ADAPTIVE RADAR

RADAR, ADAPTIVE

Effective radar was developed during World War II to detect and engage enemy aircraft, warships, and missiles, whose increasing sophistication and lethality exposed anyone without early warning to devastating attack. This motivation intensified during the Cold War; fear of nuclear weapons impelled further developments in radar and their requisite electronics and digital computers. Buderi (1) provides a compelling and thorough history. Modern weapons systems are designed to regain the advantage of surprise by attempting to thwart radar detection. The technological response to this tactic is to design radars that are more sensitive and robust against enemy countermeasures. Adaptivity is one method that provides modern radars with these advantages.

Electromagnetic interference is a principal factor affecting any radar's sensitivity. Such interference may be generated externally to the radar by an opponent seeking to jam the radar, or internally by radar reflections from the ground. External interference is called *jamming,* and internal interference is called *clutter.* This article describes adaptive methods to mitigate such interference and provides performance predictions for adaptive radar systems.

Clutter interference has always affected radar performance. The earliest and simplest technique for eliminating ground clutter is moving-target indicator (*MTI*) radar (2), where the Doppler effect is used to distinguish moving targets from the stationary ground. If the radar is mounted on an aircraft, then the ground clutter will have a nonzero Doppler shift, and an airborne MTI algorithm such as DPCA (3) must be used to cancel the clutter. Both MTI and DPCA are nonadaptive algorithms—they do not use information from the radar data to determine the filtering required to eliminate clutter interference. Instead, they use physical assumptions about the radar scenario to filter the clutter, and their performance is limited by the accuracy of these assumptions. In contrast, adaptive methods (4–8) are robust against such modeling mismatches because actual measurements from the radar itself are used in the mitigation scheme.

Jammer interference can be eliminated nonadaptively through the use of low-sidelobe antenna arrays; however, it is still possible to jam a radar through low sidelobes. To guard against false detections from jammers, an auxiliary antenna called a sidelobe blanker (*SLB*) is used to detect the presence of sidelobe jamming. Because it is desirable to detect targets in the presence of jamming, *ECCM* (electronic counter-countermeasure) techniques such as adaptive sidelobe cancelers (*SLC*s) or adaptive arrays are used to adaptively subtract the jammer energy from the radar receiver (9, 10).

Space–time adaptive processing (*STAP*) (3,4,11–13) [Guerci (59)] provides a unified approach to radar jamming and clutter interference rejection. Because jammers lie in specific directions from the radar, jammer mitigation requires spatial adaptivity. Mitigating clutter interference in an airborne radar requires space–time adaptivity because the ground's Doppler spectrum varies with its azimuth and elevation with respect to the aircraft.

This article is concerned with the description and performance of algorithms used for adaptive cancellation of radar interference. A radar's job is to receive a signal, examine the signal for the presence of targets, and provide estimates of the speeds and positions of any targets it finds. The methods used to perform each of these steps are motivated by simple physical and statistical models of the radar and its environment. The section "Adaptive Radar Fundamentals" describes the fundamentals of receiving a radar signal, including a physical description of the radar, its environment, the type of signals it is to receive, and the nature of the interference with which it must contend. Performance metrics used throughout the article are also introduced and defined. The section "Statistical Model for Adaptive Radar" provides the statistical assumptions used to formulate the adaptive radar detection problem, derive adaptive algorithms, and deduce their performance. The subsequent four sections ("Optimum Adaptive Processing," "Space–Time Adaptive Processing," "Adaptive Detection," and "Adaptive Weight Computation") all address different aspects of the problem of examining the signal for the presence of targets. The section "Optimum Adaptive Processing" describes the general theory of detecting targets in the presence of interference and provides performance bounds for this problem. The section "Space–Time Adaptive Processing" applies the theory of the previous sections to the radar problem, whose joint spatial and temporal nature imposes a special structure for interference mitigation. Several *STAP* algorithms are classified and analyzed in this section. After the interference has been mitigated, a threshold detector is used to declare the presence of any targets. The section "Adaptive Detection" provides predictions of adaptive radar detection performance for the case where the measured radar data itself is used to mitigate interference. The numerical procedures used to implement adaptive processing and detection on a computer are described in the section "Adaptive Weight Computation." Finally, bounds on the achievable accuracy of a target's position are given in the section "Adaptive Radar Estimation."

ADAPTIVE RADAR FUNDAMENTALS

Matched-Filter Signal Processing

Radars operate by transmitting a burst of radio frequency (*RF*) energy, then matched-filtering the return signal to detect the presence of any targets. In a noise-only environment, a peak in the matched-filter output indicates reflected target energy. Because the reflected signal is expected to be coherent in time, space (if there are independent receiver elements), and pulses (if several pulses are transmitted), the matched filter is a function of these variables. Matched filtering over the RF burst's time duration is considered in this section. This time scale is sometimes called *fast time* to distinguish it from the *slow time* scale of the pulse repetition interval (*PRI*) for the case of several coherent pulses. Spatial and slow-time processing, called space–time processing when combined, are consid-

ered in the first three sub-subsections of the next subsection ("Steering Vectors") and in the section "Space–Time Adaptive Processing."

All radar signals $s(t)$ are viewed as complex phasors $e^{j2\pi f_0 t}$ modulated by a bandlimited pulse $p(t)$, where the bandwidth of the pulse is less than the center frequency f_0, i.e., the radar waveform is expressed as

$$s(t) = p(t)e^{j2\pi f_0 t} \tag{1}$$

The pulse $p(t)$ and radar waveform $s(t)$ have the spectral decompositions

$$p(t) = \int_{-\infty}^{\infty} P(f)e^{j2\pi ft}\,df \tag{2}$$

$$s(t) = \int_{-\infty}^{\infty} S(f)e^{j2\pi ft}\,df \tag{3}$$

i.e., they are sums of weighted phasors over the signal bandwidth. By Eq. (1), $S(f) = P(f + f_0)$. Because $p(t)$ is bandlimited, $P(f)$ and $S(f)$ have finite support near 0 and f_0, respectively.

If the radar transmits the signal $s_t(t)$ and receives the signal $s_r(t)$, then the matched filter output at time delay τ is given by the integral

$$\mathrm{mf}(\tau) = (\text{matched filter output at } \tau) = \int_{-\infty}^{\infty} s_t^*(t-\tau)s_r(t)\,dt \tag{4}$$

where the asterisk denotes complex conjugation. We now use the integral form of the Cauchy–Schwarz inequality: For complex signals $f(t)$ and $g(t)$ with finite energy,

$$\left|\int_{-\infty}^{\infty} f^*(t)g(t)\,dt\right|^2 \le \int_{-\infty}^{\infty} |f(t)|^2\,dt \int_{-\infty}^{\infty} |g(t)|^2\,dt$$

where equality holds if and only if $f(t)$ is a scalar multiple of $g(t)$. Thus Eq. (4) has a peak at τ' if the received signal $s_r(t)$ is a multiple of the transmitted signal $s_t(t)$ delayed by τ'—almost precisely the condition of a reflected radar signal.

A pulse Doppler radar transmits coherent pulses at regular intervals called the pulse repetition interval (*PRI*). Denote the PRI by T_r. The mth pulse transmitted has the form

$$s_t(t) = p(t - mT_r)e^{j2\pi f_0 t} \tag{5}$$

$$= \int_{-\infty}^{\infty} S(f)e^{j2\pi(f-f_0)T_r m}e^{j2\pi ft}\,df \tag{6}$$

If this pulse is reflected by a target moving at velocity v toward the radar and propagates back to the radar, then under the far-field plane wave assumption, the value of the reflected field at time t and position $\boldsymbol{r} = (r_1, r_2, r_3)^{\mathrm{T}}$ is given by the signal

$$s_r(t, \boldsymbol{r}) = a s_t((1+\delta)(t - \tau + c^{-1}\boldsymbol{k}\cdot\boldsymbol{r})) \tag{7}$$

where a is a combined reflection and propagation coefficient called the target amplitude, c is the speed of light, the time dilation $1 + \delta$ due to the reflection is

$$1 + \delta = \frac{1 + v/c}{1 - v/c} \approx 1 + \frac{2v}{c} \qquad (v \ll c) \tag{8}$$

the time delay τ due to the round trip distance $2R$ to the target is

$$\tau = \frac{2R}{c} \tag{9}$$

and $\boldsymbol{k} = (k_1, k_2, k_3)^{\mathrm{T}}$ is the unit vector in the direction from the radar to the target ($\|\boldsymbol{k}\| = 1$). The dot denotes the dot or scalar product, so that $\boldsymbol{k}\cdot\boldsymbol{r} = \boldsymbol{k}^{\mathrm{T}}\boldsymbol{r} = \Sigma_{i=1}^{3} k_i r_i$. Using Eq. (6) and ignoring the time delay τ caused by the round trip distance to the target, the spectral decomposition of the reflected wave $s_r(t, \boldsymbol{r})$ is

$$s_r(t, \boldsymbol{r}) = a\int_{-\infty}^{\infty} S(f)e^{-j2\pi(f-f_0)T_r m}e^{j2\pi(f+f_d)(t+c^{-1}\boldsymbol{k}\cdot\boldsymbol{r})}\,df \tag{10}$$

where f_d is the frequency-dependent Doppler shift

$$f_d = \frac{2v}{\lambda} \tag{11}$$

at the (frequency-dependent) wavelength

$$\lambda = \frac{c}{f} \tag{12}$$

Assuming that $f_d < f_0$, $\boldsymbol{k} = -2\pi\lambda_0^{-1}\boldsymbol{k}$ is the wave vector at the center frequency, correctly implying that waves propagate from the target to the radar with speed c.

