k/\text{EC-MPE}} = \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{C} \left[\mathbf{C}^H \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{C} \right]^{-1} \mathbf{f}. \quad (12.99)$$

For the case where $\mathbf{C} = \mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{f} = g$, (12.99) takes the form

$$\mathbf{w}_{k/\text{EC-MPE}} = \mu_0 \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.100)$$

with $\mu_0 = [\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{Q}_{k/N} \Lambda_{k/N}^{-1} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d)]^{-1}$ g a scalar. The weight vector in (12.100) lies entirely in the noise subspace—as required by (12.98)—and will, by virtue of the orthonormal property of the set of eigenvectors, annihilate the clutter subspace.

The minimum norm eigencanceler (MNE) is formulated as

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{w}_k \text{ subject to } \mathbf{Q}_{k/I}^H \mathbf{w}_k = \mathbf{0} \text{ and } \mathbf{C}^H \mathbf{w}_k = \mathbf{f}. \quad (12.101)$$

The MNE weight vector is then shown in [30] to be

$$\mathbf{w}_{k/\text{EC-MNE}} = \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{C} \left[\mathbf{C}^H \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{C} \right]^{-1} \mathbf{f}. \quad (12.102)$$

Again, for the case where $\mathbf{C} = \mathbf{s}_{s-t}(\phi, \theta, f_d)$ and $\mathbf{f} = g$, and using (12.87), we see (12.102) can be written in the same form as PCI,

$$\mathbf{w}_{k/\text{EC-MNE}} = \mu_1 \left[\mathbf{I}_{NM} - \mathbf{Q}_{k/I} \mathbf{Q}_{k/I}^H \right] \mathbf{s}_{s-t}(\phi, \theta, f_d) \quad (12.103)$$

with $\mu_1 = [\mathbf{s}_{s-t}^H(\phi, \theta, f_d) \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d)]^{-1}$ g a scaling that replaces $1/\lambda_0$ in (12.96). Using (12.87), the expression in (12.103) can be written

$$\mathbf{w}_{k/\text{EC-MNE}} = \mu_1 \mathbf{Q}_{k/N} \mathbf{Q}_{k/N}^H \mathbf{s}_{s-t}(\phi, \theta, f_d), \quad (12.104)$$

which closely relates to the MPE solution in (12.100) with the inverse eigenvalue weighting given by $\Lambda_{k/N}^{-1}$ absent. It is known that limited training sample support leads to perturbed estimates of the noise eigenvalues appearing along the diagonal of $\Lambda_{k/N}$. For this reason, the MNE provides robustness relative to MPE when training data samples are lacking.

2.12.4.1.4 Hung turner projection

The Hung Turner Projection (HTP) is applicable to RFI suppression [31]. While clutter mitigation is paramount to our discussion, it is worth taking a moment to describe the HTP as a general tool in adaptive filter implementation. Reducing computational burden associated with the eigendecomposition leading to the eigenvalues and eigenvectors of the null-hypothesis covariance matrix is a primary goal.

The basic idea behind the HTP is to use the Gram-Schmidt method to characterize the interference subspace. Specifically, suppose we identify J snapshots where J RFI sources are present,

$$\begin{aligned} \mathbf{x}_k &= \mathbf{j}_k(1) + \mathbf{j}_k(2) + \cdots + \mathbf{j}_k(J) + \mathbf{n}_k, \\ \mathbf{x}_{k+1} &= \mathbf{j}_{k+1}(1) + \mathbf{j}_{k+1}(2) + \cdots + \mathbf{j}_{k+1}(J) + \mathbf{n}_{k+1}, \\ &\vdots \\ \mathbf{x}_{k+J-1} &= \mathbf{j}_{k+J-1}(1) + \mathbf{j}_{k+J-1}(2) + \cdots + \mathbf{j}_{k+J-1}(J) + \mathbf{n}_{k+J-1}. \end{aligned} \quad (12.105)$$

Next, the modified Gram-Schmidt technique is applied to (12.105) to generate the orthonormal basis, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_J\}$, where $\mathbf{v}_m \in \text{span}\{\mathbf{j}_1, \mathbf{j}_2, \dots, \mathbf{j}_J\}$. Then, in accord with [31], the HTP weight vector is chosen as $\mathbf{w}_{\text{HTP}} = \bar{\mathbf{w}}_Q^0 / \|\bar{\mathbf{w}}_Q^0\|_2$, with

$$\bar{\mathbf{w}}_Q^0 = \bar{\mathbf{w}}_Q - \sum_{m=1}^J \left(\mathbf{v}_m^H \bar{\mathbf{w}}_Q \right) \mathbf{v}_m, \quad (12.106)$$

and where $\bar{\mathbf{w}}_Q$ is the quiescent weight vector.

2.12.4.1.5 Diagonal loading

Diagonal loading involves adding a scaled diagonal matrix to the covariance matrix estimate of (12.49) [32],

$$\hat{\mathbf{R}}_k = \frac{1}{K} \sum_{m=1}^K \mathbf{x}_m \mathbf{x}_m^H + \sigma_{\text{DL}}^2 \mathbf{I}_{NM} = \hat{\mathbf{Q}}_k \hat{\Lambda}_k \hat{\mathbf{Q}}_k^H + \sigma_{\text{DL}}^2 \hat{\mathbf{Q}}_k \hat{\mathbf{Q}}_k^H = \hat{\mathbf{Q}}_k \left(\hat{\Lambda}_k + \sigma_{\text{DL}}^2 \mathbf{I}_{NM} \right) \hat{\mathbf{Q}}_k^H, \quad (12.107)$$

where σ_{DL}^2 is the diagonal loading level and $\hat{\mathbf{Q}}_k$ and $\hat{\Lambda}_k$ follow from the decomposition of the covariance matrix estimate. It is common to set $0.1\sigma_n^2 < \sigma_{\text{DL}}^2 < 10\sigma_n^2$, with $\sigma_{\text{DL}}^2 = \sigma_n^2$ most typical. Primary benefits of diagonal loading include improved conditioning of the covariance matrix estimate (which improves the numerical stability of inversion routines) and compression of the range corresponding to the smallest eigenvalues. This latter benefit leads to a reduction in spurious sidelobes of the adapted pattern at the expense of null depth. Effectively, diagonal loading removes some of the adaptivity of the system to better condition the filter's sidelobe response.

While often considered ad hoc, it turns out diagonal loading is a component of the optimal solution to the constrained optimization,

$$\min_{\mathbf{w}_k} \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k \text{ such that } \mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) = 1 \text{ and } \mathbf{w}_k^H \mathbf{w}_k \leq \beta^2. \quad (12.108)$$

The Lagrangian for (12.108) is readily found to be

$$L_c = \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k + \lambda_1 \left(\mathbf{w}_k^H \mathbf{s}_{s-t}(\phi, \theta, f_d) - 1 \right) + \lambda_2 \left(\mathbf{w}_k^H \mathbf{w}_k - \beta^2 \right), \quad (12.109)$$

where λ_1 and λ_2 are Lagrangian multipliers. Taking the gradient of (12.109) with respect to the conjugate of the weight vector and setting the result to zero yields [13]

$$\nabla_{\mathbf{w}^H} L_c = \mathbf{R}_k \mathbf{w}_k + \lambda_1 \mathbf{s}_{s-t}(\phi, \theta, f_d) + \lambda_2 \mathbf{w}_k = 0. \quad (12.110)$$

Solving for the weight vector gives

$$\mathbf{w}_k = -\lambda_1 (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.111)$$

Using the linear constraint, it is seen that

$$\lambda_1 = \frac{-1}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.112)$$

The resulting weight vector is

$$\mathbf{w}_k = \frac{(\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}{\mathbf{s}_{s-t}^H(\phi, \theta, f_d) (\mathbf{R}_k + \lambda_2 \mathbf{I}_{NM})^{-1} \mathbf{s}_{s-t}(\phi, \theta, f_d)}. \quad (12.113)$$

From (12.113) we see that diagonal loading plays a key role in the solution to the constrained optimization. In this case, λ_2 is the diagonal loading level, σ_{DL}^2 , mentioned previously. If the loading level is set to zero, (12.113) defaults to the distortionless MV beamformer solution.

2.12.4.1.6 Cross spectral metric

Consider the GSLC structure of Figure 12.5. A principal components approximation to $\mathbf{R}_{x_0/k}$ is

$$\mathbf{R}_{x_0/k} \approx \sum_{m=1}^P \lambda_{x_0/k;m} \mathbf{q}_{x_0/k;m} \mathbf{q}_{x_0/k;m}^H \quad (12.114)$$

with, ideally, $P \ll NM$, and where $\lambda_{x_0/k;m}$ and $\mathbf{q}_{x_0/k;m}$ represent the m th eigenvalue and eigenvector of the GSLC lower leg covariance matrix. Then,

$$\mathbf{w}_{MMSE-PC/k} = \sum_{m=1}^P \left(\frac{\mathbf{q}_{x_0/k;m}^H \mathbf{r}_{x_0/k} \mathbf{d}_{0/H0}}{\lambda_{x_0/k;m}} \right) \mathbf{q}_{x_0/k;m}. \quad (12.115)$$

The computational burden of the GSLC does not generally justify its use in practice. Improved statistical convergence is a potential advantage of the principal components decomposition of the GSLC (PC-GSLC).

The cross spectral metric operates in a fashion similar to the PC-GSLC method, except now the selection of the basis defining the rank-reducing transformation is target signal-dependent [33]. The error signal at the output of the GSLC follows as the difference between (12.19) and (12.21), viz.

$$\varepsilon_{o/k} = d_{o/k} - \hat{d}_{o/k}. \quad (12.116)$$

The corresponding MMSE is given as the expected value of the magnitude squared of (12.116),

$$\begin{aligned} E \left[|\varepsilon_{o/k}|^2 \right] &= E \left[\varepsilon_{o/k} \varepsilon_{o/k}^* \right] = \sigma_{d_{o/H0}}^2 - \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}^H \mathbf{R}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}^{-1} \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}} \\ &= \sigma_{d_{o/H0}}^2 - \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}^H \left(\sum_{m=1}^{NM} \lambda_{\mathbf{x}_{o/k};m}^{-1} \mathbf{q}_{\mathbf{x}_{o/k};m} \mathbf{q}_{\mathbf{x}_{o/k};m}^H \right) \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}} \\ &= \sigma_{d_{o/H0}}^2 - \sum_{m=1}^{NM} \lambda_{\mathbf{x}_{o/k};m}^{-1} \left| \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}^H \mathbf{q}_{\mathbf{x}_{o/k};m} \right|^2. \end{aligned} \quad (12.117)$$

In this case, $\sigma_{d_{o/H0}}^2 = E \left[|d_{o/k}|^2 \right]$, where no target signal energy is present in the upper leg (an assumption that carries through in the calculation of the cross correlation vector, $\mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}$). When choosing only P terms in the summation of (12.117), where $P \ll NM$, the processor attains the lowest possible mean square error (MSE) by choosing subspaces with the largest cross-spectral metric terms, given as

$$\text{CSM}_m = \frac{\left| \mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}^H \mathbf{q}_{\mathbf{x}_{o/k};m} \right|^2}{\lambda_{\mathbf{x}_{o/k};m}}. \quad (12.118)$$

The target signal influences the selection via $\mathbf{r}_{\mathbf{x}_{o/k} \mathbf{d}_{o/H0}}$. Thus, the idea behind the CSM approach is to select only those subspaces from $\mathbf{R}_{\mathbf{x}_{o/k}}$ that most greatly enhance performance.

A direct-form implementation of the CSM method is given in [34]. In this case, the objective involves maximizing SINR by choosing the appropriate, signal-dependent, reduced-rank subspace (hence, the analogous metric is called the SINR metric in [34]). Thus, decomposing output SINR in (12.7) using the inverse of the eigendecomposition in (12.86) gives

$$\text{SINR}_{\text{Max}} = \sigma_T^2 \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \left(\sum_{m=1}^{NM} \lambda_{k/m}^{-1} \mathbf{q}_{k/m} \mathbf{q}_{k/m}^H \right) \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\phi, \theta, f_d). \quad (12.119)$$

Maximizing SINR requires choosing the largest terms in (12.119) of the form,

$$\text{CSM}_{\text{DFP};m} = \frac{\left| \mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \mathbf{q}_{k/m} \right|^2}{\lambda_{k/m}}, \quad (12.120)$$

where DFP signifies “direct form processor” (i.e., the approach using a weight vector of the form given in (12.43) or (12.44)). Thus, with limited sample support and uncertainty over the interference rank, selecting the $P \ll NM$ largest CSM terms of (12.120) leads to maximal output SINR. Commonly, then, the CSM-DFP chooses the components in the noise subspace, but this is an artifact of the development and not useful: terms with the smallest eigenvalues are generally chosen, with any signal-dependent correlation meant to aid in selecting the most significant terms lost in the formulation. To avoid this scenario, [34] suggests choosing those terms maximally impacting the weight vector. In other words, considering (12.94), selecting the P largest terms of

$$\frac{\lambda_{k/m} - \lambda_0}{\lambda_{k/m}} \vartheta_{k/m} \mathbf{q}_{k/m} \quad (12.121)$$

most greatly impacts the adaptive process. Equation (12.121) is signal-dependent and forces the selection to lie in the dominant subspace.

2.12.4.1.7 Multi-Stage Wiener Filter

The Multi-Stage Wiener Filter (MWF) is a truncated decomposition of the GSLC of Figure 12.5 [35]. Its operation is best understood from the diagram in Figure 12.11, which shows a two-stage decomposition. From this figure we see that the estimation stage is broken into a series of smaller problems, where the weight vector is calculated as a scalar using the Wiener-Hopf equation [12,13] in each stage. The filter, \mathbf{h}_i , is a normalized cross-correlation vector between its input, $\mathbf{x}_{i-1/k}$, and the output of the filter in the preceding leg, $d_{i-1/k}$; the intent of choosing \mathbf{h}_i in this manner is to maximize the correlation between the interference signal in each leg. Specifically,

$$\mathbf{h}_i = \frac{\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}}}{\|\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}}\|_2}, \quad (12.122)$$

where $\mathbf{r}_{\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}} = E[\mathbf{x}_{i-1/k} \mathbf{d}_{i-1/k}^*]$. The blocking matrix, \mathbf{B}_i , lies in the null space of \mathbf{h}_i (except for $i = 0$, where $\mathbf{B}_G = \mathbf{B}_0 = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$). The weights, w_i , are scalars equal to the Wiener weight minimizing the mean square error between $d_{i-1/k}$ and $\hat{d}_{i-1/k} = w_i^* \varepsilon_i$. In this figure, for the MWF output to equal the GSLC output, the processor must calculate all quantities, including the vector weight, $\mathbf{w}_{MWF,3} = (E[\mathbf{x}_{2/k} \mathbf{x}_{2/k}^H])^{-1} \mathbf{r}_{\mathbf{x}_{2/k} \mathbf{d}_{2/k}}$. The MWF truncates the number of stages by simply dropping the lower leg of the last stage, setting the vector weight to all zeros (i.e., $\mathbf{w}_{MWF,3} = \mathbf{0}$) so that $\varepsilon_{N_{MWF}} = d_{N_{MWF}/k}$, where N_{MWF} is the number of MWF stages. So, for $N_{MWF} = 2$, $\varepsilon_2 = d_{2/k} = \mathbf{h}_2^H \mathbf{x}_{1/k}$.