The matched filter output of the reflected target signal is obtained by combining Eqs. (4), (6), and (10):

$$\mathrm{mf}(\tau) = a\int_{-\infty}^{\infty} S^*(f)S(f - f_d)e^{j2\pi f\tau}e^{j2\pi f_d T_r m}e^{j2\pi\lambda^{-1}\boldsymbol{k}\cdot\boldsymbol{r}}\,df \tag{13}$$

In general it is important to keep in mind the frequency dependence of the Doppler frequency f_d and the wavelength λ. This integral is important for wideband waveforms, but for narrowband signals [bandwidth of $p_t(t)$ less than a few percent of f_0], $\lambda \approx \lambda_0$ and $f_d \approx 2v/\lambda_0$ over the entire signal bandwidth, and the matched filter output may be approximated by

$$\mathrm{mf}(\tau) = a\chi(\tau, f_d)e^{j2\pi f_d T_r m}e^{j2\pi\lambda_0^{-1}\boldsymbol{k}\cdot\boldsymbol{r}} \tag{14}$$

where

$$\chi(\tau, f_d) = \int_{-\infty}^{\infty} P^*(f)P(f - f_d)e^{j2\pi f\tau}\,df \tag{15}$$

$$= \int_{-\infty}^{\infty} p^*(t-\tau)p(t)e^{j2\pi f_d t}\,dt \tag{16}$$

is called the ambiguity function (14) of the waveform $p(t)$. Note that the phase factor $e^{j2\pi f_0\tau}$ has been absorbed into the unknown amplitude a. The value of the ambiguity function will also be absorbed into a for the remainder of this article, but it is important to keep in mind the physical significance of this quantity for radar waveform design.

The ambiguity function is so called because it describes the tradeoff between the knowledge of a target's position and velocity. Two ideal extremes for our choice of baseband

waveform are $p(t) = \delta(t)$ (an impulse) and $p(t) \equiv 1$ (a perfect tone, or so-called continuous wave, or *CW*, signal). In the case of an impulse, $\chi(\tau, f_d) = \delta(\tau)$, i.e., the ambiguity function is an impulse in time, and we have perfect knowledge of the target's range via Eq. (9), but no knowledge of its Doppler frequency. In the case of a CW signal, $\chi(\tau, f_d) = \int_{-\infty}^{\infty} e^{j2\pi f_d t}\, dt = \delta(f_d)$, i.e., the ambiguity function is an impulse in Doppler frequency, and we have perfect knowledge of the target's velocity via Eq. (11), but no knowledge of its position. Actual waveforms, such as linear FM, are designed for specific radar missions with this tradeoff as well as implementation complexity issues in mind.

Actual radar waveforms possess a fine bandwidth B that determines the radar's range resolution

$$\Delta R = \frac{c}{2B} \tag{17}$$

The matched filter output of Eq. (4) is sampled in time at this frequency. The range samples corresponding via Eq. (9) to the sequence of discrete matched-filter samples are called *rangegates*. A peak in the sequence of outputs at a particular rangegate indicates the presence of a target at that range.

Steering Vectors

A steering vector is defined by the response of a signal coherently measured across a sensor array distributed in space, time, or both space and time. It therefore represents a discrete matched filter in space–time. Given a sensor's space–time position, these coherent matched-filter outputs are predicted by Eq. (13) (wideband case) and (14) (narrowband case). Denote the predicted response of coherent matched-filter outputs by the vector

$$\boldsymbol{v} = (v_1, v_2, \ldots, v_N)^{\mathrm{T}} \tag{18}$$

where N is the number of independent outputs. The vector $\boldsymbol{v}$ is called the steering vector because of its dependence on target angle (through $\hat{\boldsymbol{k}}$) and Doppler frequency (through T_r). For every combination of target angle and Doppler frequency, there is a specific vector $\boldsymbol{v}$ "steered" to that combination. Only the direction of $\boldsymbol{v}$ is important, not its length.

Let

$$\boldsymbol{z} = (z_1, z_2, \ldots, z_N)^{\mathrm{T}} \tag{19}$$

be the vector of matched-filter outputs measured across the sensor array in space and/or time. The vectors $\boldsymbol{v}$ and $\boldsymbol{z}$ are analogous to the functions $s_t(t)$ and $s_r(t)$ in Eq. (4). The coherent combination of the outputs of all N sensors is given by

$$\sum_{n=1}^{N} v_n^* z_n = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{z} \tag{20}$$

where superscript H denotes the Hermitian or conjugate transpose. The finite-dimensional form of the Cauchy–Schwartz inequality states that for complex vectors $\boldsymbol{v}$ and $\boldsymbol{z}$,

$$|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{z}|^2 \le \|\boldsymbol{v}\|^2 \|\boldsymbol{z}\|^2$$

where equality holds if and only if $\boldsymbol{z}$ is a scalar multiple of $\boldsymbol{v}$. Thus, the matched filter output is maximized when the measurement vector $\boldsymbol{z}$ is a scalar multiple of the steering vector $\boldsymbol{v}$.

Phased Array Radar. A phased array radar is composed of N sensors located at positions $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ (with respect to an arbitrary origin $\boldsymbol{r}_0$), each with independent, synchronous matched-filter receivers. By Eq. (14), the matched-filter outputs from the N sensors are given by the N-by-1 steering vector

$$\boldsymbol{v} = \begin{pmatrix} e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_1} \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_2} \\ \vdots \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_N} \end{pmatrix} \tag{21}$$

where the unknown target amplitude a, the ambiguity function $\chi(\tau, f_d)$, the phase factor $e^{j2\pi f_d T_r m}$ corresponding to the time delay for the mth pulse, and the arbitrary phase factor $e^{-j2\pi\lambda_0^{-1}\boldsymbol{r}_0\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_0}$ corresponding to the origin are ignored. Because the vector $\hat{\boldsymbol{k}} = \hat{\boldsymbol{k}}(\phi, \theta)$ depends upon the target's azimuth angle ϕ and elevation angle θ, the spatial steering vector

$$\boldsymbol{v} = \boldsymbol{v}(\phi, \theta) \tag{22}$$

also depends upon these angles.

The case of a rotating uniform linear array is encountered frequently; here the sensors are located at positions

$$\boldsymbol{r}_n = nd(\hat{\boldsymbol{i}}\cos\phi_a + \hat{\boldsymbol{j}}\sin\phi_a), \qquad n = 0, 1, \ldots, N-1 \tag{23}$$

where d is the interelement spacing (ideally, $d = \lambda_0/2$), ϕ_a is the array's rotation in the xy plane, and $\hat{i}$ and $\hat{j}$ are the standard unit vectors along the x and y axes, respectively. For a target located in direction $\hat{\boldsymbol{k}} = \hat{i}\cos\phi_t + \hat{j}\sin\phi_t$ with respect to the radar, Eq. (21) yields the steering vector (Fig. 1)

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ e^{j2\pi d\lambda_0^{-1}u} \\ \vdots \\ e^{j2\pi d\lambda_0^{-1}u(N-1)} \end{pmatrix}, \qquad u = \sin\phi_t' = \cos(\phi_t - \phi_a) \tag{24}$$

where $\phi'_t = \phi_a - \phi_t + 90^\circ$ is the target cone angle with respect to the array's normal. A bank of such steering vectors sensitive to many target directions is used to determine a target's position, analogously to the way a matched filter is used to determine a target's range.

If the N antenna sensors are not uniform but have distinct responses to targets at different azimuths and elevations and different polarizations, then the spatial steering vector of Eq. (21) must be replaced by

$$\boldsymbol{v}(\phi, \theta, \boldsymbol{p}) = \begin{pmatrix} G_1(\phi, \theta, \boldsymbol{p})e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_1} \\ G_2(\phi, \theta, \boldsymbol{p})e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_2} \\ \vdots \\ G_N(\phi, \theta, \boldsymbol{p})e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_N} \end{pmatrix} \tag{25}$$

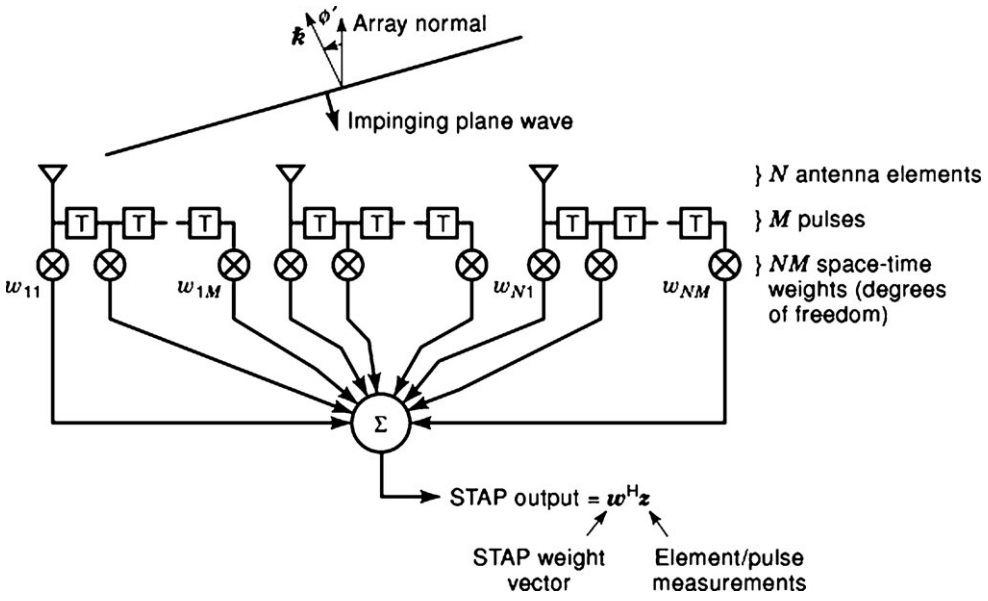

Figure 1. Space–time array.

where $G_i(\phi, \theta, \boldsymbol{p})$ $(i = 1, \ldots, N)$ is the ith sensor's complex antenna response to a signal at azimuth ϕ and elevation θ, and with polarization vector $\boldsymbol{p}$. Oftentimes $\boldsymbol{p}$ is decomposed into a vertical and horizontal component, in order to consider these two independent polarization modes considered separately.

Pulse Doppler Radar. A pulse Doppler radar transmits M coherent pulses at regular PRIs, then matched-filters the radar returns during the PRI. If the target remains coherent throughout the entire coherent pulse interval (*CPI*) of all M pulses, then the vector of all M coherent matched filter outputs is given in Eq. (14) by the M-by-1 steering vector

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ e^{j2\pi f_d T_r} \\ \vdots \\ e^{j2\pi f_d T_r (M-1)} \end{pmatrix} \tag{26}$$

where the unknown target amplitude a and the phase factor $e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}}$ corresponding to the sensor's position are ignored. The temporal steering vector

$$\boldsymbol{v} = \boldsymbol{v}(f_d) \tag{27}$$

is a function of the target Doppler frequency. A bank of steering vectors sensitive to many Doppler frequencies is used to determine a target's velocity (slow-time matched filtering), analogously to the way a matched filter is used to determine a target's range (fast-time matched filtering). Note that a target whose distance causes a round-trip delay greater than one PRI appears range-aliased with targets within the maximum range

$$R_{max} = \frac{cT_r}{2} \tag{28}$$

This range aliasing problem may be overcome by using a sequence of different pulse repetition frequencies (*PRF*s) over several CPIs, then using the Chinese remainder theorem to disambiguate target ranges over a greater unambiguous range.

Space–Time Steering Vectors. A phased array is used in conjunction with a pulse Doppler waveform to mitigate jammer and clutter interference in airborne radars, as well as providing spatial and temporal coherent gain on targets. Examining Eq. (14) at the combination of all possible pulses and element positions yields the M N-by-1 space-time steering vector (Fig. 1)

$$\boldsymbol{v} = \begin{pmatrix} \begin{pmatrix} e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_1} \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_2} \\ \vdots \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_N} \end{pmatrix} \\ e^{j2\pi f_d T_r} \begin{pmatrix} e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_1} \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_2} \\ \vdots \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_N} \end{pmatrix} \\ \vdots \\ e^{j2\pi f_d T_r (M-1)} \begin{pmatrix} e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_1} \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_2} \\ \vdots \\ e^{j2\pi\lambda_0^{-1}\hat{\boldsymbol{k}}\cdot\boldsymbol{r}_N} \end{pmatrix} \end{pmatrix} \tag{29}$$

The space–time steering vector

$$\boldsymbol{v} = \boldsymbol{v}(f_d, \phi, \theta) \tag{30}$$

is a function of the target's azimuth, elevation, and Doppler frequency. This vector is expressed conveniently using the

tensor or Kronecker product $\otimes$:

$$\boldsymbol{v}(f_d, \phi, \theta) = \boldsymbol{v}_t(f_d) \otimes \boldsymbol{v}_S(\phi, \theta) \tag{31}$$

where $\boldsymbol{v}_s$ is the spatial steering vector given in Eq. (21) and $\boldsymbol{v}_t$ is the temporal steering vector given in Eq. (26). The angle–Doppler response of a space–time steering vector is shown in the left-hand plot of Fig. 2.

The following properties of the tensor or Kronecker product are necessary to determine the properties of space–time signals. Let A and B be vectors or matrices of arbitrary dimension. Then

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix} \tag{32}$$

In addition to being bilinear, this product has the following properties:

$$\begin{aligned} (A \otimes B)(C \otimes D) &= AC \otimes BD && (C, D \text{ correct sizes}) \\ (A \otimes B)^{-1} &= A^{-1} \otimes B^{-1} && (\text{both } A, B \text{ invertible}) \\ (A \otimes B)^{H} &= A^{H} \otimes B^{H} \end{aligned}$$

Tapered Steering Vectors. As in finite-impulse response (*FIR*) filtering, the steering vectors used in practice may be tapered (or, synonymously, windowed, weighted, shaded) to reduce their sidelobe levels and thereby limit sensitivity to the presence of sidelobe targets or interference. Let $\boldsymbol{t} = (t_1, t_2, \dots, t_N)^T$ be the true target steering vector, and let $\boldsymbol{d} = (d_1, d_2, \dots, d_N)^T$ be the taper vector. Then the tapered steering vector is given by

$$\boldsymbol{v} = \boldsymbol{d} \odot \boldsymbol{t} = \begin{pmatrix} d_1 t_1 \\ d_2 t_2 \\ \vdots \\ d_N t_N \end{pmatrix} \tag{33}$$

where $\odot$ denotes the Schur–Hadamard (componentwise) product. Popular choices for $\boldsymbol{d}$ are the Chebyshev or Kaiser windows for Doppler filters or uniform linear array beamformers, or Taylor windows for uniform linear and planar array beamformers.

For space–time steering vectors of the form in Eq. (31), a space–time taper of the form

$$\boldsymbol{d} = \boldsymbol{d}_t \otimes \boldsymbol{d}_s \tag{34}$$

is used, where $\boldsymbol{d}_s$ is the spatial tapering vector and $\boldsymbol{d}_t$ is the temporal tapering vector. For a true space–time target steering vector of the form $\boldsymbol{t} = \boldsymbol{t}_t \otimes \boldsymbol{t}_s$, the tapered space–time steering vector is

$$\boldsymbol{v} = \boldsymbol{d} \odot \boldsymbol{t} = (\boldsymbol{d}_t \odot \boldsymbol{t}_t) \otimes (\boldsymbol{d}_s \odot \boldsymbol{t}_s) \tag{35}$$

The angle–Doppler response of a space–time steering vector with a 30 dB Chebyshev spatial taper and 50 dB Chebyshev temporal taper is shown in the right-hand plot in Fig. 1.

Dispersion. The spatial, temporal, and space–time steering vectors defined in this section all use the narrowband assumption, that is, Eq. (14). If the radar waveform has a fractional bandwidth greater than a few percent, then Eq. (13) must be used to define the steering vectors. The effect of wideband signals is called *dispersion* because, in accordance with Eqs. (11), (12), (13), a wideband signal from a single direction and a single Doppler frequency is equivalent to a sum of narrowband signals from many directions and Doppler frequencies. Let $\boldsymbol{v}(f)$ be the frequency-dependent steering vector (i.e., both the Doppler frequency f_d and the wavelength λ depend on f according to Eqs. (11) and (12), respectively). If the range of Doppler frequencies f_d over the signal bandwidth is small relative to this bandwidth and the matched-filter time offset τ is small relative to the signal duration, then the wideband steering vector $\boldsymbol{v}$ is approximated by the integral

$$\boldsymbol{v} = \int_{\text{bandwidth}} |S(f)|^2 \boldsymbol{v}(f)\, df \tag{36}$$

that is, $\boldsymbol{v}$ is a weighted sum of narrowband steering vectors across the signal bandwidth.

Signal-, Jammer-, and Clutter-to-Noise Ratios

The signals measured by a radar are composed of reflections from targets as well as undesired reflections from the environment (*clutter*), interference transmitted inadvertently or purposely from other electromagnetic sources (*jamming*), and noise from both receiver and sky. The signal-to-noise ratio (*SNR*), clutter-to-noise ratio (*CNR*), and jammer-to-noise ratio (*JNR*) are required to predict the radar's performance. Typically, these ratios are computed via the radar equation at the output of a matched filter coherently combining all antenna elements and radar pulses; however, it is also oftentimes convenient to express them on a per-element and per-pulse basis by dividing by the numbers of coherent elements (N) and coherent pulses (M). We shall compute the full coherent ratios.

The SNR is determined by the radar's average power P_t (W) coherently integrated over a CPI of T_{CPI} (s), transmitted with directivity G_t, reflected from a target of size σ (m^2), and received with an antenna whose effective area is $G_r\lambda^2/4\pi$. There is a $1/(4\pi R^2)$ propagation loss in both directions, and the noise power is kTF, where $k = 1.38 \times 10^{-23}$ W/Hz·K is Boltzmann's constant, T is the receiver temperature (typically about 290 K) and F is the receiver noise factor (typically about 10 dB). Combining these gains and losses into one equation, we have

$$\text{SNR} = \frac{P_t T_{CPI} G_t \sigma \lambda^2 G_r}{(4\pi)^3 R^4 kTF} \tag{37}$$

By a slight modification, Eq. (38) also provides the CNR:

$$\text{CNR} = \frac{P_t T_{CPI} G_t \sigma_c \lambda^2 G_r}{(4\pi)^3 R^4 kTF} \tag{38}$$

where σ_c is the effective area of a clutter patch appearing at the output of a rangegate, beam, and Doppler filter bank. A

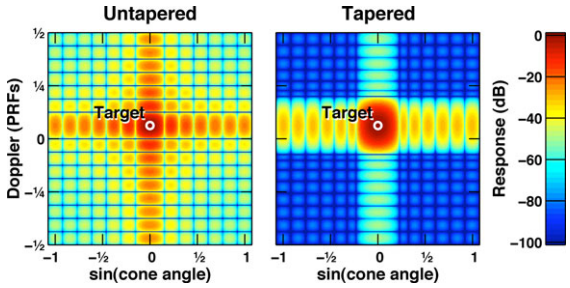

Figure 2. The response patterns of untapered and tapered space–time steering vectors. The radar parameters $N = 14$, $M = 16$ from the Mountaintop experimental radar system (31) are used. The cone angle ϕ'_t is defined after Eq. (24).

"constant gamma" model for the clutter is oftentimes used, in which the clutter area is assumed to be

$$\sigma_c = \gamma \tan\theta A_c \tag{39}$$

where γ is a constant ranging from about -15 dB (*light* clutter) to -3 dB (*heavy* clutter), and θ is the grazing angle of the radar's beam on the clutter patch of area A_c, assumed to be approximately rectangular. The downrange dimension is determined by the radar's range resolution $\Delta R = c/2B$ of Eq. (16), and the crossrange resolution is determined either by $R\,\Delta\phi$, the effective beamwidth at range R, or by $R\lambda/(2vT_{CPI})$, the effective Doppler resolution cell size at range R, whichever is smaller. Note that v is the radar platform's velocity. Also note that for steep grazing angles (θ near 90°), the approximation $\tan\theta A_c$ is inappropriate and must be replaced by the radar's effective beam area $R^2\,\Delta\theta\,\Delta\phi$ at range R.

The JNR is qualitatively different from the SNR and CNR because the jammer signal is generated externally to the radar and propagates only in a single direction from the jammer to the radar. Assuming that a broadband jammer transmits a waveform having power spectral density (*PSD*) Ψ_t (W/Hz) over a bandwidth B (Hz) during the radar's CPI of T_{CPI} (s), the JNR is

$$\text{JNR} = \frac{\Psi_t B T_{CPI} G_t \lambda^2 G_r}{(4\pi)^2 R^2 k T F} \tag{40}$$

where G_t is the jammer's transmit directivity and $G_r\lambda^2/4\pi$ is the radar's effective receive area. Note that there is only a single factor $1/(4\pi R^2)$ due to one-way propagation.