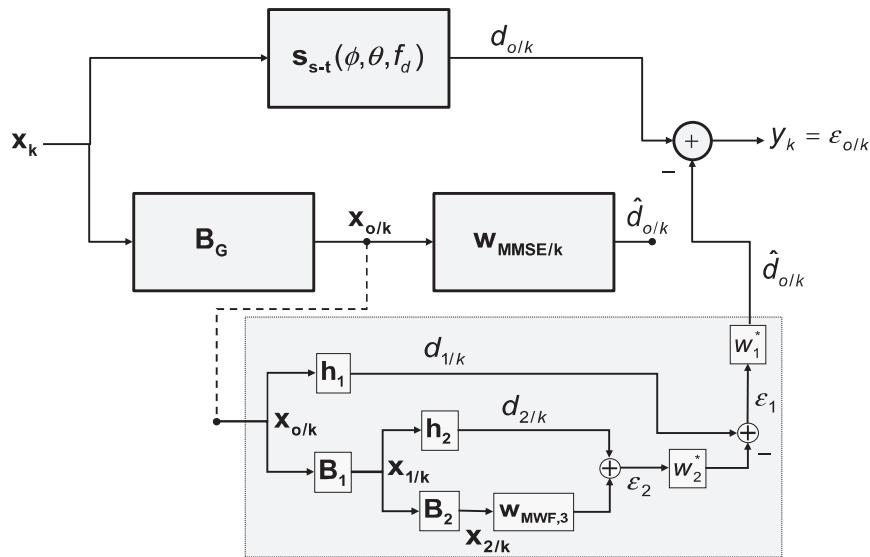


FIGURE 12.11

Multi-Stage Wiener Filter flow diagram.

In the adaptive version of the MWF, the processor replaces the known covariance matrices and steering vectors with estimates. The quantities, \mathbf{h}_i , \mathbf{B}_i , w_i , and, naturally, ε_i , are all expressible as linear transformations applied to \mathbf{R}_k and $\mathbf{s}_{s-t}(\phi, \theta, f_d)$; in the adaptive filter, $\mathbf{R}_k \rightarrow \hat{\mathbf{R}}_k$ and $\mathbf{s}_{s-t}(\phi, \theta, f_d) \rightarrow \mathbf{v}_{s-t}(\phi, \theta, f_d)$. The implementation requires first calculating and applying $\mathbf{B}_G = \mathbf{B}_0 = \text{null}(\mathbf{s}_{s-t}(\phi, \theta, f_d))$; next, the processor implements the first stage, calculating \mathbf{h}_1 , \mathbf{B}_1 , and $d_{1/k}$; with these first stage quantities in hand, the processor calculates the second stage quantities, \mathbf{h}_2 , \mathbf{B}_2 , and $d_{2/k}$, then third stage quantities, and so forth, working out to the last stage; then, as noted above, the error signal in the last stage is set to $\varepsilon_L = d_{L/k} = \mathbf{h}_L^H \mathbf{x}_{L-1/k}$, where $i = L$ is the last stage; and, finally, the scalar weights, w_i , are solved from the outer (last) stage into the first stage.

While the computational loading of the MWF is generally high, the target signal-dependent nature of the stage decomposition identifies the interference subspace most greatly impacting the MMSE through the cross-correlation process, estimates it, and subtracts it out. This allows the MWF to converge towards the optimal solution with minimal use of available training data.

2.12.4.2 Reduced-dimension STAP

Reduced-dimension STAP (RD-STAP) methods take advantage of the structure of the clutter angle-Doppler region of support to minimize filter length, thus reducing computational burden and training sample requirements. A taxonomy of RD-STAP methods, organized similarly to the discussion in [10], is given in Figure 12.12.

RD-STAP methods apply deterministic, linear filtering operations and dimensionality reduction prior to adaptive filtering. A linear transformation, $\mathbf{T} \in C^{NM \times \tilde{P}}$, where \tilde{P} is the length of the reduced

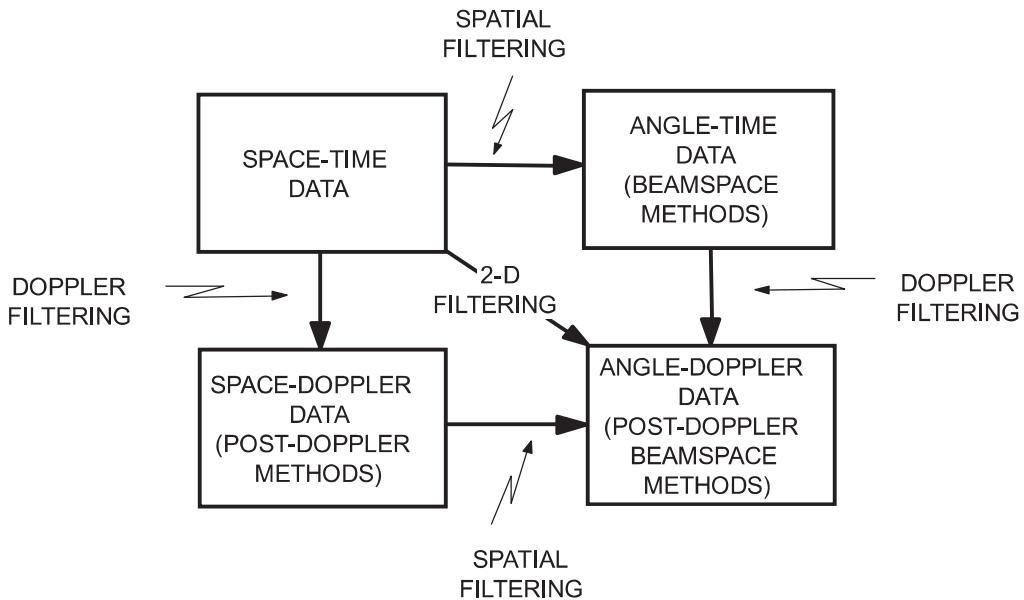


FIGURE 12.12

RD-STAP taxonomy.

dimension, describes these deterministic operations. Starting with the space-time snapshot, \mathbf{x}_k , the reduced-dimension snapshot is

$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \in C^{\tilde{P}_x 1}. \quad (12.123)$$

Via this transformation, the null-hypothesis covariance matrix is

$$\tilde{\mathbf{R}}_k = \mathbf{T}^H \mathbf{R}_k \mathbf{T} \in C^{\tilde{P}_x \tilde{P}} \quad (12.124)$$

and the target steering vector becomes

$$\tilde{\mathbf{s}}_{RD} = \mathbf{T}^H \mathbf{s}_{s-t}(\phi, \theta, f_d). \quad (12.125)$$

The corresponding optimal weighting follows from (12.43) as

$$\tilde{\mathbf{w}}_{RD/k} = \tilde{\mu} \left(\mathbf{T}^H \mathbf{R}_k \mathbf{T} \right)^{-1} \left(\mathbf{T}^H \mathbf{s}_{s-t}(\phi, \theta, f_d) \right). \quad (12.126)$$

The adaptive weighting usually follows by forming the sample covariance matrix of (12.49) from the reduced-dimension data,

$$\hat{\mathbf{R}}_k = \frac{1}{\tilde{k}} \sum_{m=1}^{\tilde{k}} \mathbf{T}^H \mathbf{x}_m \mathbf{x}_m^H \mathbf{T}, \quad (12.127)$$

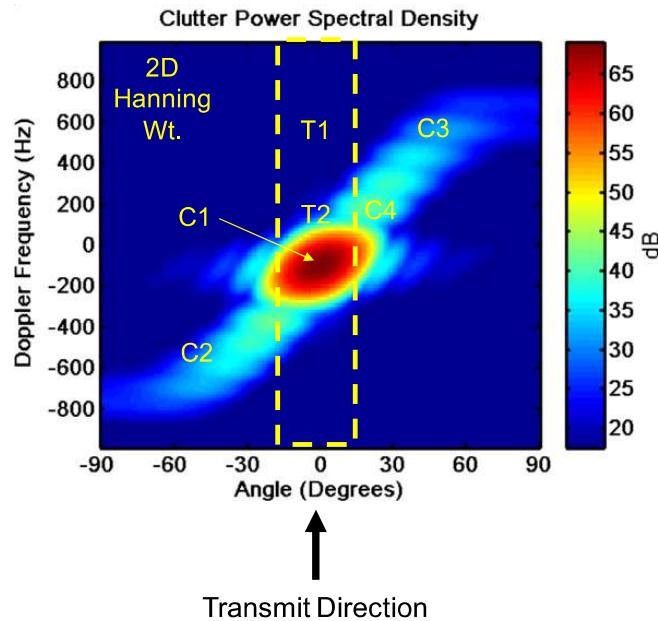
forming its inverse, and using $\tilde{\mathbf{v}}_{RD} = \mathbf{T}^H \mathbf{v}_{s-t}(\phi, \theta, f_d)$ in place of (12.125), viz.

$$\hat{\mathbf{w}}_{RD/k} = \hat{\mu} \hat{\mathbf{R}}_k^{-1} \tilde{\mathbf{v}}_{RD}. \quad (12.128)$$

In this case, $\hat{\mu}$ is a scalar that usually depends on estimated quantities.

Motivations for RD-STAP include the following:

- The clutter exhibits a highly structured angle-Doppler region of support, as shown in Figure 12.13. In this figure, T1 and T2 are target locations and C1–C4 are clutter regions. Target T1 predominantly competes with clutter located at C3, and a spatial null within the corresponding Doppler filter suffices to mitigate clutter. Clutter at other locations—such as at C1, C2, and C4—has no significant bearing on target detection. Alternately, target T2 is impacted by clutter in the region of C1 and C4; a two-dimensional null, positioned in angle and Doppler, is needed to suppress mainlobe and near-in sidelobe clutter.
- The RMB rule of Section 2.12.2.4.2 applies to RD-STAP, with $NM \rightarrow \tilde{P}$. It is not uncommon to have $9 \leq \tilde{P} \leq 50$, whereas $500 \leq NM \leq 5000$ is reasonable. Thus, limited and potentially heterogeneous or nonstationary training sample support is reduced by a factor of 10–100 or more.
- The implicit inverse of the sample covariance matrix is $O(N^3 M^3)$, being reduced to $O(\tilde{P}^3)$ in the RD-STAP case. For each halving of the processor’s DoFs, computational burden decreases by a factor of eight. Real-time operation requires RD-STAP methods.

**FIGURE 12.13**

Description of RD-STAP benefits.

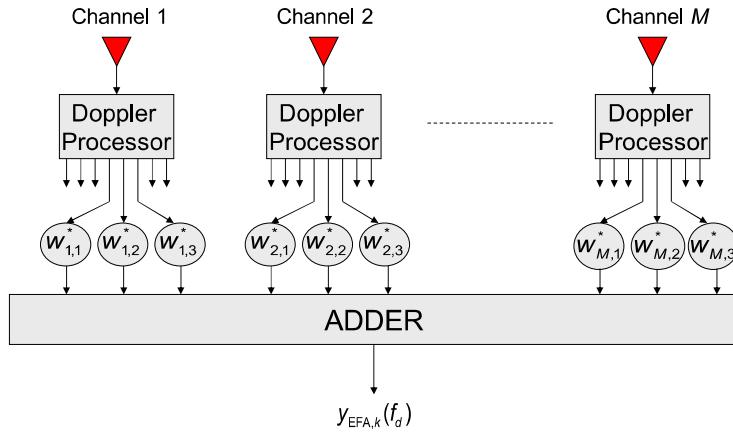
2.12.4.2.1 Post-Doppler STAP

Of the various RD-STAP methods, post-Doppler STAP techniques are most popular. Post-Doppler STAP is shown in the lower left and right of Figure 12.12, where the lower left box corresponds to the so-called “channel space,” post-Doppler methods that operate on space-Doppler data, and the lower right box characterizes the post-Doppler, (spatial) beamspace methods. The Extended Factored Algorithm (EFA) is an example of a “channel space,” post-Doppler method [36]. In EFA, each channel is Doppler filtered, and then the adaptive weighting of (12.128) is applied to the transformed data vector,

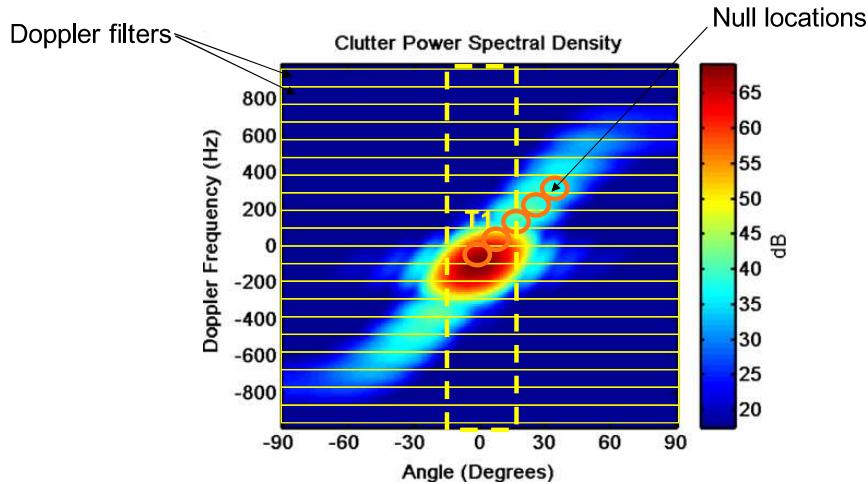
$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \stackrel{\text{e.g.}}{=} [\mathbf{f}_{s,k}^T(n-2) \quad \mathbf{f}_{s,k}^T(n-1) \quad \mathbf{f}_{s,k}^T(n) \quad \mathbf{f}_{s,k}^T(n+1) \quad \mathbf{f}_{s,k}^T(n+2)]^T, \quad (12.129)$$

where $\mathbf{f}_{s,k}(q) \in \mathbb{C}^{M \times 1}$ is the spatial snapshot corresponding to the q th Doppler bin output (i.e., the q th Doppler bin output from channel 1 through M stacked in a vector). Equation (12.129) shows, as an example, the case of five adjacent Doppler bins. The EFA output for the k th range bin and n th Doppler bin is then given by $y_{\text{EFA}/k}(f_d) = \hat{\mathbf{w}}_{\text{RD}/k}^H \tilde{\mathbf{x}}_k$, where f_d corresponds to the n th Doppler bin output. Figure 12.14 depicts the EFA processing architecture, where $w_{m,q}$ is the EFA weight applied to the m th channel and q th Doppler bin. This figure shows the case of processing three adjacent Doppler bins; EFA typically employs three or five adjacent bins. The special case of one Doppler bin—where the processor only generates a spatial null—is known as Factored Time-Space (FTS) [10].

Figure 12.15 describes EFA operation in the angle-Doppler domain. Each horizontal, rectangular box corresponds to a Doppler filter extent. The circles correspond to null locations EFA is able to generate. EFA can generate up to $N_t M - 1$ nulls, where N_t is the number of temporal DoFs (the number of Doppler

**FIGURE 12.14**

Extended Factored Algorithm (EFA) implementation.