Noise, Jammer, and Clutter Covariance Matrices

The interference plus noise, represented by the random vector $\boldsymbol{n}$, can be decomposed into three statistically independent components:

$$\boldsymbol{n} = \boldsymbol{n}_n + \boldsymbol{n}_j + \boldsymbol{n}_c \tag{41}$$

—the noise component $\boldsymbol{n}_n$, the jammer component $\boldsymbol{n}_j$, and the clutter component $\boldsymbol{n}_c$. The interference-plus-noise covariance matrix $\boldsymbol{R}$ is defined as

$$\boldsymbol{R} = E[\boldsymbol{n}\boldsymbol{n}^H] \tag{42}$$

Because these three sources are statistically independent, the interference-plus-noise covariance matrix takes the form

$$\boldsymbol{R} = \boldsymbol{R}_n + \boldsymbol{R}_j + \boldsymbol{R}_c \tag{43}$$

where $\boldsymbol{R}_n = E[\boldsymbol{n}_n\boldsymbol{n}^H{}_n]$ is the noise covariance, $\boldsymbol{R}_j = E[\boldsymbol{n}_j\boldsymbol{n}^H{}_j]$ is the jammer covariance, and $\boldsymbol{R}_c = E[\boldsymbol{n}_c\boldsymbol{n}^H{}_c]$ is the clutter covariance.

Noise Covariance. The receiver and sky noise are spatially and temporally uncorrelated; therefore, the covariance matrix of the noise at the matched filter outputs is a diagonal matrix:

$$\boldsymbol{R}_n = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{MN}^2) \tag{44}$$

where $\sigma^2{}_i$ is the noise power at the ith matched filter. If all receivers have identical gains, as is often the case, then the noise covariance is a multiple of the identity matrix $\boldsymbol{I}$:

$$\boldsymbol{R}_n = \sigma_n^2 \boldsymbol{I} \tag{45}$$

where $\sigma^2{}_n$ is the noise power per element per pulse. For space–time signals, it is also possible to express the noise in the form

$$\boldsymbol{n} = \boldsymbol{n}_t \otimes \boldsymbol{n}_s \tag{46}$$

where $\boldsymbol{n}_s$ is the spatial noise and $\boldsymbol{n}_t$ is the temporal noise. The space–time noise covariance is therefore expressed as

$$\boldsymbol{R}_n = \boldsymbol{R}_t \otimes \boldsymbol{R}_s \tag{47}$$

where $\boldsymbol{R}_s = E[\boldsymbol{n}_s\boldsymbol{n}^H{}_s]$ and $\boldsymbol{R}_t = E[\boldsymbol{n}_t\boldsymbol{n}^H{}_t]$.

Jammer Covariance

Jammer interference is typically uncorrelated temporally because its waveform spans the radar's full operational bandwidth, but correlated spatially because it propagates from a specific direction (Fig. 3). It is this spatial coherence that is used to null jammer interference. A broadband jammer at the output of the radar's match filters is represented by the waveform

$$\boldsymbol{n}_j = \boldsymbol{n}_t \otimes \boldsymbol{v}_j \tag{48}$$

where $\boldsymbol{v}_j$ is the spatial steering vector corresponding to the jammer's position, and $\boldsymbol{R}_t = E[\boldsymbol{n}_t\boldsymbol{n}^H{}_t]$ is the temporal covariance matrix of the jammer waveform; typically $\boldsymbol{R}_t =$

$\sigma^2_j \boldsymbol{I}$, where σ^2_j/σ^2_n is the JNR per element per pulse. For the case of several independent jammers,

$$\boldsymbol{n}_j = \sum_{i=1}^{J} \boldsymbol{n}_i \otimes \boldsymbol{v}_i \tag{49}$$

where the $\boldsymbol{n}_i$ ($i = 1, \ldots, J$) are the temporal components of the jammer signal, and the $\boldsymbol{v}_i$ ($i = 1, \ldots, J$) are the spatial steering vectors corresponding to the jammers' positions. Assuming that the jammers are statistically independent and that $E[\boldsymbol{n}_i \boldsymbol{n}^H{}_i] = \sigma^2{}_i \boldsymbol{I}$, the jammer covariance matrix is

$$\boldsymbol{R}_j = \boldsymbol{I} \otimes \boldsymbol{J}\boldsymbol{J}^H \tag{50}$$

where

$$\boldsymbol{J} = (\sigma_1 \boldsymbol{v}_1, \sigma_2 \boldsymbol{v}_2, \ldots, \sigma_J \boldsymbol{v}_J) \tag{51}$$

Note that for spatial-only signals, Eq. (51) reduces to $\boldsymbol{R}_j = \boldsymbol{J}\boldsymbol{J}^H$.

The rank of $\boldsymbol{R}_j$ given in Eq. (51) is MJ; as will be discussed in the subsection "Adaptive Degrees of Freedom and Nulling Performance," an adaptive weight vector requires at least this many degrees of freedom to null those jammers.

Clutter Covariance. Clutter interference is best understood by representing it as the sum of many infinitesimal reflections from discrete points on the ground. Reflection from a discrete point, called a *clutter discrete,* is targetlike because it appears from a specific direction and has a fixed Doppler shift. If the radar platform is airborne, there is a relationship between the clutter's direction and Doppler frequency:clutter in front of the radar is approaching and has a positive Doppler shift, while clutter behind the radar is receding and has a negative Doppler shift. This angle-dependent Doppler frequency is given by the expression

$$f_d = 2vT_r \lambda \cos \theta_c \sin \phi'_c \tag{52}$$

where v is the radar platform's speed, θ_c is the lookdown angle to the clutter patch, and ϕ'_c is the complementary angle between the clutter patch and the radar platform's velocity. This locus of angle–Doppler positions is called the *clutter ridge* (Fig. 3); oftentimes the mainlobe clutter is shifted to zero Doppler frequency using TACCAR (3). Matched filtering localizes clutter reflections to a fixed-range ring on the ground; if the radar is range-unambiguous (R_{max} > radar horizon), clutter interference may be assumed to come only from reflections around this range ring. If the radar is range-ambiguous (R_{max} < radar horizon), clutter interference comes from the set of ambiguous range rings. In either case, the clutter covariance matrix for a single range ring is determined by integrating the infinitesimal return from the clutter patch between the azimuth ϕ and $\phi + d\phi$ over the full ring:

$$\boldsymbol{n}_c = \oint \boldsymbol{v}_t(\phi) \otimes \boldsymbol{v}_s(\phi) A(\phi)\, d\gamma(\phi) \tag{53}$$

where $\boldsymbol{v}_t(\phi)$ is the temporal steering vector of the clutter at azimuth ϕ, $\boldsymbol{v}_s(\phi)$ is the spatial steering vector of the clutter at azimuth ϕ, $A(\phi)$ is the antenna pattern, and $d\gamma(\phi)$ is the infinitesimal clutter reflectivity. It is assumed that this reflectivity is a stochastic process with orthogonal increments (15), so that

$$E[d\gamma^*(\phi)\, d\gamma(\phi')] = \sigma^0(\phi)\delta(\phi - \phi')\, d\phi \tag{54}$$

that is, clutter reflectivity at different points on the ground is independent. From Eqs. (54) and (55) the space–time clutter covariance matrix is

$$\boldsymbol{R}_c = \oint |A(\phi)|^2 \sigma^0(\phi) [\boldsymbol{v}_t(\phi)\boldsymbol{v}_t^H(\phi)] \otimes [\boldsymbol{v}_s(\phi)\boldsymbol{v}_s^H(\phi)]\, d\phi \tag{55}$$

Equation (55) may be evaluated using Riemann summation.

In the case of uniform linear arrays, this integral may be evaluated very efficiently using a Fourier–Bessel series and the Toeplitz–block-Toeplitz structure (16) of $\boldsymbol{R}_c$ (17) (Fig. 4). The rank of the clutter covariance matrix $\boldsymbol{R}_c$ in this case when the array's axis is aligned with the velocity vector is approximated by Brennan's rule (12),

$$\operatorname{rank} \boldsymbol{R}_c \approx N + \beta(M - 1) \tag{56}$$

where

$$\beta = 2vT_r/d \tag{57}$$

(in which v is the radar platform's velocity, T_r is the PRI, and d is the interelement spacing) is the number of half interelement spacings traversed per PRI. If β is an integer, Brennan's rule holds exactly (12). As will be discussed in the subsection "Adaptive Degrees of Freedom and Nulling Performance," an adaptive weight vector requires at least this many degrees of freedom to null the clutter.

The PSD of radar clutter measured at White Sands, NM, and of modeled radar clutter is illustrated in Fig. 5. The angle–Doppler dependence of the clutter is evident. Note that these PSDs are computed from the clutter output power of tapered space–time steering vectors at a particular angle and Doppler shift. The clutter itself is localized to a locus along the broad ridges seen in these PSD estimates.

Dispersive Effects on Covariance. In addition to the dispersion of a target response considered in the sub-section "Dispersion" above, the dispersion of the interference must also be considered. Except for the unknown target amplitude, the matched filter output of Eq. (13) assumes a nonrandom radar return. If the radar measurements are due to a broadband jammer or clutter, the radar return $s_r(t)$ is more accurately modeled as a stationary stochastic process with PSD $\Psi(f)$, i.e., $s_r(t)$ has the spectral decomposition

$$s_r(t) = \int_{-\infty}^{\infty} e^{j2\pi ft}\, d\psi(f), \qquad E|d\psi(f)|^2 = \Psi(f)\, df \tag{58}$$

where $\psi(f)$ is a stochastic process with orthogonal increments. For broadband jammers, $\Psi(f)$ is the jammer signal's power spectrum, and for clutter in the absence of intrinsic clutter motion $\Psi(f)$ is modeled as the transmit waveform's power spectrum $|S(f)|^2$. From Eqs. (4), (43), and (59), the covariance matrix $\boldsymbol{R}$ of the matched filter outputs due to a

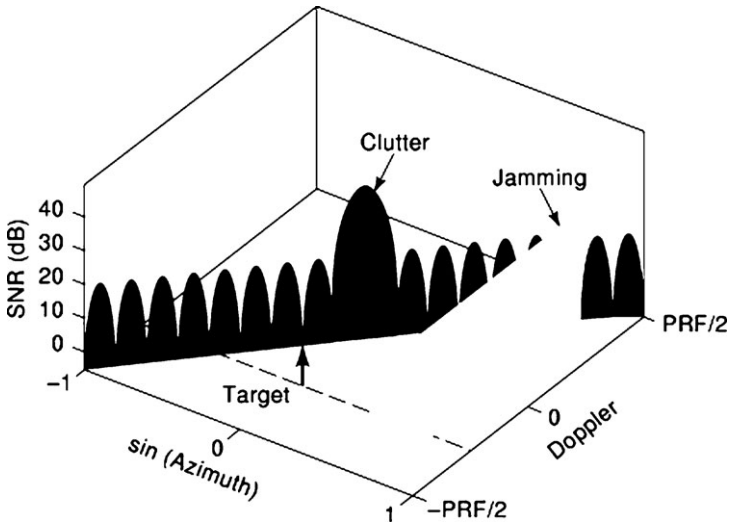

Figure 3. Clutter and jammer ridges in angle–Doppler space.The mainlobe clutter at a cone angle of 0° has been shifted to zero Doppler frequency using TACCAR.

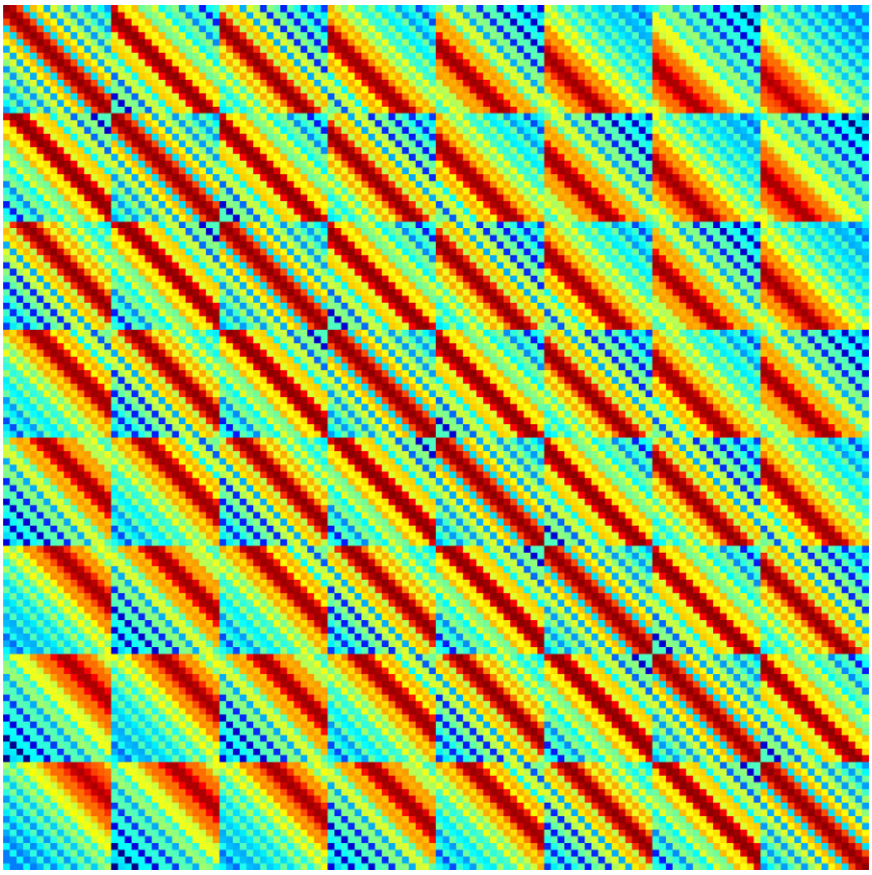

Figure 4. Structure of an ideal Toeplitz–block-Toeplitz clutter covariance matrix ($M = 8, N = 16, \beta = 2$). The log magnitudes of the elements of $\boldsymbol{R}_c$ are shown using a relative color scale.

broadband waveform is given by

$$\boldsymbol{R} = \int_{-\infty}^{\infty} |S(f)|^2 \Psi(f) \boldsymbol{R}(f)\, df \tag{59}$$

where $\boldsymbol{R}(f)$ is the narrowband interference covariance at frequency f. For jammer-only covariance matrices, $\boldsymbol{R}(f)$ takes the form of Eq. (51), and for clutter-only covariance matrices, $\boldsymbol{R}(f)$ takes the form of Eq. (56). Therefore, the wideband covariance matrix is a weighted sum of narrowband covariance matrices. Because of this fact, the rank of a wideband interference covariance matrix may exceed that of a narrowband interference matrix, which has an

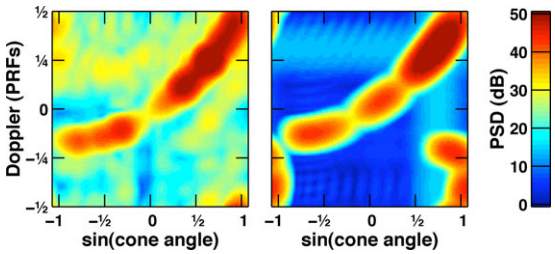

Figure 5. Measured and modeled clutter interference PSD (relative to noise) versus Doppler frequency and angle. The measured PSD on the left is estimated using data from the Mountaintop database (31), collected on March 9, 1994 at White Sands, NM. The modeled PSD on the right is obtained using Eq. (56) assuming the constant-gamma clutter model in the preceding subsection.

effect on the adaptive nulling performance, as discussed in the subsection "Adaptive Degrees of Freedom and Nulling Performance" below.

Equation (60) also quantifies the effect of electronic countermeasures (*ECM*) on radar performance. Because radars may select to transmit an instantaneous frequency f_0 from a range of frequencies in the radar's operational bandwidth, a broadband jammer must spread its energy $\Psi(f)$ over this operational bandwidth to ensure that the product $|S(f)|^2\Psi(f)$ is large enough to interfere with target signals regardless of the choice of f_0. The radar's electronic counter-countermeasure (*ECCM*) is to use spatial diversity to adaptively null the jammer interference in the radar returns.

Intrinsic Clutter Motion. Clutter motion is caused by windblown trees or ocean waves whose velocities are sensed by the radar. Radar reflections from ocean waves can exhibit complicated behavior (18); however, simple, empirically supported models exist for the Doppler spectra of windblown ground clutter. For CPI lengths less than a second or so, the windblown clutter's random Doppler component is constant over the CPI. Therefore, assume that windblown ground clutter imparts a random Doppler shift to all pulses in the CPI, modeled as the random space–time vector

$$\boldsymbol{n}_{\text{icm}} = \begin{pmatrix} x(0) \\ x(T_r) \\ \vdots \\ x((M-1)T_r) \end{pmatrix} \otimes \mathbf{1}_N \tag{60}$$

where $x(t)$ is a wide-sense stationary stochastic process with spectral decomposition

$$x(t) = \int_{-\infty}^{\infty} e^{j2\pi ft}\, d\xi(f), \qquad E|d\xi(f)|^2 = \Xi(f)\, df \tag{61}$$

$\Xi(f)$ is the Doppler power spectrum of $x(t)$, and $\xi(f)$ has orthogonal increments. The Kronecker product with the N-by-1 vector $\mathbf{1}_N = (1, 1, \ldots, 1)^{\mathrm{T}}$ corresponds to the fact that the random Doppler shift is sensed identically on all N elements. The matched-filter output of the clutter interference for pulses $m = 0, 1, \ldots, M-1$ may be modeled as

$$\boldsymbol{n}_{\text{c}} = \boldsymbol{n}_{\text{icm}} \odot \boldsymbol{n}_{\text{no icm}} \tag{62}$$

where $\boldsymbol{n}_{\text{noicm}}$ is the clutter interference vector without intrinsic clutter motion (*ICM*), defined in Eq. (54). The clutter reflectivity defined by $\gamma(\phi)$ in Eq. (54) and the ICM defined by $x(t)$ in Eq. (62) are assumed to be independent stochastic processes.

The clutter covariance matrix with ICM is

$$\boldsymbol{R}_{\text{c}} = E[\boldsymbol{n}_{\text{c}}\boldsymbol{n}_{\text{c}}^{\text{H}}] = \boldsymbol{R}_{\text{icm}} \odot \boldsymbol{R}_{\text{no icm}} \tag{63}$$

where $\boldsymbol{R}_{\text{noicm}}$ is the clutter covariance matrix without ICM, defined in Eq. (56), and $\boldsymbol{R}_{\text{icm}} = E[\boldsymbol{n}_{\text{icm}}\boldsymbol{n}^{\text{H}}_{\text{icm}}]$. That is, to include the effects of internal clutter motion, we need only to compute the clutter covariance matrix without ICM, then multiply it componentwise by the matrix $\boldsymbol{R}_{\text{icm}}$. This matrix is given by the time samples of the ICM's autocorrelation function

$$X(\tau) = \int_{-\infty}^{\infty} \Xi(f) e^{j2\pi f\tau}\, df \tag{64}$$

From Eqs. (62), (61), and (65),

$$\boldsymbol{R}_{\text{icm}} = \begin{pmatrix} X(0) & X(-T_r) & \ldots & X(-(M-1)T_r) \\ X(T_r) & X(0) & \ldots & X(-(M-2)T_r) \\ \vdots & \vdots & & \vdots \\ X((M-1)T_r) & X((M-2)T_r) & \ldots & X(0) \end{pmatrix} \otimes \mathbf{1}_{N\times N} \tag{65}$$

where $\mathbf{1}_{N\times N} = \mathbf{1}_N\mathbf{1}^{\mathrm{T}}{}_N$ is an N-by-N matrix whose elements are all unity. Note that the matrix on the left-hand side of the Kronecker product is Toeplitz.

Billingsley's (19) empirical model of windblown ground clutter applicable from VHF to X-band frequencies is

$$X(\tau) = \frac{1}{r+1}\left(r + \frac{1}{1+(\omega/\beta)^2}\right), \qquad \omega = 2\pi\lambda_0^{-1}\tau \tag{66}$$

where

$$r = \frac{10^{6.32}}{v_{\text{w}}^{1.55} f_0^{1.21}} \tag{67}$$

is the empirically derived ratio of dc to ac power, v_{w} is the wind speed in miles per hour, f_0 is the center frequency in megahertz, λ_0 is the wavelength in meters, and β is an exponential shape parameter loosely dependent on the wind speed. For example, in breezy conditions, v_{w} is 7 mi/h to 15 mi/h (3 m/s to 7 m/s) and a typical value for β is 8 $(\text{m/s})^{-1}$.