**FIGURE 12.15**

EFA operation in the angle-Doppler domain.

bins used in the adaptive combiner). The target location, T1, is shown close to mainlobe clutter. Nulling along the section of the clutter ridge in proximity to T1 is required for target detection. EFA is able to effectively suppress the clutter local to the target position. Its performance benchmarks very close to that of the space-time optimal solution, as will be shown in Section 2.12.5. FTS can generate up to $M - 1$ spatial nulls within the Doppler bin; as seen in this case, given the closeness of T1 to mainlobe clutter, only nulling sidelobe clutter is insufficient.

References [37,38] describe post-Doppler, (spatial) beamspace methods. The Joint Domain Localized (JDL) algorithm generalizes EFA by applying a beamspace transformation prior to Doppler filtering in Figure 12.14. JDL then combines M_s spatial beams and N_t Doppler bins to suppress clutter, in a

manner similar to a two-dimensional sidelobe canceler. For example, a common configuration is $M_s = 3$ and $N_t = 3$; in this three beam configuration, the processor uses the eight angle-Doppler beams surrounding the center beam to estimate and coherently remove the clutter signal in this ninth angle-Doppler direction of interest, thereby exploiting the structure of the clutter local to the target position. JDL performance tends to benchmark well relative to the space-time optimal solution. A special case of JDL is given in [38], where the Doppler filtering is applied to sum and difference beams, and then “auxiliary” beams about the sum beam and Doppler bin of interest are adapted to coherently remove the clutter signal. This latter method has been called Sigma-Delta STAP. Reference [39] also includes germane discussion.

Additional post-Doppler STAP methods based on PRI designs are given in [10].

2.12.4.2.2 Pre-Doppler STAP

Pre-Doppler STAP—sometimes called adaptive displaced phase center antenna (ADPCA)—involves processing overlapped sub-CPIs to mitigate clutter, then applying Doppler processing to the aggregated output [10, 40]. It has particular application in those situations where the clutter response decorrelates during the course of the CPI, such as when the antenna array rotates. Figure 12.16 shows the pre-Doppler STAP, or ADPCA, architecture for the three pulse case.

The corresponding ADPCA data snapshot for three pulses is simply

$$\tilde{\mathbf{x}}_k = \mathbf{T}^H \mathbf{x}_k \stackrel{\text{e.g.}}{=} [\mathbf{x}_{s,k}^T(n-1) \quad \mathbf{x}_{s,k}^T(n) \quad \mathbf{x}_{s,k}^T(n+1)]^T. \quad (12.130)$$

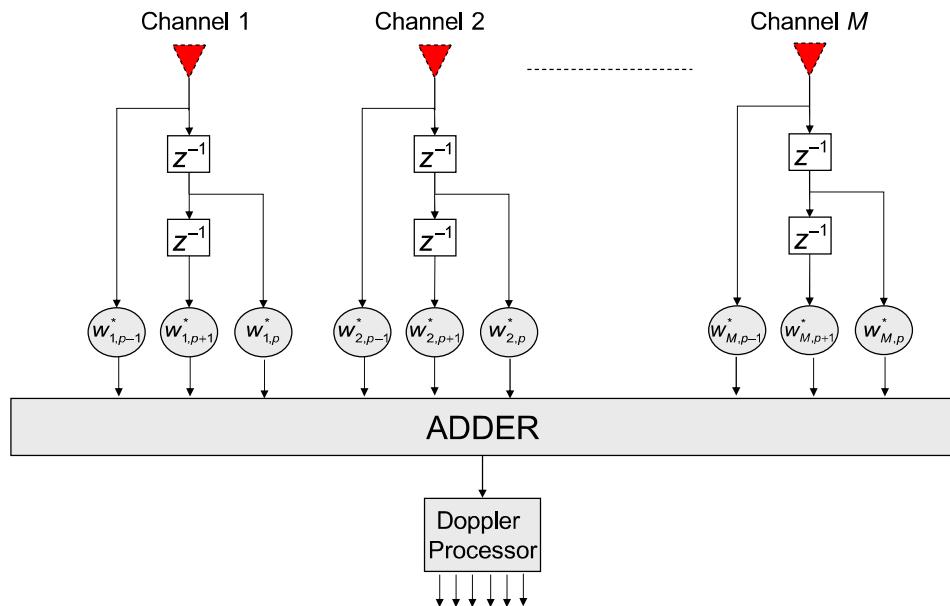


FIGURE 12.16

Pre-Doppler STAP.

The ADPCA weight vector is

$$\hat{\mathbf{w}}_{\text{adpca}/k} = \hat{\mathbf{R}}_{\text{adpca}/k}^{-1} \left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \odot \begin{pmatrix} 1 \\ \exp(j2\pi f_1) \\ \exp(j2\pi 2f_1) \end{pmatrix} \right) \otimes \mathbf{v}_s(\phi, \theta) \quad (12.131)$$

with f_1 the peak clutter Doppler times the PRI. The $3M$ -by- $3M$ ADPCA covariance matrix estimate, $\hat{\mathbf{R}}_{\text{adpca}/k}$, is an approximation to

$$\mathbf{R}_{\text{adpca}/k} = E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^H], \quad (12.132)$$

where $\tilde{\mathbf{x}}_k$ is given in (12.130). The covariance inverse in the ADPCA weight vector provides a dynamic response to whiten ground clutter returns. The steering vector term in parentheses suppresses mainlobe clutter with the binomial weights, identified as $[1 \ -2 \ 1]^T$, while forming a beam in a specified angular direction; the steering vector incorporates an additional linear phase variation over the temporal pulses to steer the Doppler null in cases where clutter is not centered at 0 Hz.

The performance of ADPCA is generally not as good as post-Doppler STAP techniques, thus limiting its use in practice. Performance benchmark results are given in Section 2.12.5.

2.12.4.3 Parametric adaptive matched filter

The parametric adaptive matched filtering (PAMF) employs multichannel, linear prediction to estimate the covariance matrix inverse [41]. We now discuss the PAMF based on [41,42].

Consider the LDU decomposition of the covariance matrix inverse,

$$\mathbf{R}_k^{-1} = (\mathbf{A}_k^{-1})^H \mathbf{D}_k^{-1} \mathbf{A}_k^{-1}, \quad (12.133)$$

where $\mathbf{A}_k = \mathbf{L}$, $\mathbf{A}_k^H = \mathbf{U}$, and \mathbf{D}_k is diagonal. Since we are considering space-time data, a block LDU decomposition is required. In this case, \mathbf{A}_k^{-1} takes the form, for the example of $M = 4$ and $N = 8$,

$$\mathbf{A}_k^{-1} = \begin{bmatrix} \mathbf{I}_4 & \mathbf{0} \\ \mathbf{A}_1^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_3^H(3) & \mathbf{A}_3^H(2) & \mathbf{A}_3^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_4^H(4) & \mathbf{A}_4^H(3) & \mathbf{A}_4^H(2) & \mathbf{A}_4^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_5^H(5) & \mathbf{A}_5^H(4) & \mathbf{A}_5^H(3) & \mathbf{A}_5^H(2) & \mathbf{A}_5^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_6^H(6) & \mathbf{A}_6^H(5) & \mathbf{A}_6^H(4) & \mathbf{A}_6^H(3) & \mathbf{A}_6^H(2) & \mathbf{A}_6^H(1) & \mathbf{I}_4 & \mathbf{0} \\ \mathbf{A}_7^H(7) & \mathbf{A}_7^H(6) & \mathbf{A}_7^H(5) & \mathbf{A}_7^H(4) & \mathbf{A}_7^H(3) & \mathbf{A}_7^H(2) & \mathbf{A}_7^H(1) & \mathbf{I}_4 \end{bmatrix}, \quad (12.134)$$

where $\mathbf{A}_p^H(n)$ are the 4-by-4 multi-channel linear prediction (MCLP) coefficients for the p -order filter and the n th lag, $\mathbf{0}$ is the 4-by-4 zero matrix, and the realization index, k , is dropped for notational

convenience [42]. The matrix \mathbf{D}_k^{-1} is

$$\mathbf{D}_k^{-1} = \begin{bmatrix} \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_5 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_6 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_7 \end{bmatrix}, \quad (12.135)$$

where \mathbf{D}_p is the 4-by-4 inverse of the MCLP error for the p -order filter. As noted in [42], the MCLP development, while not necessary, is useful and provides a framework to calculate $\mathbf{A}_p^H(n)$ and \mathbf{D}_p .

The PAMF uses the LDU decomposition to approximate \mathbf{R}_k^{-1} . For example, let the maximum MCLP filter order $P = 2$, then

$$\mathbf{A}_k^{-1} \approx \begin{bmatrix} \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_2^H(2) & \mathbf{A}_2^H(1) & \mathbf{I}_4 \end{bmatrix}. \quad (12.136)$$

Extending the Wiener-Hopf equation to calculate the linear prediction coefficients, $\mathbf{A}_p^H(n)$, and augmenting with the identity matrices shown in (12.136), gives an approximation to \mathbf{A}_k^{-1} . The linear prediction error terms are given as

$$\mathbf{D}_k^{-1} \approx \begin{bmatrix} \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_2 \end{bmatrix}. \quad (12.137)$$

Generally, each of the diagonal blocks is inverted after solving for $\mathbf{D}_k = \mathbf{A}_k^{-1} \mathbf{R}_k (\mathbf{A}_k^{-1})^H$. The approximation to the covariance matrix inverse follows from (12.133).

2.12.4.4 Summary of adaptive filter implementations

Table 12.3 summarizes the adaptive filter implementations discussed in this section. Performance assessment of a number of these methods is given in Section 2.12.5.2.

Table 12.3 Summary of Adaptive Filter Implementations

Type	Method	Section	Comments
RR-STAP	Adaptive beam pattern synthesis (principal components)	2.12.4.1.1	Weight calculation approach, interprets STAP as pattern synthesis, where notches are incorporated into the quiescent beam pattern based on information in the dominant subspace. Insightful view of STAP, not as practical as other methods, especially due to high computational burden associated with the eigendecomposition
RR-STAP	Principal Components Inverse (PCI)	2.12.4.1.2	Weight calculation approach, involves forming a projection operation orthogonal to the clutter and interference subspaces. May be used with RD-STAP methods. Computationally burdensome due to eigendecomposition
RR-STAP	Eigencanceler	2.12.4.1.3	Two formulations, one leading to Principal Components Inverse weight vector solution
RR-STAP	Hung-Turner Projection (HTP)	2.12.4.1.4	Uses Gram-Schmidt to approximate PCI solution at much lower computational cost. Developed for use with large adaptive DoF systems. Best performance against stronger interference sources
Other	Diagonal loading	2.12.4.1.5	Adds scaled diagonal matrix to covariance matrix estimate, enhancing numerical stability and mitigating spurious sidelobe effects due to perturbed noise subspace
RR-STAP	Cross Spectral Metric (CSM)	2.12.4.1.6	Signal-dependent approach to select interference subspaces in weight calculation most greatly enhancing output SINR. Generally limited to the GSCLC architecture, which tends to greatly increase computational burden and limits practical application
Other	Multi-Stage Wiener Filter (MWF)	2.12.4.1.7	Truncated decomposition of the GSCLC architecture, involves calculating a series of scalar weights for each stage, with each subsequent stage maximally correlated with the “desired signal” (residual interference) of the prior stage. Very high computational burden relative to competing methods

Continued

Table 12.3 *Continued*

Type	Method	Section	Comments
RD-STAP	Extended Factored Algorithm (EFA)	2.12.4.2.1	Post-Doppler, “element” space STAP. Typically involves adaptive weighting three to five adjacent Doppler bins and all spatial channels to suppress clutter. Efficient, effective technique: reasonable computational burden and low training sample support requirements
RD-STAP	Factored Time-Space (FTS)	2.12.4.2.1	Special case of EFA when using only one Doppler bin. Places a spatial null within a given Doppler bin. Generally performs poorly at low range rates, since it only has spatially adaptive DoFs
RD-STAP	Joint-Domain Localized (JDL)	2.12.4.2.1	Post-Doppler, beamspace STAP. Essentially a two-dimensional, linear prediction technique: employ angle-Doppler beams surrounding the target beam to estimate and coherently remove competing interference. Low computational burden and minimal training sample support requirements
RD-STAP	Adaptive Displaced Center Antenna (ADPCA)	2.12.4.2.2	Pre-Doppler STAP method, involves adaptively combining all spatial channels and a small collection of pulses from within the current CPI. Adaptive canceler is applied to overlapped segments of pulses from the CPI. Collection of clutter canceled outputs then fed through a Doppler filter bank to achieve temporal integration gain. Useful for situations where interference may appear slightly nonstationary over the CPI, such as in the case with a rotating array. Computational burden and training sample requirements similar to EFA
Other	Parametric Adaptive Matched Filter (PAMF)	2.12.4.3	Low order, multi-channel linear prediction approach to estimate the covariance matrix inverse. Computationally efficient. Model order selection is generally ad hoc

Table 12.4 Radar Simulation Parameters

Parameter	Value	Parameter	Value
Transmit frequency	L-Band (1240 MHz)	Aperture size	~1.2 m width × 1 m height
Receiver bandwidth	0.8 MHz	Receive aperture configuration	11 channels configured as a uniform linear array
Peak power	15 kW	Array attitude	Sidelooking, canted 5° down, 6.78° yaw, 0° roll
PRF	1984 Hz	Platform altitude	3000 m
CPI length	32 pulses	Platform velocity	100 m/s
Clutter model	–20 dB reflectivity, Constant Gamma Model	ICM	Gaussian, 0.25 m ² /s ²

2.12.5 Application

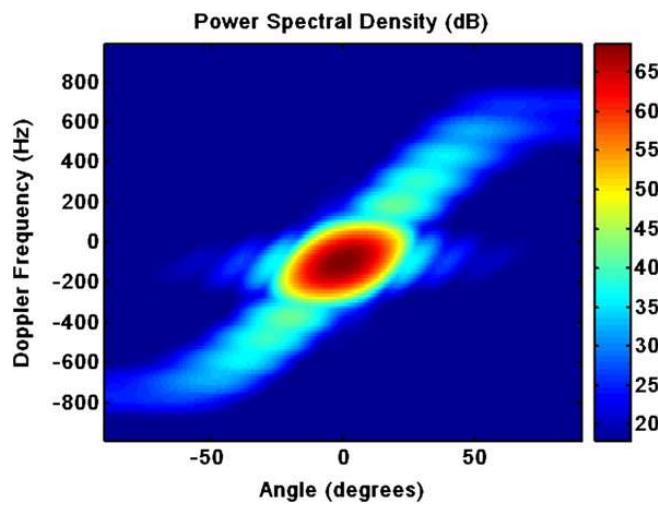
In this section we consider the application of STAP to the important problem of radar detection in clutter. We employ the simulation models of Section 2.12.3 and some of the metrics from Section 2.12.2 to characterize the interference, algorithm performance, and STAP improvement over the non-adaptive processor.

Table 12.4 provides some of the parameters of the simulated, multi-channel radar system. Chosen parameters are similar to those of the Multi-Channel Airborne Radar Measurements (MCARM) system described in Section 2.12.8. Also, these are the same parameters used to generate Figures 12.1–12.3 in Section 2.12.1.

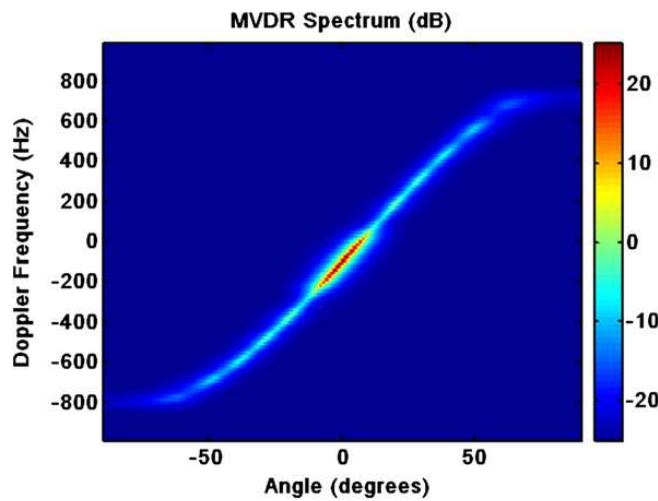
2.12.5.1 Interference characteristics

Using the models described in Section 2.12.3, we simulate clutter and noise for an eleven channel, thirty-two pulse airborne radar scenario. The clutter environment is homogeneous. Additional parameters are given in Table 12.4 and are similar in nature to a MCARM data collection referenced in Section 2.12.8. In this case, we simulated one CPI, where the aircraft experiences 6.78° of yaw. We focus our attention somewhat arbitrarily at a 32 km slant range for the analysis. Additionally, we simulated the eleven channel uniform linear array without array amplitude and phase errors. The antenna is steered to broad side: 0° azimuth and 0° elevation angle.

Figure 12.17 shows the PSD for this scenario. The PSD uses a Hanning weight in both space and time. The slight offset in peak clutter Doppler frequency in the look-direction from 0 Hz is due to platform yaw. The overall CNR is in the range of slightly greater than 50 dB. For comparison, Figure 12.18 shows the MVDR spectrum, clearly outlining the clutter angle-Doppler region of support. (Note: while MVDR provides units of power, it is widely known to be a poor estimator of the signal strength.) Figure 12.19 shows the eigenspectrum calculated using the known covariance matrix; the eigenspectrum displays the

**FIGURE 12.17**

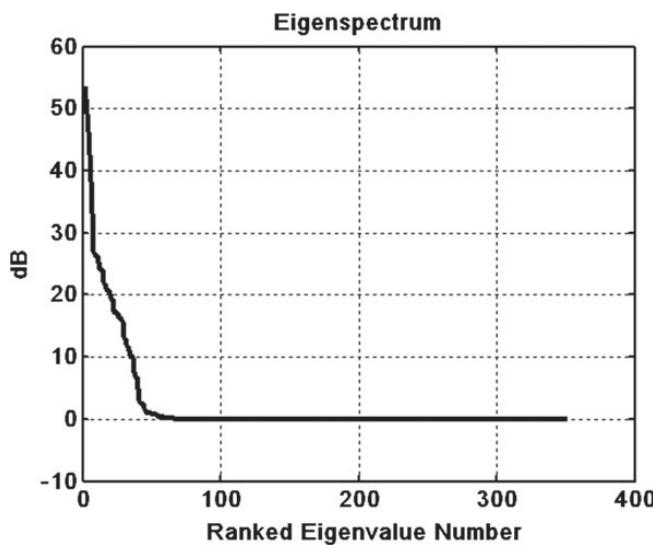
Power spectral density for sidelooking array radar example.

**FIGURE 12.18**

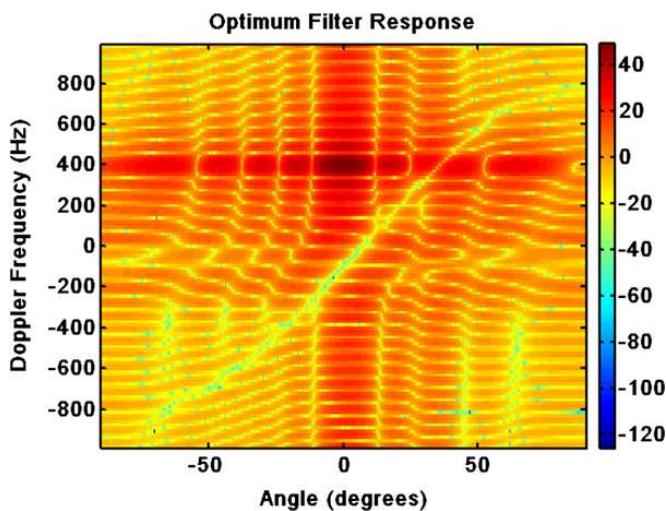
MVDR spectrum for sidelooking array radar example.

covariance matrix eigenvalues, sorted from largest to smallest. The noise eigenvalues are set to 0 dB, and those values greater than the noise floor correspond to the clutter signal. The largest eigenvalue is indicative of the CNR to within a few decibels, suggesting CNR in the range of 53–55 dB.

Considering the MVDR spectrum in Figure 12.18, we anticipate an optimal filter response that drives a deep null along the clutter angle-Doppler region of support. Figure 12.20 shows the optimal space-time filter response, steered to broadside and 400 Hz Doppler; the clutter null and peak gain is evident in the figure.

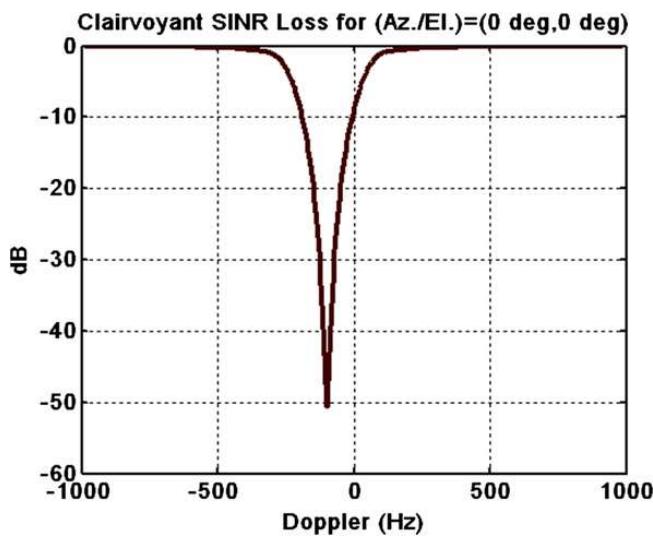
**FIGURE 12.19**

Eigenspectrum for example scenario.

**FIGURE 12.20**

Optimal space-time frequency response, steered to array broadside and 400 Hz Doppler frequency.

Finally, Figure 12.21 shows the clairvoyant (known covariance) SINR loss, $L_{s,1}$, over unambiguous Doppler ($-\text{PRF}/2 \leq f_d \leq \text{PRF}/2$) in the broadside direction. The loss cut shows the impact of clutter on detection performance, with the null offset from 0 Hz due to the platform yaw. As seen from the figure, the null depth is slightly greater than 50 dB, consistent with the mainlobe CNR and the eigenspectrum of Figure 12.19. We will consider the performance of various STAP techniques relative to the clairvoyant SINR loss curve momentarily.

**FIGURE 12.21**

Clairvoyant SINR loss.

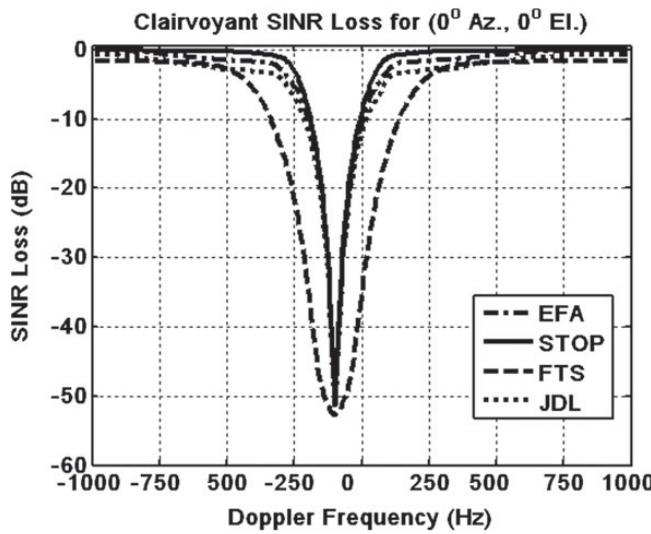
2.12.5.2 STAP algorithm performance

Based on the discussion in Section 2.12.4, we benchmark the performance of a number of STAP techniques using the scenario from the prior section. Again, we focus on a slant range of 32 km. Also, we consider clairvoyant loss, $L_{s,1}$. For the homogeneous clutter scenario considered herein, the adaptive loss, $L_{s,2}$, precisely follows the RMB rule: using sample support equal to twice the processor's DoF yields, on average, 3 dB loss. The different techniques have different sample support requirements, based on their adaptive DoFs. An exception in this case involves the PAMF; the PAMF implementation averages over the pulse domain, and so can rely on less training data over range. However, this advantage is not unique to the PAMF, and other methods can also average over pulses (like the ADPCA technique) or employ smoothing techniques to effectively increase the sample support [43].

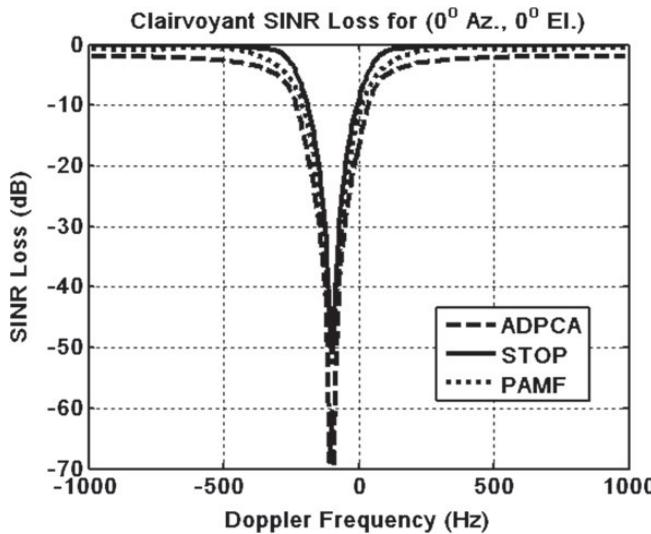
Figure 12.22 compares the clairvoyant SINR loss of EFA, FTS, and JDL to the bound provided by the space-time optimal processor (STOP; the optimal space-time filter using the clairvoyant covariance matrix). In this case, the processors are configured as follows:

- *FTS*—uses all eleven channels and a Hanning weighting on the Doppler filters. Eleven adaptive spatial DoFs.
- *EFA*—uses all eleven spatial channels and three adjacent, Hanning-weighted Doppler filters. Thirty-three adaptive space-Doppler DoFs.
- *JDL*—uses three adjacent, uniform weighted spatial beams and three adjacent, Hanning-weighted Doppler filters. Nine total adaptive angle-Doppler DoFs. The beam spacing is three degrees (a little less than one-third of the full aperture 3 dB beamwidth).

From Figure 12.22, we find that EFA and JDL provide excellent performance potential relative to the achievable bound set by STOP. The FTS performance is disappointing, but not a surprise since this algorithm does not adaptively combine any temporal DoFs. One then might conclude JDL to be a better selection, since it only requires nine DoFs. However, the additional spatial DoFs afforded by EFA may

**FIGURE 12.22**

Clairvoyant SINR loss for post-Doppler STAP techniques.

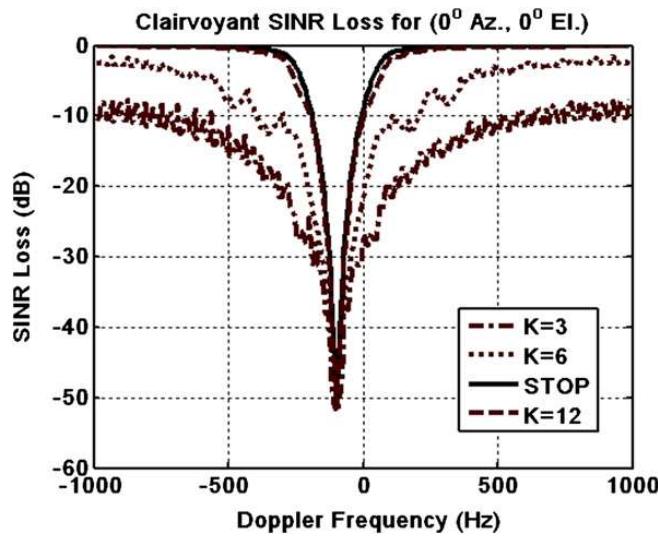
**FIGURE 12.23**

Clairvoyant SINR loss for pre-Doppler STAP and PAMF.

prove beneficial when RFI or Doppler ambiguities are present. Thus, a number of analyses should take place before final algorithm selection.

Figure 12.23 duplicates the aforementioned analysis for pre-Doppler STAP (ADPCA) and the PAMF. The processors are configured as follows:

- *ADPCA*—uses three adjacent pulses and all spatial channels for cancellation, then Doppler filters the resulting outputs using a Hanning weight. Thirty-three adaptive space-time DoFs. The

**FIGURE 12.24**

Clairvoyant SINR loss for the MWF.

implementation steers the temporal gain response to center the temporal steering vector null at -100 Hz, the center of mainlobe clutter.

- *PAMF*—uses a fourth-order filter model. No weighting is used on the Doppler steering vector.

The clairvoyant SINR loss curves in Figure 12.23 indicate that both ADPCA and PAMF provide very good performance, rivaling that of EFA and JDL in Figure 12.22.

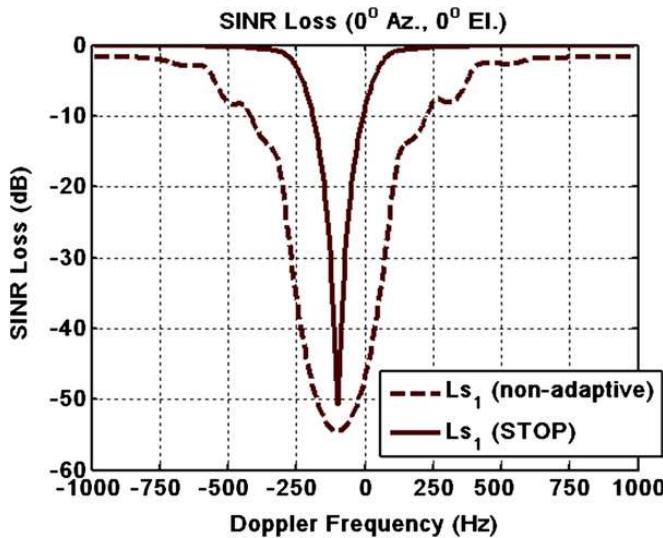
As a last example, Figure 12.24 shows the benchmark performance of the MWF using three, six, and twelve stages ($K = 3, 6, 12$). As seen from the figure, the performance for $K = 3$ or $K = 6$ is poor, but very good capability is observed when $K = 12$. The computational loading of MWF, for the implementation used by the author, is significantly higher than the other methods examined.