Another choice for $X(\tau)$ is Barlow's Gaussian model (2)

$$X(\tau) = e^{-(\sigma_{\text{c}}\tau)^2/2} \tag{68}$$

where

$$\sigma_{\text{c}} = 2\sigma_{\text{v}}/\lambda \tag{69}$$

is the standard deviation (in hertz) of the clutter power spectrum $\Xi(f)$, σ_v is the standard deviation (in meters per second) of the intrinsic clutter motion, and λ is the transmitted wavelength.

The Adaptive Weight Vector

Adaptive radar processing consists fundamentally of three steps: (1) compute an adaptive weight vector $\boldsymbol{w}$; (2) apply $\boldsymbol{w}$ to the vector $\boldsymbol{z}$ of radar measurements, and compare the output power (Fig. 1)

$$z = |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}|^2 \tag{70}$$

with a threshold to determine the presence of a target, and (3) estimate the target's speed and bearing. All three steps are dependent upon the adaptive weight vector $\boldsymbol{w}$.

The adaptive weight vector's purpose is to provide coherent gain on any targets present in the measurement $\boldsymbol{z}$ while simultaneously canceling any interference (Fig. 6). If the interference-plus-noise covariance matrix $\boldsymbol{R}$ were known, the optimum choice (relative to the criteria given in the section "Optimum Adaptive Processing") for the adaptive weight vector $\boldsymbol{w}$ to detect a target with steering vector $\boldsymbol{v}$ would be

$$\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v} \tag{71}$$

This equation is the consequence of various optimality criteria discussed in the section "Optimum Adaptive Processing" below. Because the precise jammer and clutter characteristics are unknown, the interference covariance matrix $\boldsymbol{R}$ is also unknown, and is typically estimated using the sample covariance matrix

$$\hat{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{z}_k\boldsymbol{z}_k^{\mathrm{H}} \tag{72}$$

where $\boldsymbol{z}_k$ ($k = 1, \ldots, K$) are independent measurements of the interference in the absence of target returns. The adaptive weight vector becomes

$$\hat{\boldsymbol{w}} = \hat{\boldsymbol{R}}^{-1}\boldsymbol{v} \tag{73}$$

Much of the performance of an adaptive radar can be predicted using the known covariance adaptive weight vector $\boldsymbol{w}$ of Eq. (72), and the detection performance using the estimated covariance adaptive weight vector $\hat{\boldsymbol{w}}$ of Eq. (74) can be predicted by an extension of classical matched-filter detection theory provided in the section "Adaptive Detection" below.

Performance Metrics

To compare the performance of different adaptive algorithms with each other, as well as with nonadaptive approaches, it is necessary to use some standard benchmarks. Important radar performance metrics are the probability of detecting a target of a given size, the probability (or rate) of declaring a false alarm, and the accuracy with which target speed and/or bearing may be measured. Useful intermediate quantities for the probability of detection are the signal-to-interference-plus-noise ratio (*SINR*) and SINR loss, defined below. Finally, the filter response itself is important—it should have a distinct mainlobe that is as narrow as possible as well as low sidelobes.

Probabilities of Detection and False Alarm. A target is declared to be present if the output power z in Eq. (71) exceeds a set threshold τ. Therefore, the probability of detection (*PD*) and probability of false alarm (*PFA*) both depend on the statistics of z in the presence and absence of a target. If $f(z)$ is the probability density of this output power, then the PD and PFA are given by the equations

$$\mathrm{PD} = \int_{\tau}^{\infty} f(z|\text{target present})\,dz \tag{74}$$

$$\mathrm{PFA} = \int_{\tau}^{\infty} f(z|\text{target absent})\,dz \tag{75}$$

In the section "Optimum Adaptive Processing" below, it is shown that the adaptive weight vector of Eq. (72) optimizes the probability of detection for a fixed PFA. Example PD curves are shown in Fig. 13.

Target Parameter Estimation Accuracy. The accuracy with which a radar can determine a target's bearing and speed is bounded by the radar's aperture in space and/or time. Given a vector of radar measurements $\boldsymbol{z}$, the root-mean-square (*rms*) accuracy of an estimate is defined to be

$$\text{rms accuracy} = (E|\text{estimate}(\boldsymbol{z}) - E[\boldsymbol{z}]|^2)^{1/2} \tag{76}$$

SINR and SINR Loss. Given a true target steering vector $\boldsymbol{t}$, the SINR at the output of an adaptive filter weight using the weight vector $\boldsymbol{w}$ is

$$\mathrm{SINR} = \frac{|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{t}|^2}{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w}} \tag{77}$$

It is important to note that the vector $\boldsymbol{t}$ represents the steering vector of a target in the radar returns, whereas the vector $\boldsymbol{v}$ will be used to represent an arbitrary steering vector, which may be the true target vector $\boldsymbol{t}$, a hypothesized steering vector, or a tapered steering vector. Compared with the maximum SNR achievable in the absence of interference, the SINR loss due to adaptivity is given by the equation

$$\text{SINR loss} = \frac{|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{t}|^2}{(\boldsymbol{t}^{\mathrm{H}}\boldsymbol{R}_n\boldsymbol{t})(\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w})} \tag{78}$$

Example SINR loss curves for several STAP algorithms are shown in Fig. 10 below.

Minimum Detectable Velocity. The width of the SINR loss notch near mainlobe clutter (as seen in Fig. 11) determines the lowest velocity detectable by the radar.

Response Patterns and Sidelobe Levels. The response

$$F(f_d, \phi, \theta) = |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}(f_d, \phi, \theta)|^2 \tag{79}$$

of a fixed adaptive weight vector to steering vectors at various Doppler frequencies and angles defines the weight vector's response pattern and sidelobe levels. The response

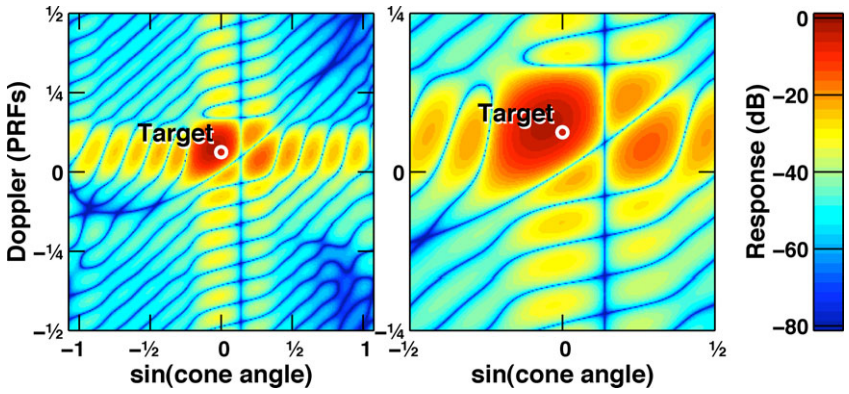

Figure 6. Space–time adaptive weight vector response pattern versus Doppler frequency and angle. The response is computed using Eq. (72), where $\boldsymbol{v}$ is a space–time steering vector pointed to the target location shown, $\boldsymbol{R} = \boldsymbol{R}_n + \boldsymbol{R}_j + \boldsymbol{R}_c$ is a space–time covariance matrix, $\boldsymbol{R}_n = \boldsymbol{I}$, $\boldsymbol{R}_j$ is a jammer covariance given by Eq. (51) for a jammer at a cone angle of 8°, and $\boldsymbol{R}_c$ is a clutter covariance using parameters similar to those shown in Figs. 2 and 4. A space–time null in the directions of the clutter and jammer interference is visible. A close-up of the pattern is shown on the right. Compare this adaptive weight response with the nonadaptive steering vector response of Fig. 1.

patterns of a tapered and untapered space–time steering vector are shown in Fig. 2; the patterns of several adaptive weight vectors are shown in Figs. 6 and 12.

STATISTICAL MODEL FOR ADAPTIVE RADAR

All fundamental performance limits of adaptive radar are derived from a statistical model of the radar measurements. Such a model provides an optimum detection strategy, predicts detection performance (i.e., probability of detection and false alarm rates), and predicts estimation performance (i.e., the best achievable accuracy). An additive Gaussian model is the prevailing choice, because it is for the most part physically accurate and because it yields tractable solutions to detection and estimation problems. Of course, a Gaussian model is not completely accurate; however, the detection and estimation techniques derived from it have been shown to be robust within a wide class of different models. See Ref. 20a–f.

Adaptive Radar Hypothesis (Known Covariance)

The radar detection problem is posed as distinguishing between the hypotheses

$$\begin{aligned} \mathrm{H}_0&: \quad \boldsymbol{z} = \boldsymbol{n} \qquad \text{(target absent)} \\ \mathrm{H}_1&: \quad \boldsymbol{z} = a\boldsymbol{v} + \boldsymbol{n} \qquad \text{(target present)} \end{aligned} \tag{80}$$

The vector $\boldsymbol{z}$ is assumed to be distributed as a complex Gaussian random variable with distribution

$$f(\boldsymbol{z}|a) = \frac{1}{\pi^N \det \boldsymbol{R}} \exp[-(\boldsymbol{z} - a\boldsymbol{v})^{\mathrm{H}} \boldsymbol{R}^{-1} (\boldsymbol{z} - a\boldsymbol{v})] \tag{81}$$

where the covariance $\boldsymbol{R} = E[\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}]$ and mean direction $\boldsymbol{v}$ are assumed to be known, the complex amplitude a is assumed to be unknown and possibly random, and N is the dimension of the vector $\boldsymbol{z}$. For space–time problems, N must be replaced by MN—the product of the number of coherent pulses M and the number of coherent antenna elements N. The vector $\boldsymbol{n}$ refers to both interference and noise, which appear additively. The precise structure of the interference and noise encountered in adaptive radar is considered in the subsection "Noise, Jammer, and Clutter Covariance Matrices" above. Of course, the whole point of adaptive radar is that we do not know the interference covariance beforehand, violating a key assumption above, and must therefore estimate it. Nevertheless, most of the results from the known-covariance case are directly applicable to the unknown case.