2.12.5.3 STAP Comparison with nonadaptive solution

To conclude this section, we simply compare the performance of the space-time optimal filter to a nonadaptive processing scheme involving beamforming and Doppler processing. The nonadaptive filter implementation uses a Hanning temporal weight and uniform illumination spatially on receive. (Note: the transmit illumination uses a 30 dB Taylor weighting.) As seen from Figure 12.25, the space-time optimal filter provides a very significant performance advantage over the nonadaptive processor, thereby indicating the tremendous benefits of STAP.

2.12.5.4 Application summary

In this section we considered a typical airborne radar example. Clutter and noise models from Section 2.12.3 were used to simulate data snapshots and covariance matrices. The covariance matrices were used to assess the performance potential of several STAP techniques discussed in Section 2.12.4. From this analysis, we find that a number of the methods perform similarly well relative to the achievable

**FIGURE 12.25**

Comparison of the optimal and nonadaptive filter performance potential.

performance bound. We caution the reader, however, that many other practical matters drive algorithm selection. Of the methods discussed, EFA often provides the best performance and greatest flexibility when considering a number of issues.

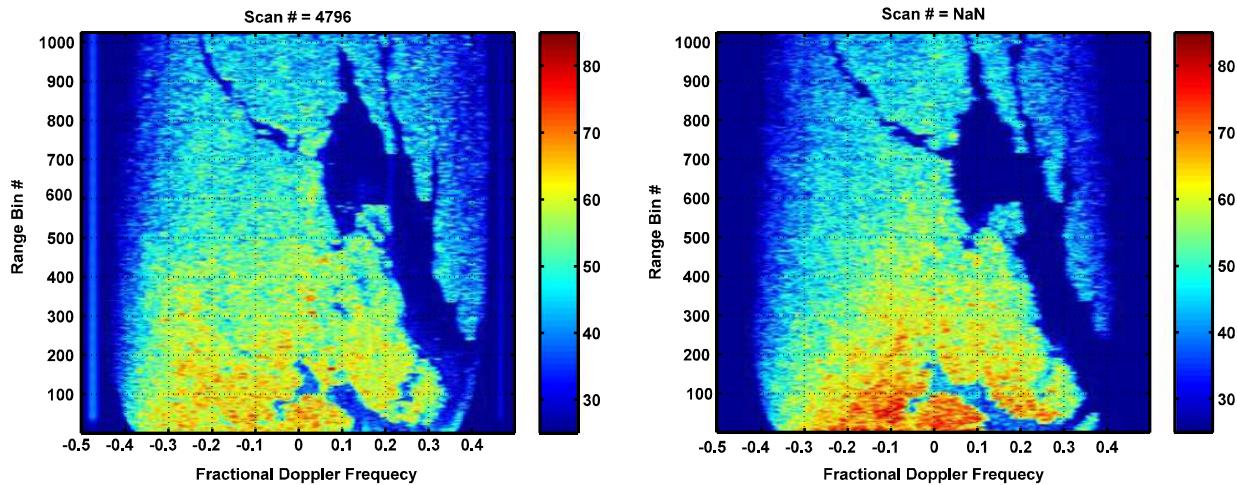
2.12.6 Challenges

Contemporary STAP research topics generally focus on challenges associated with covariance matrix estimation or mitigating computational burden. In this section we provide commentary on the former.

2.12.6.1 Heterogeneous clutter

Real world clutter environments are heterogeneous, thus tacitly undermining the IID assumption central to the development of (12.49) [44–46]. Heterogeneous clutter is a result of culturally-varying terrain type. Examples include: variation in clutter amplitude or spectral spread due to a mixture of clutter types; abrupt edges characteristic of clutter interfaces (e.g., between urban and rural regions); the presence of target-like signals in the training data (TSD) [45]; and, stationary, manmade objects resulting in strong, discrete responses. Each of the aforementioned effects is localized within the training set. Hence, the estimate of (12.49) captures the average behavior of the training data, thus potentially appearing mismatched to any particular cell under test. This mismatch translates to an erroneous, adaptive response relative to the optimal condition. SINR loss or threshold bias results in such instances. References [44–46] characterize in detail the nature of such system degradation.

As an example, distributed clutter is typically site-specific: the clutter reflectivity varies as a function of range and angle due to changes in the terrain features. Equation (12.70) accommodates site-specific

**FIGURE 12.26**

Comparison of measured (left) and simulated (right) multichannel UHF radar data (colorbar in decibels) (after [47], © 2006 IEEE).

simulation by modifying the term, $v_k(p, q)$, through access of a database to determine the clutter RCS, $\sigma_{k,p,q}$, in (12.67)–(12.69) (where the reference to the q th range ambiguity is added). Figure 12.26 compares actual and site-specific simulation of UHF radar data taken in the Delaware-Maryland-Virginia (Delmarva) Peninsula region; this region is dominated by rural clutter, bodies of water and rivers, and also covered by a number of roadways. As seen from Figure 12.26, the range and cross-range (Doppler, or angle) variation of the clutter response is evident and predictable. The clutter variation shown is one source of heterogeneity leading to covariance matrix estimation error.

Table 12.5 provides a summary of various sources of heterogeneous clutter and their impact on STAP performance. Of the effects listed, it is observed that target-like signals corrupting the secondary (training) data (TSD) and clutter discretes tend to lead to the greatest performance loss.

A variety of techniques have been developed to enhance STAP detection performance given the challenges of heterogeneous clutter. We summarize some of the available methods in Table 12.6. As discussed in [47], the different approaches fall in one of two categories: indirect or direct. The indirect methods attempt to manipulate the training set to improve the covariance matrix estimate, oftentimes using knowledge of the platform motion or surrounding terrain or operating environment, whereas the direct methods attempt to modify the filter response using models and ancillary information. Both indirect and direct methods attempt to improve the instantaneous adaptive filter response to minimize heterogeneous clutter residue at the filter output; in this manner, STAP performance in complex, heterogeneous environments can approach that attainable in a homogeneous setting.

As an example of the degradation of heterogeneous interference effects, and the improvement potential of available STAP techniques, Figure 12.27 compares detection rates using MCARM Flight 5, Acquisition 575 data using block training (bins 200–320) versus intelligent training and filter selection (ITFS, [47]), where maps facilitate training data excision in regions overlaying certain roadways. As seen from the figure, TSD leads to excessive, additional loss in the lobed regions near ± 20 m/s;

Table 12.5 Summary of Heterogeneous Clutter and STAP Impact

Source	Description	Impact on STAP
Spatially-varying, distributed clutter RCS	Range-angle variation in clutter RCS leads to varying power across clutter angle-Doppler region of support	Over- or undernulled clutter. Overnullled clutter can lead to signal cancellation, whereas undernulling leads to clutter residue and degraded SINR
Spatially-varying, distributed clutter spectral spread	Range-angle variation of intrinsic clutter motion varies spectral spread across clutter angle-Doppler region of support	Training on regions where ICM is less than that of the application region leads to insufficient null width and an increase in clutter residual, whereas regions with ICM less than that present in the training set experience suppression of low speed targets
Clutter edges and shadowing	Gross variation in clutter type (e.g., land-sea interface) or regions of extended obscuration, predominantly in the mainlobe direction	Over- or undernulled clutter
Clutter discretes	Stationary objects with relatively large RCS, often manmade objects such as cars, utility poles, etc., also includes extended discretes, such as fence lines and train tracks	Increased false alarm rate, upward threshold bias leading to increase in missed detections
Target-like signals	Vehicles on roadways and within airspace, predominantly through the mainbeam	Signal cancellation due to nulling off the clutter ridge in angle-Doppler locations consistent with targets of interest

removing a specific highway using ITFS—as noted in the figure’s legend—fully mitigates this loss, raising the detection rate from roughly 55% to 98%.

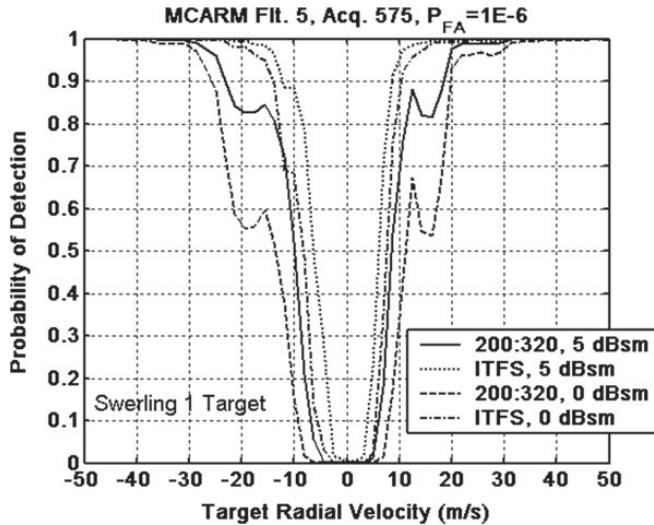
2.12.6.2 Nonstationary clutter

STAP maximizes SINR by filtering ground clutter in the angle-Doppler domain. Radar geometry determines the filter null location in this higher-dimensional space. When the null location varies over range, the adaptive filter produces an incorrect frequency response. In the monostatic radar case, the angle-Doppler region of support can exhibit range variation when the velocity vector and array normal are non-orthogonal. This occurs for forward-looking arrays, or under conditions when the platform is yawed [48,49].

In sidelooking array radar, the spatial frequency measured by a uniform linear array is proportional to the corresponding Doppler frequency of a stationary clutter patch on the Earth’s surface. Specifically,

Table 12.6 STAP Techniques to Mitigate the Impact of Heterogeneous Clutter on Detection Performance

Technique	Description	Reference
Nonhomogeneity Detector (NHD)	Data-dependent screening measure, tests training data for similarity to the average, selects homogeneous training samples	[50–52]
Power Selected Training (PST)	Data-dependent screening measure, selects training samples strongest in power to drive null as deeply as possible	[53, 54]
Power Comparable Training (PCT)	Data-dependent screening measure, sorts data into tiles of similar power levels to match training data to cell under test	[55]
Power Variable Training (PVT)	Scale the power of the estimated covariance matrix's dominant (clutter) subspace to match the power level in the cell under test, requires eigendecomposition	[56]
Map-aided training	Uses mapping data and distance measures to screen training samples	[24, 57]
Covariance matrix taper (CMT)	Purposely spreads the estimated clutter response to increase null width	[58]
Subaperture smoothing	Exploits similar covariance structure among sub-apertures of a uniformly sampled spatial or temporal aperture to enhance the quantity of homogeneous training samples	[43, 59]
Nonlinear, nonadaptive STAP	Employs a model of the clutter covariance matrix to suppress clutter	[60]
Color loading	Adds a scaled clutter covariance model to the covariance matrix estimate to emphasize the anticipated clutter response and enable localized training	[61, 62]
Discrete Matched Filter (DMF)	Coherently removes clutter discrete signals using a modified version of the CLEAN algorithm	[47]
Signal and Clutter as Highly Independent Structured Modes (SCHISM)	Coherently removes distributed clutter returns using a modified version of the CLEAN algorithm	[63]
Adaptive Coherence Estimator (ACE)	Incorporates a statistical measure of the “whiteness” of the clutter residue in a particular cell under test to suppress clutter discretes?	[64]
Knowledge-Aided Parametric Covariance Estimation (KAPE)	Estimates parameters of a validated covariance matrix model, including clutter amplitude, spread, mainlobe centroid, and error components of the array manifold	[16, 65]

**FIGURE 12.27**

Estimated detection probability improvement over Doppler frequency using KA training on MCARM data (after [24], © 2004 IEEE).

normalized Doppler, $\tilde{f}_d = f_d T$, can be written

$$\tilde{f}_d = \frac{2v_p T}{\lambda} \cos \phi_{\text{cone}} = \left(\frac{v_p T}{d/2} \right) f_s, \quad (12.138)$$

where v_p is the platform velocity in the direction orthogonal to the array normal, T is PRI, λ is wavelength, ϕ_{cone} is the cone angle measurement from the platform centerline to the clutter patch location, d is channel spacing, and f_s is spatial frequency. Spatial frequency is given by

$$f_s = \frac{d}{\lambda} \cos \phi_{\text{cone}}. \quad (12.139)$$

Physically, the clutter iso-Doppler contours and the array beam traces align, translating to angle-Doppler behavior that is range invariant.

In the non-sidelooking radar case, the beam traces and iso-Dopplers of the ground clutter returns misalign at some ranges, mainly those where the range divided by the altitude is less than five [18, 19]. The variation of the clutter response through the training set leads to an adaptive filter response with incorrect null placement, leading to increased clutter residue and degraded detection performance. References [48, 49] discuss effective compensation methods.

Range varying angle-Doppler loci further exacerbates culturally induced heterogeneous clutter effects. For example, residue from clutter discretes further increase in the presence of adaptive filter null migration associated with the non-stationary clutter mechanism described in this section.

2.12.6.3 Bistatic STAP

Bistatic STAP development received considerable attention in the early-to-mid 2000s. Airborne bistatic radar involves a moving transmitter and receiver separated a considerable distance. The bistatic geometry

Table 12.7 Summary of Bistatic STAP Methods

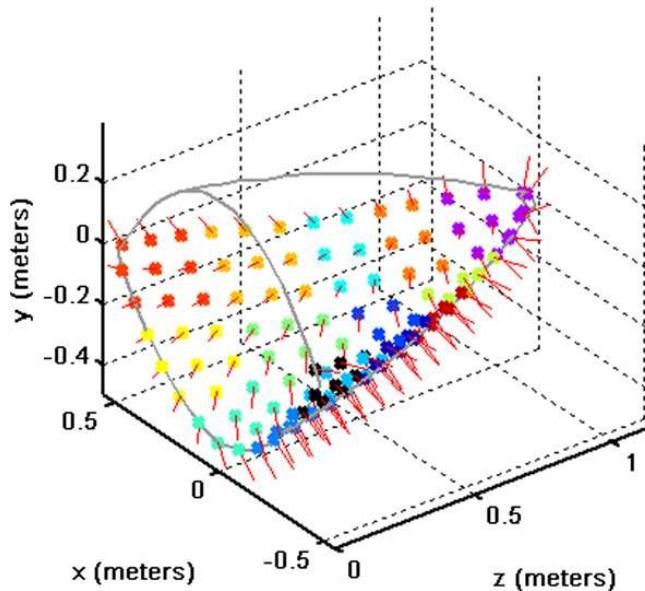
Bistatic STAP approach	Description	Reference
Localized processing	Simple strategy, attempts to choose training data in the vicinity of the cell under test to minimize nonstationary impacts	[66, 67]
Time-varying weights	Employs truncated Taylor series expansion of the weight vector, presumes linear evolution in the weight vector elements over range	[68, 69]
Doppler warping	Aligns a point of the clutter ridge for each bistatic range to a designated reference point (e.g., 0 Hz Doppler) using a range-varying, complex modulation prior to STAP application	[49]
Higher-order Doppler warping	Aligns sections of the clutter ridge to a reference ridge using a range varying modulation, a multi-point extension of the Doppler warping method prior to STAP application	[70]
Angle-Doppler compensation	Aligns “spectral centers,” or regions of maximum angle-Doppler return, over range using a complex, range-varying, space-time modulation prior to STAP application	[71]
Adaptive angle-Doppler compensation	Derives key information on clutter range variation directly from the data, applies a complex, space-time, range-varying modulation to the data to align dominant clutter subspaces prior to STAP application	[72, 73]
Registration of “direction-Doppler curves”	Uses curve fitting methods to warp the power spectral density of a given range sum to a reference prior to STAP application	[74]

and independent motion between transmitter and receiver leads to significant angle-Doppler variation over range [75, 72].