Adaptive Radar Hypothesis (Unknown Covariance)

In practice the interference-plus-noise covariance matrix $\boldsymbol{R}$ in Eq. (82) is unknown and must be estimated. Fortunately for the radar problem, the large number of range gates to be tested for the presence of targets usually provides a sufficient number of samples that may be assumed to be identically distributed. It is important that these range gates be free of target reflections, because such targets will be (partially) nulled by the adaptive weight vector. Assume that there is a vector of radar measurements $\boldsymbol{z}_0$ from the range gate of interest, and K other vectors of radar measurements $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K$ from nearby range gates. The vector $\boldsymbol{z}_0$ is called the primary data, and the collection of vectors $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K$ are called the secondary data. Further assume that the secondary data are independent of the primary data, that they have zero mean (i.e., there are no target reflections in the secondary data), and that they have the same covariance matrix as the primary data. This final assumption ensures that we can use the secondary data to estimate the interference found in the primary data.

By the assumptions of independent and identically distributed data and Eq. (82), the primary and secondary data have the joint probability distribution

$$f(\boldsymbol{z}_0: \boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_K|a) = f(\boldsymbol{z}_0|a) \prod_{k=1}^{K} f(\boldsymbol{z}_k|a = 0) \tag{82}$$

$$= f(\boldsymbol{z}_0|a) \cdot [(\pi^N \det \boldsymbol{R})^{-1} e^{-\mathrm{tr}\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}}]^K \tag{83}$$

where $\hat{\boldsymbol{R}}$ is the sample covariance matrix (*SCM*) given in Equation (73). The SCM is the maximum likelihood estimate of $\boldsymbol{R}$ given the secondary data.

The proof of this last fact is a short computation. The log likelihood of the secondary data is $l(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K|\boldsymbol{R}) = -K(\mathrm{tr}\,\boldsymbol{R}^{-1}\hat{\boldsymbol{R}} + \log\det\boldsymbol{R}) + \text{constants}$. From the first-order terms of the Taylor series

$$\mathrm{tr}\,(\boldsymbol{R}+\boldsymbol{\Delta})^{-1}\boldsymbol{A} = \mathrm{tr}\,\boldsymbol{R}^{-1}\boldsymbol{A} - \mathrm{tr}\,\boldsymbol{\Delta}\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R}^{-1} + \mathrm{tr}\,(\boldsymbol{\Delta}\boldsymbol{R}^{-1})^2\boldsymbol{A}\boldsymbol{R}^{-1} + \cdots$$

$$\log\det(\boldsymbol{R}+\boldsymbol{\Delta}) = \log\det\boldsymbol{R} + \mathrm{tr}\,\boldsymbol{\Delta}\boldsymbol{R}^{-1} - \tfrac{1}{2}\,\mathrm{tr}\,(\boldsymbol{\Delta}\boldsymbol{R}^{-1})^2 + \cdots$$

the derivative of the log likelihood function with respect to $\boldsymbol{R}$ is

$$\frac{\partial}{\partial\boldsymbol{R}}\,l(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_K|\boldsymbol{R}) = K(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\boldsymbol{R}^{-1} - \boldsymbol{R}^{-1})$$

Setting this derivative to zero and solving for $\boldsymbol{R}$ yields the maximum likelihood estimate $\boldsymbol{R} = \hat{\boldsymbol{R}}$. The second-order terms of the Taylor series above establish that this point is indeed a maximum.

Fluctuating Target Statistics

The statistics of the unknown amplitude a strongly affect the radar's detection performance. The magnitude and phase of $a = re^{j\phi}$ are treated separately. In almost all cases, the phase component $e^{j\phi}$ is assumed to be unknown and distributed uniformly on the unit circle from $-\pi$ to π. Conditioning the density of Eq. (82) over the random phase component (21, 22) yields the probability of radar measurements conditioned on the unknown target magnitude r:

$$f(\boldsymbol{z}|r) = \int f(\boldsymbol{z}|a)f_\Phi(\varphi)\,d\varphi = \frac{1}{2\pi}\oint f(\boldsymbol{z}|a)\,d\varphi = \frac{1}{\pi^N\det\boldsymbol{R}}\exp(-\boldsymbol{z}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z} - r^2\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v})\,I_0(2r|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}|) \tag{84}$$

where I_0 is the modified Bessel function of the first kind with order 0—a monotonic function for positive real numbers.

There are several standard statistical models for the target magnitude r, each applicable for specific targets. The simplest is Marcum's (21) nonfluctuating target model, in which r is fixed. Swerling (23) introduced a class of target fluctuation models based upon the complex $\chi^2{}_\nu$ probability distribution with ν complex degrees of freedom:

$$f_{\mathrm{R}}(r) = (\nu/\bar{r})f_{\chi^2_\nu}(\nu r/\bar{r}) \tag{85}$$

$$f_{\chi^2_\nu}(z) = \Gamma(\nu)^{-1}z^{\nu-1}e^{-z} \tag{86}$$

where $\bar{r} = E\,r$ is the mean amplitude. In Swerling model I, the target magnitude is assumed to be fixed over a limited dwell of several CPIs, but fluctuates according to a $\chi^2{}_1$ model ($\hat{r}^{-1}e^{-r/\hat{r}}$) from scan to scan, where scans may occur on the order of tens of seconds apart. For Swerling model II, the target is assumed to fluctuate every CPI according to a $\chi^2{}_1$ model. Swerling called this "pulse-to-pulse" fluctuation, but in the context of pulse Doppler radar, it is appropriate to regard it as "CPI-to-CPI" fluctuation. Swerling I is appropriate for determining the probability of detecting a target over several scans, whereas Swerling II is appropriate for determining the probability of detecting the target over several CPIs. Of course, if one CPI is used per scan, there is no difference between the two. Swerling models III and IV are identical except that a $\chi^2{}_2$ model is used. Others have suggested using more degrees of freedom for some targets, based on experimental data. Sometimes Marcum's nonfluctuating model is called a Swerling model with infinite degrees of freedom because $f_{\mathrm{R}}(r) \to \delta(r - \bar{r})$ as $\nu \to \infty$.

OPTIMUM ADAPTIVE PROCESSING

The goal of adaptive radar is to use spatial and temporal coherence to mitigate the clutter and jamming interference and achieve radar performance limited only by noise. This is accomplished by constructing an optimum filter that incorporates a steering vector and the interference and noise covariance. This filter is optimum in several senses; perhaps the most compelling is that under some ideal assumptions about the signal, interference, and noise statistics, it optimizes the probability of detecting a target given a fixed probability of false alarm.

Detection Optimality

The Neyman–Pearson criterion of maximizing the probability of detection given a fixed probability of false alarm yields the test (24)

$$\frac{f(\boldsymbol{z}|\text{target present})}{f(\boldsymbol{z}|\text{target absent})} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \text{threshold} \tag{87}$$

for some fixed threshold that is determined by the probability of false alarm. The notation means: if the likelihood ratio on the left exceeds the given threshold, decide that a target is present; otherwise, decide that no target is present. Inserting Eq. (86) with unspecified target magnitude r for H_1 and $r = 0$ for H_0, and taking the logarithm (which gives the log likelihood function on the left), this test reduces to the equivalent test

$$\log I_0(2r|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}|) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \text{threshold} \tag{88}$$

after absorbing the constant term $r^2\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}$ into the threshold on the right. Because $\log(\cdot), I_0(\cdot), (\cdot)^2$, and scaling by positive constants are all monotonic functions for positive real arguments, this test may be reduced to the form

$$\frac{|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}|^2}{\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \tau \tag{89}$$

for some positive threshold τ. This is called the matched-filter (*MF*) test. We shall use this form of the test even when the steering vector $\boldsymbol{v}$ differs from the true target steering vector $\boldsymbol{t}$, i.e., when $E[\boldsymbol{z}|g] = a\boldsymbol{t}$. The effect of this steering vector mismatch on detection performance is given in the section "Adaptive Detectors." The squared (power) form of this test is used instead of the unsquared (voltage) form for consistency with the SINR analysis in the next subsection. Note that the exact threshold must change for each

of the equivalent tests of Eqs. (89), (90), and (91) despite the fact that this is hidden typographically. One important quality of the MF test is that it has a constant false-alarm rate (*CFAR*), that is, its probability of false alarm is independent of the interference-and-noise scenario defined by $\boldsymbol{R}$. This is an important practical feature because it allows the detection threshold to be set without regard to a specific interference scenario.

For a known covariance matrix, the probabilities of detection and false alarm are given by the classical matched-filter formulae. Assume that $\boldsymbol{z}$ is a Gaussian random variable with covariance $\boldsymbol{R}$ and mean $a\boldsymbol{t}$ (the complex target amplitude $a = re^{j\phi}$ multiplied by the true target steering vector $\boldsymbol{t}$). By using the whitened coordinates $\tilde{z} = \boldsymbol{R}^{-1/2}\boldsymbol{z}$, $\tilde{v} = \boldsymbol{R}^{-1/2}\boldsymbol{v}$, and $\tilde{t} = \boldsymbol{R}^{-1/2}\boldsymbol{t}$, then, choosing a unitary matrix $\boldsymbol{Q}$ such that $\alpha\boldsymbol{e}_1 = \boldsymbol{Q}\tilde{v}$, where $\boldsymbol{e}_1$ is the canonical unit vector $(1, 0, \ldots, 0)^{\mathrm{T}}$ and α is a complex constant whose magnitude is the length of $\tilde{v}$, the left-hand side of Eq. (91) is recognized as a noncentral complex χ^2 statistic with one complex degree of freedom and noncentrality parameter

$$s = r^2 \frac{|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{t}|^2}{\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}} \tag{90}$$

which is the SINR of the target. Therefore, the PFA and PD are given by the (known covariance) formulae

$$\mathrm{PFA}_{\mathrm{MF}} = \int_{\tau}^{\infty} f_{\chi_1^2}(z|\mathrm{H}_0)\,dz = e^{-\tau} \tag{91}$$

$$\begin{aligned}\mathrm{PD}_{\mathrm{MF}} &= \int_{\tau}^{\infty} f_{\chi_1^2}(z|\mathrm{H}_1)\,dz \\ &= \int_{\tau}^{\infty} e^{-s-z} I_0(2\sqrt{sz})\,dz = Q(\sqrt{2s}, \sqrt{2\tau})\end{aligned} \tag{92}$$

where $Q(\alpha, \beta)$ is the Marcum Q-function (21, 25). The threshold τ is determined by Eq. (93) for a given PFA. Because of the large number of rangegates, Doppler bins, and beam directions to be examined for targets, radar PFAs typically lie in the range 10^{-6} to 10^{-12}. Conditioning the PD over a target fluctuation model for r yields detection probabilities for fluctuating targets. These known-covariance probabilities must be modified as described in the subsection "Adaptive Detectors" below when sample covariance matrices are used.