Table 12.7 summarizes a number of recently developed bistatic STAP techniques, including a description of the approach and germane references. As seen from the table, the methods generally fall in one of three categories: localized processing (or training); data pre-warping, prior to STAP application; and, time-varying weights. Several of the bistatic STAP methods in Table 12.7 have proven extraordinarily effective, restoring performance to levels similar to that achievable in homogeneous clutter environments.

2.12.6.4 Conformal array STAP

A conformal array’s shape is compliant with the contours of the radar bearing platform. For example, Figure 12.28 depicts conformal antenna elements mounted to a chined nose cone; the short lines emanating from each dot representing an antenna element indicate the surface normal.

**FIGURE 12.28**

Example of conformal antenna elements mounted to a chined nose cone.

The curvature of the conformal array, coupled with the look-direction relative to the velocity vector, results in a nonstationary clutter angle-Doppler response [76–78]. SINR loss resulting from the nonstationary clutter response rivals that seen in bistatic radar examples. Hersey et al. [76] discusses several solutions to mitigate nonstationary clutter behavior based on adaptations of solutions developed for bistatic STAP implementation, including localized processing, localized processing with time-varying weights, angle-Doppler warping, and higher-order angle-Doppler warping. Additionally, Hersey et al. [76] develops a method called equivalent uniform linear array transformation which attempts to resample the data to a linear configuration. The conformal array STAP methods in [76] lead to significant performance enhancement.

2.12.7 Implementation

We discuss a generic, STAP-based, airborne radar detection architecture in this section. A flow diagram of the detection process is given in Figure 12.29.

STAP is applied at the CPI level. As a result of target fading, the MTI radar typically transmits bursts of pulses at several different frequencies, where each burst comprises the CPI. Being a coherent signal processing technique, STAP operates on each CPI as indicated in the figure, essentially generating a range-Doppler map (RDM) at a receive angle consistent with the transmit direction: a space-time snapshot for each range bin, \mathbf{x}_k , is processed using a space-time weighting steered to a given angle and Doppler to yield a single pixel in the RDM, $|y_k|^2$. Each CPI is processed, generating aligned RDMs; the alignment may require additional processing steps, such as motion compensation. Then, the RDMs

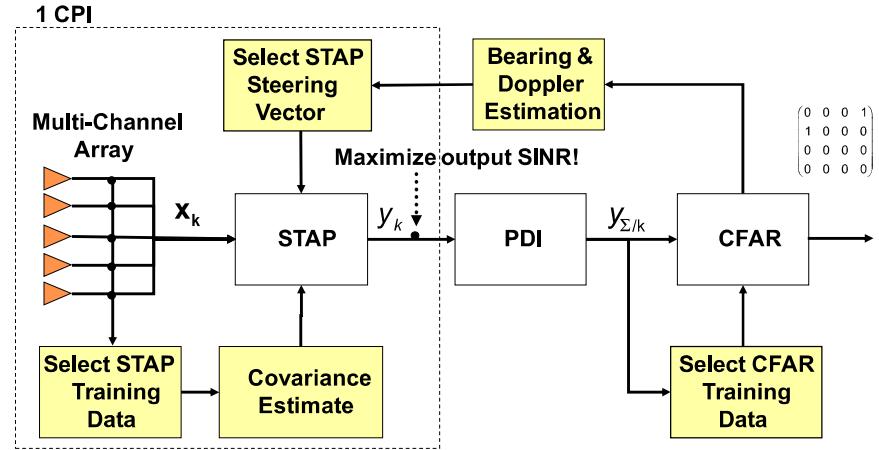


FIGURE 12.29

Detection processing flow.

at the different frequencies are noncoherently summed in a step called post-detection integration (PDI; also called noncoherent addition, NCA). PDI boosts detection performance by taking advantage of the purposeful decorrelation created by frequency hopping to avoid target fading that adversely affects operation using a single coherent dwell. Frequency hopping forces the target to take on a Swerling 2 characteristic [1, 7].

After PDI, a detection threshold is set using a constant false alarm rate (CFAR) algorithm; the CFAR algorithm multiplies an estimate of the residual interference power by a threshold multiplier in an attempt at achieving a constant false alarm rate [1]. Pixel values crossing the detection threshold are declared targets. The signal processor then estimates the bearing and Doppler frequency of the threshold crossing generally using a maximum likelihood estimator. The corresponding SINR, bearing, Doppler, and range—among other potential characteristics—for each detection is then passed to an analyst or an automatic tracker. Figure 12.29 is meant to show the basic processing steps; it is common to incorporate additional functionality to cope with clutter heterogeneity or target motion through resolution cells, for instance.

The selection of the scalar, $\hat{\mu}$, in (12.44) is an important, practical consideration. It is common to use the adaptive matched filter (AMF) normalization [79],

$$\hat{\mu} = \frac{1}{\sqrt{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)}}. \quad (12.140)$$

The AMF normalization sets the residual interference-plus-noise floor to unity as $\hat{\mathbf{R}}_k \rightarrow \mathbf{R}_k$, viz.

$$P_{I+N} = \hat{\mathbf{w}}_k^H \mathbf{R}_k \hat{\mathbf{w}}_k = \frac{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{R}_k \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)}{\left(\sqrt{\mathbf{v}_{s-t}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_k^{-1} \mathbf{v}_{s-t}(\phi, \theta, f_d)} \right)^2} \rightarrow 1. \quad (12.141)$$

The AMF normalization accommodates variable training intervals and other segmentation of the weight estimation and application over range, minimizing shifts in the mean, residual interference-plus-noise

power that otherwise leads to threshold bias and increased false alarm rate. More specifically, assuming the segmented interference-plus-noise environment is IID, the AMF normalization possesses CFAR properties: as power fluctuates from one weight application region to the next, the filter output is scaled by the inverse of the residual interference-plus-noise power so that a fixed detection threshold maintains a constant false alarm rate. This is seen by examining the AMF decision statistic,

$$\eta_{\text{AMF}} = \frac{\left| \mathbf{v}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}} \right|^2}{\mathbf{v}_{\mathbf{s}-\mathbf{t}}^H(\phi, \theta, f_d) \hat{\mathbf{R}}_{\mathbf{k}}^{-1} \mathbf{v}_{\mathbf{s}-\mathbf{t}}(\phi, \theta, f_d)} \stackrel{H_1}{>} T_1, \stackrel{H_0}{<} T_1, \quad (12.142)$$

where T_1 is a fixed decision threshold. Moving the denominator of the decision statistic to the right, (12.142) is interpreted as the square of the filter output (the numerator term) compared to a threshold multiplier, T_1 , scaled by the residual interference-plus-noise output power. This is the conventional view of CFAR in radar signal processing.

A virtually identical formulation of the AMF result is given in [80] and referred to as Modified Sample Matrix Inversion (MSMI).

A maximum likelihood estimate for target bearing and Doppler integrates seamlessly with the architecture in Figure 12.29 and use of the AMF normalization. This can be seen by expressing the alternative hypothesis of (12.2) as

$$\mathbf{x}_{\mathbf{k}} = \alpha \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p}) + \mathbf{n}_{\text{tot}}; \mathbf{p} = \begin{bmatrix} \phi \\ \theta \\ f_d \end{bmatrix}, \quad (12.143)$$

where α is a complex constant, and $\mathbf{n}_{\text{tot}} = \mathbf{c}_{\mathbf{k}} + \mathbf{j}_{\mathbf{k}} + \mathbf{n}_{\mathbf{k}}$ is the interference-plus-noise (total noise) vector. The objective is to estimate \mathbf{p} with $\hat{\mathbf{p}}$. Since $\mathbf{n}_{\text{tot}} \sim \text{CN}(\mathbf{0}, \mathbf{R}_{\mathbf{k}})$, the joint probability density function (pdf) is

$$p(\mathbf{x}_{\mathbf{k}} | \mathbf{p}) = \frac{1}{\pi^M |\mathbf{R}_{\mathbf{k}}|} \exp \left(-(\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p}))^H \mathbf{R}_{\mathbf{k}}^{-1} (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p})) \right). \quad (12.144)$$

Given the likelihood function defined by (12.144), we first require an estimate for the complex constant. Equation (12.144) is maximal when

$$Q(\alpha, \mathbf{p}) = (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p}))^H \mathbf{R}_{\mathbf{k}}^{-1} (\mathbf{x}_{\mathbf{k}} - \alpha \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p})) \quad (12.145)$$

is minimal. Differentiating (12.145) and setting the result to zero yields

$$\hat{\alpha} = \frac{\mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}}}{\mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p})}. \quad (12.146)$$

Substituting (12.146) into (12.145) leads to

$$Q(\hat{\alpha}, \mathbf{p}) = \mathbf{x}_{\mathbf{k}}^H \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{x}_{\mathbf{k}} - \frac{\left| \mathbf{x}_{\mathbf{k}}^H \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p}) \right|^2}{\mathbf{s}_{\mathbf{s}-\mathbf{t}}^H(\mathbf{p}) \mathbf{R}_{\mathbf{k}}^{-1} \mathbf{s}_{\mathbf{s}-\mathbf{t}}(\mathbf{p})}. \quad (12.147)$$

Differentiating (12.147) with respect to \mathbf{p} and setting the result to zero yields the maximum likelihood estimate (MLE), $\hat{\mathbf{p}}$. Consequently, the estimator is [81]

$$\hat{\mathbf{p}} = \begin{bmatrix} \hat{\phi} \\ \hat{\theta} \\ \hat{f}_d \end{bmatrix} = \arg \max_{\mathbf{p}=[\phi \ \theta \ f_d]^T} \left(\frac{\left| \mathbf{s}_{s-t}^H(\mathbf{p}) \mathbf{R}_k^{-1} \mathbf{x}_k \right|^2}{\mathbf{s}_{s-t}^H(\mathbf{p}) \mathbf{R}_k^{-1} \mathbf{s}_{s-t}(\mathbf{p})} \right). \quad (12.148)$$

The MLE takes the form of an AMF normalized STAP cost surface and in practice \mathbf{v}_{s-t} replaces \mathbf{s}_{s-t} . Implementing the estimator requires a very fine grid search to find the peak of the likelihood function (via the MLE approach). The grid search—which essentially amounts to stepping the space-time steering vector over potential target angles and Dopplers—is computationally burdensome and can be sub-optimum. Specifically, if the step size is too large, the estimator can exhibit bias and increased variance. For this reason, numerical approximations to the MLE, which exhibit good practical performance, are commonly used.

When processing multiple CPIs, the MLE appears as the sum of the individual cost surfaces at each frequency, where each cost surface takes the form of the term in parentheses in (12.148). An equivalent development for the multistatic radar case is given in [82].

STAP remains computationally challenging to implement. The weight calculation generally takes place in either the voltage domain using QR decomposition (QRD) or in the power domain using Cholesky factorization. Calculating the weight vector using QRD in the voltage domain seems favorable, since it essentially doubles the dynamic range in decibels over the power domain solution and avoids explicitly forming outer products in the covariance estimation step. Cholesky factorization in the power domain is also a common approach when considering all the details of real-time implementation on specific computing devices.

2.12.8 STAP data collection programs

In this section we highlight some measured data collection programs.

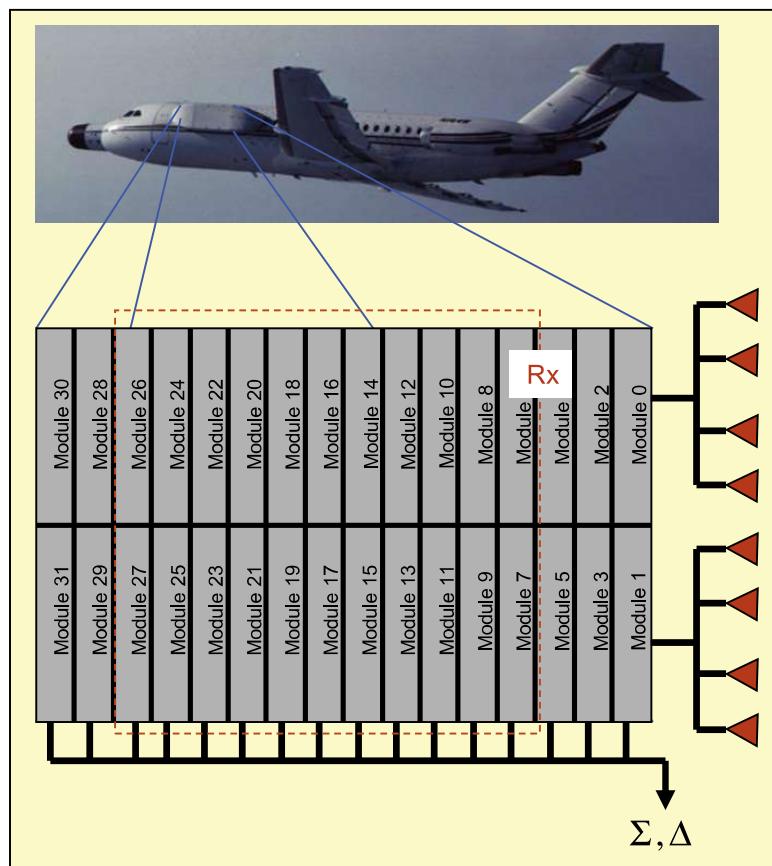
2.12.8.1 Multichannel Airborne Radar Measurements (MCARM)

In the mid-1990s, the US Air Force Rome Laboratory (now the Air Force Research Laboratory) contracted with Westinghouse Corporation (now Northrop Grumman) to build and fly a multi-channel, *L*-band radar under the Multi-Channel Airborne Radar Measurements (MCARM) Program [83]. The MCARM program produced a meaningful collection of multi-channel radar data over a range of clutter conditions to evaluate the performance potential of STAP and support continued technology development.

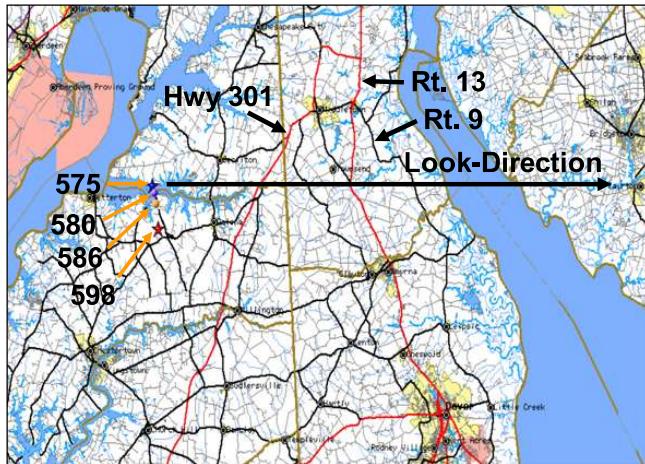
The salient parameters of the MCARM system are given in Table 12.8. As seen from the table, the system has relatively high peak power, an electrically modest aperture at *L*-band, relatively low range resolution, and built-in features to calibrate the antenna and radio frequency front-end. Additionally, as Figure 12.30 shows, MCARM has a highly flexible array configuration, supporting a range of study options, including planar array configuration, uniform linear array, and sum and difference channel analyses. Figure 12.31 shows the region of the Eastern United States—the Delmarva Peninsula—where

Table 12.8 MCARM Nominal System Parameters

Parameter	Value	Parameter	Value
Transmit frequency	L-band (1240 MHz)	Beamwidth	7.5°, and 3× “blob” pattern
Bandwidth	1 MHz Transmit, 0.8 MHz Receive	Array configuration	Most collections use 11-over-11 receive channels, plus sum and difference (see figure)
Peak power	15 kW	Sample rate	5 MHz
Waveform	Linear frequency modulated or gated RF	CPI length	Nominally 128 pulses
PRF	0.5 kHz, 2 kHz, 7 kHz	Other	Test manifold for channel balance, range measured steering vectors

**FIGURE 12.30**

MCARM data collection system.

**FIGURE 12.31**

CPI locations for several acquisitions during MCARM Flight 5.

several acquisitions were collected during MCARM Flight 5. As previously mentioned, this region is relatively flat, dominated by rural clutter and water, and covered by a number of roadways, including several highways. Results from Acquisition 575, in particular, have been presented in a number of papers, starting with [50].

It is known that some of the MCARM acquisitions were previously made publicly releasable, as described in [84]. The Air Force Research Laboratory, Sensors Directorate, maintains the MCARM database.

2.12.8.2 Naval Research Laboratory (NRL) database

The Naval Research Laboratory (NRL) Adaptive Array Flight Test Database is described in [85]. Table 12.9 provides some flight test parameters. According to [85], the effort remarkably consisted of over thirty flights, collecting in excess of 2500 data files. Data were collected predominantly in the mountainous regions of Virginia and West Virginia, with urban clutter data collected in the vicinity of Washington, DC, and rural clutter data collected in North Carolina, Georgia, Ohio, and Indiana. Sea clutter data were collected over the Atlantic Ocean, east of Virginia, with some littoral clutter data also recorded.

Lee and Staudaher [85] indicates a database was constructed, with data formatted for MATLAB® analysis, with the intent “that the database will provide the adaptive signal processing community with a valuable tool for validating both existing and future adaptive algorithms.”

2.12.8.3 Mountaintop database

According to [86,87], the Mountaintop Program goal sought to advance airborne early warning radar technology. Collection of surveillance radar clutter and jamming data from mountaintops at White Sands Missile Range, New Mexico, and Pacific Missile Range Facility, Hawaii, served as central objectives in this effort. The Mountaintop radar, known as the Radar Surveillance Technology Experimental Radar (RSTER), employed an inverse displaced phase center antenna (IDPCA) technique to give the

Table 12.9 NRL Adaptive Array Flight Test System Parameters (from [85])

Parameter	Value	Parameter	Value
Transmit frequency	UHF	Beamwidth	15.7° azimuth, 80° elevation (16.5 dBi gain)
Bandwidth	2.6 MHz	Array configuration	Eight channel, uniform linear array
Peak power	In excess of 1 MW	Sample rate	5 MHz (separate in-phase and quadrature channels)
Waveform	Unstated, presumably linear frequency modulated or gated RF	CPI length	3–18 pulses at low PRF, 34–66 pulses for high PRF
PRF	Typically 300 Hz or 750 Hz	Other	Test manifold for channel balance

Table 12.10 Mountaintop RSTER Parameters (from [87, 88])

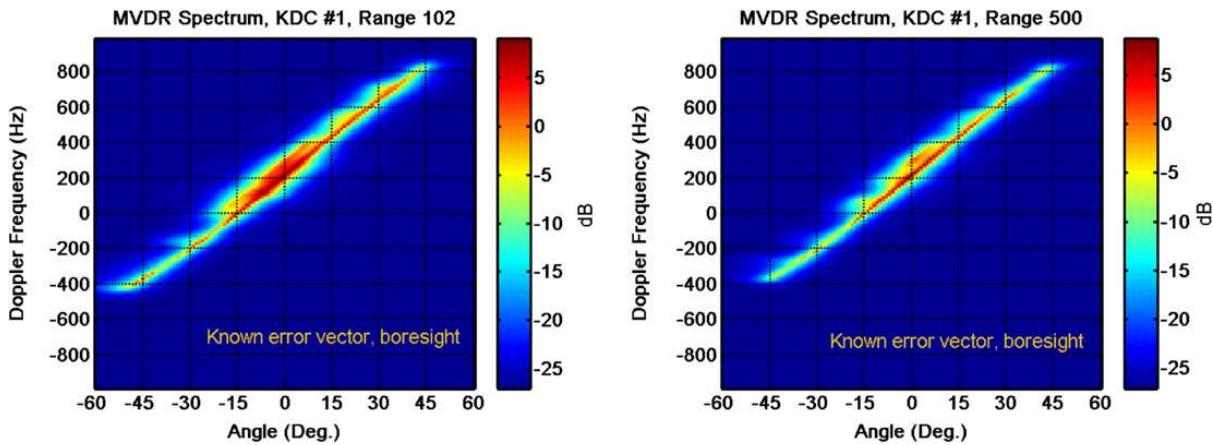
Parameter	Value	Parameter	Value
Transmit frequency	UHF (400–500 MHz)	Beamwidth	9° azimuth, 6° elevation
Bandwidth	200 kHz	Array configuration	14 channel uniform linear array
Peak power	100 kW	Sample Rate	1 MHz [88]
Waveform	Linear frequency modulated [yy]	CPI length	16 pulses
PRF	250–1500 Hz	Other	Horizontally polarized, relocatable

appearance of platform motion. RSTER parameters are given in Table 12.10. In IDPCA, the system steps the transmit location to one (or three) of eighteen total transmit antenna columns.

According to [87], Mountaintop RSTER data are archived using the Common Research Environment for STAP (CREST) database. Additionally, some data is accessible at the IEEE Signal Processing Information Base at Rice University [88].

2.12.8.4 Knowledge-aided sensor signal processing and expert reasoning (KASSPER) data

Improving the performance of STAP in complex clutter environments served as a primary aim of the DARPA KASSPER Program [89]. Under this program, DARPA funded the creation of several sets of synthetic, multichannel, multi-pulse, multi-range datacubes at *X*-band and *L*-band. The *L*-band datacube—KASSPER datacube #1, or KDC #1—is publicly releasable and was provided to attendees at the 2002 DARPA KASSPER Workshop [90]. Many of the *L*-band data parameters resemble those of

**FIGURE 12.32**

MVDR spectra for KDC #1, range bins 102 (left) and 500 (right).

the MCARM system described in Table 12.8, the exception mainly being the use of an eleven channel uniform linear array and typically fewer pulses in the CPI. Each of the KASSPER datacubes is a result of high-fidelity, site-specific clutter simulation. Dense moving targets are present in the data. A unique aspect of the simulation is the availability of truth information, including known covariance matrices on a range cell basis and target location data for some of the datasets (including the *L*-band datacube and some of the *X*-band cases).

Figure 12.32 shows MVDR spectra for KDC #1 range bins 102 and 500, with pre-steering applied to the antenna boresight. Additionally, the KASSPER datacubes included simulated antenna errors, with the true error vector provided with the data sets; the two-dimensional MVDR spectra in Figure 12.32 employ knowledge of the precise array manifold.

2.12.9 Summary

This tutorial describes a number of key aspects of STAP. Naturally, it is challenging to fully cover all the details of a topic so vast. Our goals for this effort are simpler: provide a concise, single reference to augment the available text books, reports, and papers on this important subject and highlight some of the more significant developments within the research community.

In this tutorial, we describe a number of approaches to design optimal filters. Of the methods described, all end up being closely related and ultimately maximize the probability of detection for a fixed probability of false alarm. We then discuss signal models for clutter, noise, radio frequency interference and target. These models form the basis for STAP algorithm development and analysis. STAP, as it turns out, refers to a collection of a number of practical techniques; we discuss several major algorithm themes, include post-Doppler STAP, pre-Doppler STAP, and parametric STAP. We further discuss a number of eigen- or subspace-based cancellation architectures, including Principle Components Inverse, the cross spectral metric, and the Multi-Stage Wiener Filter.

Using the simulation models and algorithm descriptions, we then characterize the performance of the various methods using SINR loss, a key performance metric in the STAP community, using a homogeneous clutter scenario. We find that, for the approach taken, all but one of the methods provides similarly good performance relative to the bound set by the optimal space-time filter.

Key challenges in the STAP research community center on modifying textbook principles to support real-world application. Invariably, the accurate estimation of the unknown clutter covariance matrix and implementing the adaptive filter in a real-time computing environment are driving considerations. In this vein, we briefly discuss the challenges of STAP application in heterogeneous clutter environments, for bistatic or conformal array configurations, and in the presence of otherwise nonstationary clutter.

A number of concepts are concisely summarized in Section 2.12.7, where we describe the end-to-end implementation of a basic MTI radar processor. STAP normalization and target parameter estimation is considered in this section.

We conclude the discussion by summarizing several STAP data collection programs, the characteristics of the various radar systems, and the type of data collected or generated.

Symbols and notation

We provide a key to some common notation used in this paper.

f_s	spatial frequency
f_d	Doppler frequency (Hz)
ϕ, θ	azimuth and elevation (rads)
ϕ_{cone}	cone angle (rads)
ψ	grazing angle (rads)
λ	wavelength (m)
M	number of channels
N	number of pulses
L	number of available range bins
T	pulse repetition interval (s)
σ_n^2	noise variance (w)
σ_T^2	target signal variance (w)
\mathbf{v}_p	platform velocity vector (m/s, m/s, m/s)
\mathbf{s}_s	spatial steering vector
\mathbf{v}_s	hypothesized spatial steering vector
\mathbf{s}_t	temporal steering vector
\mathbf{v}_t	hypothesized temporal steering vector
\mathbf{s}_{s-t}	space-time steering vector
\mathbf{v}_{s-t}	hypothesized space-time steering vector
$\mathbf{x}_{s/k}(n)$	spatial data snapshot, k th range cell, n th pulse
\mathbf{x}_k	space-time data snapshot, k th range cell
\mathbf{c}_k	clutter space-time snapshot, k th range cell
\mathbf{j}_k	RFI space-time snapshot, k th range cell

\mathbf{n}_k	uncorrelated noise space-time snapshot, k th range cell
\mathbf{t}_k	target space-time snapshot, k th range cell
\mathbf{R}_k	null-hypothesis covariance matrix (k th range cell)
$\mathbf{R}_{c/k}$	clutter covariance matrix
\mathbf{R}_{RFI}	RFI covariance matrix
\mathbf{R}_n	receive noise covariance matrix
$\hat{\mathbf{R}}_k$	null-hypothesis covariance estimate, k th range cell
\mathbf{w}_k	space-time weight vector
$\hat{\mathbf{w}}_k$	adaptive space-time weight vector
P_d	probability of detection
P_{fa}	probability of false alarm.

Acronyms

Following is a list of common acronyms:

ADPCA	adaptive displaced phase center antenna
AMF	adaptive matched filter
CFAR	constant false alarm rate
CNR	clutter-to-noise ratio
CPI	coherent processing interval
CREST	Common Research Environment for STAP
CUT	cell-under-test
DNR	discrete-to-noise ratio
DoFs	degrees of freedom
FIR	finite impulse response
FLAR	forward-looking array radar
GLSC	generalized sidelobe canceler
GMTI	ground moving target indication
HTP	Hung-Turner projection
ICM	intrinsic clutter motion
IDPCA	inverse displaced phase center antenna processing
IF	improvement factor
IID	independent and identically distributed
JDL	joint domain localized
KASSPER	Knowledge-Aided Sensor Signal Processing and Expert Reasoning
MCARM	Multi-Channel Airborne Radar Measurements
MCLP	multichannel linear prediction
MDV	minimum detectable velocity
MLE	maximum likelihood estimate
MMSE	minimum mean square error
MSE	mean square error

MSMI	modified sample matrix inversion
MPE	minimum power eigencanceler
MNE	minimum norm eigencanceler
MTI	moving target indication
MV	minimum variance
MVDR	minimum variance distortionless response
MWF	multi-stage Wiener filter
NCA	noncoherent addition
pdf	probability density function
PDI	post-detection integration
PRI	pulse repetition interval
PSD	power spectral density
RDM	range-Doppler map
RFI	radio frequency interference
RD-STAP	reduced-dimension STAP
RR-STAP	reduced-rank STAP
RSTER	Radar Surveillance Technology Experimental Radar
SAR	synthetic aperture radar
SLAR	sidelooking array radar
SMI	sample matrix inversion
SNR	signal-to-noise ratio
SINR	signal-to-interference-plus-noise ratio
STAP	space-time adaptive processing
STOP	space-time optimal processing
TNR	target-to-noise ratio
TSD	targets in the secondary data
WGN	white Gaussian noise
WSS	wide-sense stationary

Relevant Theory: Statistical Signal Processing and Array Signal Processing

See [Vol. 3, Chapter 5](#) Distributed Signal Detection

See [Vol. 3, Chapter 19](#) Array Processing in the Face of Nonidealities

References

- [1] M. Richards, et al. (Eds.), *Principles of Modern Radar: Basic Principles*, SciTech Publishing, Inc., North Carolina, 2010.
- [2] J.V. DiFranco, W.L. Rubin, *Radar Detection*, Artech-House, Dedham, MA, 1980.
- [3] L.E. Brennan, I.S. Reed, *Theory of adaptive radar*, IEEE Trans. AES 9 (2) (1973) 237–252.
- [4] I.S. Reed, *A brief history of adaptive arrays*, Sudbury/Wayland Lecture Series (Raytheon Div. Education), Notes, 23 October 1985.

- [5] R. Klemm, Doppler properties of airborne clutter, in: Proceedings of the Research and Technology Organization, North Atlantic Treaty Organization (RTO-NATO) Lecture Series 228—Military Applications of Space-Time Adaptive Processing, RTO-ENP-027, September 2002, pp. 2-1–2-24.
- [6] J.N. Entzminger, C.A. Fowler, W.J. Kenneally, JointSTARS and GMTI: past, present and future, IEEE Trans. AES 35 (2) (1999) 748–761.
- [7] M.I. Skolnik, Introduction to Radar Systems, second ed., McGraw Hill, New York, NY, 1980.
- [8] N. Levanon, Radar Principles, John Wiley & Sons, New York, 1988.
- [9] W.L. Melvin, A STAP overview, in: Peter Willett (Ed.), IEEE AES Systems Magazine—Special Tutorials Issue, 19 (1) (2004) 19–35.
- [10] J. Ward, Space-Time Adaptive Processing for Airborne Radar, Lincoln Laboratory Technical Report, ESC-TR-94-109, December 1994.
- [11] J.R. Guerci, Space-Time Adaptive Processing for Radar, Artech House, Norwood, MA, 2003.
- [12] S. Haykin, Adaptive Filter Theory, third ed., Prentice-Hall, Upper Saddle River, NJ, 1996.
- [13] D.H. Johnson, D.E. Dudgeon, Array Signal Processing: Concepts and Techniques, Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [14] L.J. Griffiths, C.W. Jim, An alternative approach to linearly constrained adaptive beamforming, IEEE Trans. Antenn. Propag. 30 (1) (1982) 27–34.
- [15] L.M. Novak, M.B. Sechtin, M.J. Cardullo, Studies of target detection algorithms that use polarimetric radar data, IEEE Trans. AES 25 (2) (1989) 150–165.
- [16] W.L. Melvin, G.A. Showman, Knowledge-aided parametric covariance estimation, IEEE Trans. AES (2006) 1021–1042.
- [17] I.S. Reed, J.D. Mallett, L.E. Brennan, Rapid convergence rate in adaptive arrays, IEEE Trans. AES 10 (6) (1974) 853–863.
- [18] R. Klemm, Space-time adaptive processing: principles and applications, IEE Radar, Sonar, Navigation and Avionics 9, IEE Press, 1998.
- [19] R. Klemm, Principles of space-time adaptive processing, second ed., IEE Radar, Sonar, Navigation and Avionics 12, IEE Press, UK, 2002. (Note: a 3rd edition of this text, published in 2006, is also available.)
- [20] D. Barton, Land clutter models for radar design and analysis, Proc. IEEE 73 (2) (1985) 198–204.
- [21] J.B. Billingsley, Low-Angle Radar Land Clutter: Measurements and Empirical Models, William Andrew Publishing, Inc., 2002.
- [22] R.J. Mailloux, Phased Array Antenna Handbook, Artech House, Boston, MA, 1994.
- [23] M. Zatman, How narrow is narrowband? IEE Proc. Radar Sonar Navig. 145 (2) (1998) 85–91.
- [24] W.L. Melvin, G.A. Showman, J.R. Guerci, A knowledge-aided GMTI detection architecture, in: Proceedings of the 2004 IEEE Radar Conference, Philadelphia, PA, 26–29 April, 2004, ISBN: 0-7803-8235-8.
- [25] D.V. Rabinkin, N.B. Pulsone, Subband-domain signal processing for radar array systems, in: Proceedings of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations IX, Denver, CO., July 1999, pp. 174–187.
- [26] R. Klemm, Adaptive clutter suppression for airborne phased array radar, Proc. IEE 130 (1) (1983) 125–132.
- [27] C.H. Gierull, B. Balaji, Minimal sample support space-time adaptive processing with fast subspace techniques, IEE Proc. Radar Sonar Navig. 149 (5) (2002) 209–220.
- [28] W.F. Gabriel, Using spectral estimation techniques in adaptive processing antenna systems, IEEE Trans. Antenn. Propag. 34 (3) (1986) 291–300.
- [29] D.W. Tufts, I. Kirssteins, R. Kumaresan, Data-adaptive detection of a weak signal, IEEE Trans. AES 19 (2) (1983) 313–316.
- [30] A.M. Haimovich, The Eigencanceler: adaptive radar by eigenanalysis methods, IEEE Trans. AES 32 (2) (1996) 532–542.

- [31] E.K.L. Hung, R.M. Turner, A fast beamforming algorithm for large arrays, *IEEE Trans. AES* 19 (4) (1983) 598–607.
- [32] B.D. Carlson, Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans. AES* 24 (4) (1988) 397–401.
- [33] J.R. Guerci, J.S. Goldstein, I.S. Reed, Optimal and adaptive reduced-rank STAP, *IEEE Trans. AES* 36 (2) (2000) 647–661.
- [34] S.D. Berger, B.M. Welsh, Selecting a reduced-rank transformation for STAP—a direct form perspective, *IEEE Trans. AES* 35 (2) (1999) 722–729.
- [35] J.S. Goldstein, I.S. Reed, P.A. Zulch, Multistage partially adaptive STAP CFAR detection algorithm, *IEEE Trans. AES* 35 (2) (1999) 645–661.
- [36] R.C. DiPietro, Extended factored space-time processing for airborne radar, in: Proceedings of the 26th Asilomar Conference, Pacific Grove, CA, October 1992, pp. 425–430.
- [37] H. Wang, L. Cai, On adaptive spatial-temporal processing for airborne surveillance radar systems, *IEEE Trans. AES* 30 (3) (1994) 660–670.
- [38] R. Brown, M. Wicks, Y. Zhang, Q. Zhang, H. Wang, A space-time adaptive processing approach for improved performance and affordability, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 321–326.
- [39] R. Klemm, Antenna design for airborne MTI, in: Proceedings of the Radar 92, October 1992, Brighton, UK, pp. 296–299.
- [40] R. Blum, W. Melvin, M. Wicks, An analysis of adaptive DPCA, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 303–308.
- [41] J.R. Roman, M. Rangaswamy, D.W. Davis, Q. Zhang, B. Himed, J.H. Michels, Parametric adaptive matched filter for airborne radar applications, *IEEE Trans. AES* 36 (2) (2000) 677–692.
- [42] G.A. Showman, Personal communication, 13 July 2003.
- [43] R.L. Fante, E.C. Barile, T.P. Guella, Clutter covariance smoothing by subaperture averaging, *IEEE Trans. AES* 30 (3) (1994) 941–945.
- [44] W.L. Melvin, Space-time adaptive radar performance in heterogeneous clutter, *IEEE Trans. AES* 36 (2) (2000) 621–633.
- [45] W.L. Melvin, J.R. Guerci, Adaptive detection in dense target environments, in: Proceedings of the 2001 IEEE Radar Conference, Atlanta, GA, 1–3 May 2001, pp. 187–192.
- [46] W.L. Melvin, STAP in heterogeneous clutter environments, in: R. Klemm (Ed.), *The Applications of Space-Time Processing*, IEE Radar, Sonar, Navigation and Avionics 9, IEE Press, 2004.
- [47] W.L. Melvin, J.R. Guerci, Knowledge-aided sensor signal processing: a new paradigm for radar and other sensors, *IEEE Trans. AES* (2006) 983–996.
- [48] O. Kreyenkamp, R. Klemm, Doppler compensation in forward-looking STAP radar, *IEE Proc. Radar Sonar Navig.* 148 (5) (2001) 252–258.
- [49] G. Borsari, Mitigating effects on STAP processing caused by an inclined array, in: Proceedings of the 1998 IEEE Radar Conference, Dallas, Tx, May 1998, pp. 135–140.
- [50] W.L. Melvin, M.C. Wicks, R.D. Brown, Assessment of multichannel airborne radar measurements for analysis and design of space-time processing architectures and algorithms, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, Michigan, May 13–16, 1996, pp. 130–135.
- [51] W.L. Melvin, M.C. Wicks, Improving practical space-time adaptive radar, in: Proceedings of the 1997 IEEE National Radar Conference, Syracuse, New York, May 13–15, 1997, pp. 48–53.
- [52] P. Chen, W.L. Melvin, M.C. Wicks, Screening among multivariate normal data, *J. Multivariate Anal.* 69 (1999) 10–29.

- [53] D.J. Rabideau, A.O. Steinhardt, Improving the performance of adaptive arrays in non-stationary environments through data-adaptive training, in: Proceedings of the 30th Asilomar Conference, Pacific Grove, CA, 3–6 November 1996, pp. 75–79.
- [54] D.J. Rabideau, A.O. Steinhardt, Improved adaptive clutter cancellation through data-adaptive training, *IEEE Trans. Aerosp. Electron. Syst.* 35 (3) (1999) 879–891.
- [55] G.A. Showman, W.L. Melvin, Knowledge-aided discrete removal techniques, in: Proceedings of the 2005 DARPA ISIS/KASSPER Workshop, Las Vegas, NV, 22–24 February 2005, CD ROM.
- [56] N. Pulsone, Improving ground moving target indication performance, in: Proceedings of the DARPA/AFRL KASSPER Workshop, 14–16 April 2003, Las Vegas, NV, 2003.
- [57] S.D. Berger, W.L. Melvin, G.A. Showman, Map-aided secondary data selection, in: Proceedings of the 2007 IEEE Radar Conference, Boston, MA, April 2007.
- [58] J.R. Guerci, Theory and application of covariance matrix tapers for robust adaptive beamforming, *IEEE Trans. Signal Process.* 47 (4) (1999) 977–985.
- [59] S.U. Pillai, Y.L. Kim, J.R. Guerci, Generalized forward/backward subaperture smoothing techniques for sample starved STAP, *IEEE Trans. Signal Process.* 48 (12) (2000) 3569–3574.
- [60] A. Farina, P. Lombardo, M. Pirri, Nonlinear nonadaptive space-time processing for airborne early warning radar, *IEE Proc. Radar Sonar Navig.* 145 (1) (1998) 9–18.
- [61] J.S. Bergin, C.M. Teixeira, P.M. Techau, J.R. Guerci, STAP with knowledge-aided data pre-whitening, in: Proceedings of the 2004 IEEE Radar Conference, April 2004, pp. 289–294.
- [62] J.S. Bergin, C.M. Teixeria, P.M. Techau, J.R. Guerci, Space-time beamforming with knowledge-aided constraints, in: Proceedings of the 2003 Adaptive Sensor Array Processing Workshop, MIT Lincoln Laboratory, Lexington, MA, March, 2003.
- [63] G.R. Letgers, J.R. Guerci, Physics-based airborne GMTI radar signal processing, in: Proceedings of the 2004 IEEE Radar Conference, Philadelphia, PA, April 2004, pp. 283–288, ISBN: 0-7803-8234-X.
- [64] C.D. Richmond, Statistical performance analysis of the adaptive sidelobe blanker detection algorithm, in: Proceedings of the 31st Annual Asilomar Conference—Signal, Systems & Computers, Pacific Grove, CA, 2–5 November 1997, pp. 872–876.
- [65] W.L. Melvin, G.A. Showman, Knowledge-aided, physics-based signal processing for next-generation radar, in: Proceedings of the Asilomar Conference on Signals, Systems, Computers, Pacific Grove, CA, November 2007.
- [66] W.L. Melvin, M.J. Callahan, M.C. Wicks, Adaptive clutter cancellation in bistatic radar, in: Proceedings of the 34th Asilomar Conference, Pacific Grove, CA, October 2000, pp. 1125–1130.
- [67] B. Himed, J.H. Michels, Y. Zhang, Bistatic STAP performance analysis in radar applications, in: Proceedings of the 2001 IEEE Radar Conference, Atlanta, GA, May 2001, pp. 198–203.
- [68] S.D. Hayward, Adaptive beamforming for rapidly moving arrays, in: Proceedings of the CIE International Conference on Radar, IEEE Press, Beijing, CN, 8–10 October 1996, pp. 480–483.
- [69] S.M. Kogon, M.A. Zatman, Bistatic STAP for airborne radar systems, in: Proceedings of the IEEE SAM, Lexington, MA, March 2000.
- [70] F. Pearson, G. Borsari, Simulation and analysis of adaptive interference suppression for bistatic surveillance radars, in: Proceedings of the 2001 ASAP Symposium, Lexington, MA, 13 March 2001.
- [71] B. Himed, Y. Zhang, A. Hajjari, STAP with angle-Doppler compensation for bistatic airborne radars, in: Proceedings of the 2002 IEEE Radar Conference, Long Beach, CA, 22–25 April 2002, ISBN: 0-7803-7358-8.
- [72] W.L. Melvin, M.E. Davis, Adaptive cancellation method for geometry-induced non-stationary bistatic clutter environments, *IEEE Trans. AES* (2007) 651–672.
- [73] W.L. Melvin, B. Himed, M.E. Davis, Doubly-adaptive bistatic clutter filtering, in: Proceedings of the 2003 IEEE Radar Conference, Huntsville, AL, 5–8 May 2003, pp. 171–178.

- [74] F.D. Lapierre, J.G. Verly, M. Van Droogenbroeck, New solutions to the problem of range dependence in bistatic STAP radars, in: Proceedings of the 2003 IEEE Radar Conference, Huntsville, AL, 5–8 May 2003, pp. 452–459.
- [75] W.L. Melvin, Adaptive moving target indication, in: N. Willis, H. Griffiths (Eds.), *Advances in Bistatic Radar*, Sci-Tech Publishing, 2007 (Chapter 11).
- [76] R.K. Hersey, W.L. Melvin, J.H. McClellan, E. Culpepper, Adaptive ground clutter suppression for conformal array radar systems, *IET Radar Sonar Navig.* 3 (4) (2009) 357–372.
- [77] R.K. Hersey, W.L. Melvin, E. Culpepper, Adaptive filtering for conformal array radar, in: Proceedings of the IEEE Radar Conference, May 2008, Rome, Italy.
- [78] R.K. Hersey, W.L. Melvin, J.H. McClellan, Clutter-limited detection performance of multi-channel conformal arrays, *Signal Process.* 84 (2004) 1481–1500 (special issue on New Trends and Findings in Antenna Array Processing for Radar).
- [79] F.C. Robey, D.R. Fuhrman, E.J. Kelly, R. Nitzberg, A CFAR adaptive matched filter detector, *IEEE Trans. AES* 28 (1) (1992) 208–216.
- [80] W.S. Chen, I.S. Reed, A new CFAR detection test for radar, *Digital Signal Processing*, vol. 1, Academic Press, 1991, pp. 198–214.
- [81] R.C. Davis, L.E. Brennan, I.S. Reed, Angle estimation with adaptive arrays in external noise fields, *IEEE Trans. AES* AES-12 (2) (1976) 179–186.
- [82] W.L. Melvin, R. Hancock, M. Rangaswamy, J. Parker, Adaptive distributed radar, in: Proceedings of the 2009 International Radar Conference, Bordeaux, France, October 2009.
- [83] D.K. Fenner, W.F. Hoover, Test results of a space-time adaptive processing system for airborne early warning radar, in: Proceedings of the 1996 IEEE National Radar Conference, Ann Arbor, MI, 13–16 May 1996, pp. 88–93.
- [84] V. Cavo, MCARM/STAP Data Analysis, Final Technical Report, Air Force Research Laboratory, AFRL-SN-RS-TR-1999-48, May, Volume One (of Two), 1999.
- [85] F. Lee, F. Staudaher, The NRL adaptive array flight test database, in: Proceedings of the IEEE Adaptive Antenna Systems Symposium, 1992, pp. 101–104.
- [86] G.W. Titi, An overview of the ARPA/Navy mountaintop program, in: Proceedings of the IEEE Adaptive Antenna Systems Symposium, 1994.
- [87] G.W. Titi and D.F. Marshall, The ARPA/Navy Mountaintop Program: adaptive signal prcessing for airborne early warning radar, in: Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 2, pp. 1165–1168.
- [88] <http://spib.rice.edu/spib/mtn_top.html>.
- [89] J.R. Guerci, Knowledge-aided sensor signal processing and expert reasoning, in: Proceedings of the 2002 Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER) Workshop, Washington, DC, 3 April, 2002, CD ROM.
- [90] J.S. Bergin, P.M. Techau, (Electronic data) Workshop datacube, in: Proceedings of the 2002 KASSPER Conference, Washington, DC, 3 April 2002, CD ROM.