The Neyman–Pearson optimum detection test of Eq. (91) may be interpreted in the following way. First, construct an adaptive filter

$$\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v} \tag{93}$$

second, compute

$$\mathrm{SINR} = \frac{|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}|^2}{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w}} \tag{94}$$

at the output of this adaptive filter, and third, compare the SINR with a fixed threshold, yielding a desired PFA. Equation (83) is called the adaptive filter equation; the elements of $\boldsymbol{w}$ are called adaptive weights, and the number of elements in $\boldsymbol{w}$ is called the number of adaptive degrees of freedom (*dof*).

SINR Optimality

An optimum adaptive weight vector may be obtained with relaxed conditions on the statistical model of the radar measurements. Given the signal and additive noise vector

$$\boldsymbol{z} = a\boldsymbol{v} + \boldsymbol{n} \tag{95}$$

where $\boldsymbol{n}$ is a zero-mean random vector with covariance $\boldsymbol{R} = E[\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}]$, compute the linear filter

$$\text{linear filter output} = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z} \tag{96}$$

that maximizes the SINR. Note that only the first- and second-order statistics of the noise have been specified. The interference-plus-noise power is

$$E|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{n}|^2 = E[\boldsymbol{w}^{\mathrm{H}}\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\boldsymbol{w}] = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w} \tag{97}$$

The signal-plus-interference-plus-noise power is

$$E|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}|^2 = |a|^2|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}|^2 + \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w} \tag{98}$$

Therefore, the SINR is

$$\mathrm{SINR} = |a|^2 \frac{|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}|^2}{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w}} \tag{99}$$

We now utilize a generalization of the finite-dimensional Cauchy–Schwartz inequality.

For complex vectors $\boldsymbol{w}$ and $\boldsymbol{v}$ and nondegenerate Hermitian inner product $\langle , \rangle$,

$$|\langle \boldsymbol{w}, \boldsymbol{v}\rangle|^2 \le \langle \boldsymbol{w}, \boldsymbol{w}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle \tag{100}$$

and equality holds if and only if $\boldsymbol{w}$ is a scalar multiple of $\boldsymbol{v}$. For the problem of optimizing the SINR, take $\langle \boldsymbol{w}, \boldsymbol{v}\rangle$ def $\equiv \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{v}$; then $\mathrm{SINR} = |a|^2|\langle \boldsymbol{w}, \boldsymbol{R}^{-1}\boldsymbol{v}\rangle|^2/\langle \boldsymbol{w}, \boldsymbol{w}\rangle$. Thus the SINR is maximized if and only if

$$\boldsymbol{w} = \alpha\boldsymbol{R}^{-1}\boldsymbol{v} \tag{101}$$

for an arbitrary complex constant α. The maximum SINR is

$$\mathrm{SINR}_{\max} = |a|^2\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} \tag{102}$$

Like Eq. (95), Eq. (103) is called the adaptive filter equation, and the components of $\boldsymbol{w}$ are called adaptive weights.

Maximum Likelihood Estimate of Target Amplitude

The adaptive filter equation is also a consequence of the maximum likelihood estimate of the target amplitude. The log likelihood function (ignoring constants) from Eq. (82) is

$$l(\boldsymbol{z}|a) = -\boldsymbol{z}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z} + a^*\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z} + a\boldsymbol{z}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} - aa^*\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} \tag{103}$$

A necessary condition for optimality is that the first derivatives with respect to a^* and a vanish. Recall that if $a = x + jy$ is a complex variable, the derivatives with respect to a and a^* are

$$\frac{\partial}{\partial a} = \frac{1}{2}\left(\frac{\partial}{\partial x} - j\frac{\partial}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial a^*} = \frac{1}{2}\left(\frac{\partial}{\partial x} + j\frac{\partial}{\partial y}\right)$$

These derivatives yield the desired formulae

$$\frac{\partial a}{\partial a} = 1, \qquad \frac{\partial a^*}{\partial a} = 0, \qquad \frac{\partial a}{\partial a^*} = 0, \quad \text{and} \quad \frac{\partial a^*}{\partial a^*} = 1$$

Note that $f(a, a^*)$ is an analytic function if and only if $\partial f/\partial a^* = 0$, which simply expresses the Cauchy–Riemann equations. As we are considering the log likelihood function—a real-valued function everywhere—analyticity is irrelevant here.

Solving the equation

$$\frac{\partial l}{\partial a^*} = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z} - a\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} = 0 \tag{104}$$

for a provides its maximum likelihood (*ML*) estimate

$$\hat{a} = \frac{\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}}{\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}} \tag{105}$$

The Hessian matrix

$$\begin{pmatrix} (\partial^2/\partial^2 a)l & (\partial^2/\partial a\,\partial a^*)l \\ (\partial^2/\partial a^*\,\partial a)l & (\partial^2/\partial^2 a^*)l \end{pmatrix} = \begin{pmatrix} 0 & -\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} \\ -\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} & 0 \end{pmatrix} \tag{106}$$

is negative definite with respect to the Hermitian metric

$$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

therefore, this estimate provides a maximum. Inserting the ML estimate of Eq. (109) into the maximum SINR of Eq. (104) yields the Neyman–Pearson statistic of Eq. (91).

Adaptive Nulling and Interference Subspaces

The optimum properties of adaptive filters are theoretically and practically important; however, perhaps the most intuitive explanation of their function comes from considering the effect of adaptivity on the interference and noise subspaces. These are defined by the weight vectors $\boldsymbol{u}$ that are stationary values of the interference-plus-noise-to-noise ratio (*INNR*),

$$\mathrm{INNR} = \frac{\boldsymbol{u}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{u}}{\boldsymbol{u}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}\boldsymbol{u}} \tag{107}$$

By the Courant–Fisher minimax theorem, these vectors and the corresponding INNR values are, respectively, generalized eigenvectors and generalized eigenvalues of the generalized eigenvalue problem

$$\boldsymbol{R}\boldsymbol{u} = \lambda\boldsymbol{R}_{\mathrm{n}}\boldsymbol{u} \tag{108}$$

Its solution yields the decomposition of the interference-plus-noise covariance

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{R}_{\mathrm{n}}^{1/2}\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{1/2} \\ &= \boldsymbol{R}_{\mathrm{n}}^{1/2}(\boldsymbol{U}_{\mathrm{i}} \quad \boldsymbol{U}_{\mathrm{n}})\begin{pmatrix} \boldsymbol{\Lambda}_{\mathrm{i}} & \\ & \boldsymbol{\Lambda}_{\mathrm{n}} \end{pmatrix}\begin{pmatrix} \boldsymbol{U}_{\mathrm{i}}^{\mathrm{H}} \\ \boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}} \end{pmatrix}\boldsymbol{R}_{\mathrm{n}}^{1/2} \end{aligned} \tag{109}$$

where $\Lambda = \mathrm{diag}(\Lambda_{\mathrm{i}}, \Lambda_{\mathrm{n}})$ is a diagonal matrix of ordered eigenvalues and $\boldsymbol{U} = (\boldsymbol{U}_{\mathrm{i}}, \boldsymbol{U}_{\mathrm{n}})$ is a unitary matrix of eigenvectors. The generalized eigenvalues in the matrix Λ_{n} are called the noise eigenvalues. For ideal covariance matrices, $\Lambda_{\mathrm{n}} = \boldsymbol{I}$. The generalized eigenvalues in the matrix $\Lambda_{\mathrm{i}} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r)$ are called the interference eigenvalues; any eigenvalue that exceeds the greatest noise eigenvalue is an interference eigenvalue. For ideal covariance matrices, this distribution is straightforward; however, for sample covariance matrices, there is an arbitrary cutoff set between interference and noise eigenvalues, because it is difficult or impossible to measure the interference and noise separately. In practice, one uses a sample covariance matrix in place of $\boldsymbol{R}$ (an assumed ideal matrix $\boldsymbol{R}_{\mathrm{n}}$, which is typically $\sigma^2{}_{\mathrm{n}}\boldsymbol{I}$) and declares all generalized eigenvalues of Eq. (113) near or below the 0 dB level to be noise eigenvalues. Note well that the 0 dB level for *generalized* noise eigenvalues is implied by Eq. (113) and not the noise covariance $\boldsymbol{R}_{\mathrm{n}}$ itself. The column spans of the generalized eigenvector matrices $\boldsymbol{U}_{\mathrm{i}} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r)$ and $\boldsymbol{U}_{\mathrm{n}}$ are called the interference and noise subspaces, respectively. The dimension r of the interference subspace is called the *rank* of the interference. Each generalized eigenvector in $\boldsymbol{U}$ is called an interference or noise eigenvector. The response of each of these vectors to a bank of steering vectors is called an eigenbeam. For the case of pulse Doppler radar or a uniform linear array, an eigenbeam is simply the discrete Fourier transform of an eigenvector.

The adaptive weight equation (95) using the interference and noise subspaces from Eq. (114) becomes

$$\begin{aligned} \boldsymbol{w} &= \boldsymbol{R}^{-1}\boldsymbol{v} \\ &= \boldsymbol{R}_{\mathrm{n}}^{-1/2}\boldsymbol{U}_{\mathrm{i}}[\boldsymbol{\Lambda}_{\mathrm{i}}^{-1}(\boldsymbol{U}_{\mathrm{i}}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1/2}\boldsymbol{v})] + \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{U}_{\mathrm{n}}[\boldsymbol{\Lambda}_{\mathrm{n}}^{-1}(\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1/2}\boldsymbol{v})] \end{aligned} \tag{110}$$

Interpreting the adaptive weight equation in the form of Eq. (115), the optimum weights are obtained by first whitening the steering vector, second projecting the whitened steering vector $\boldsymbol{R}^{-1/2}{}_{\mathrm{n}}\boldsymbol{v}$ onto the interference and noise subspaces, third scaling the projected, whitened steering vectors $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{R}^{-1/2}{}_{\mathrm{n}}\boldsymbol{v}$ by the reciprocals of the interference and noise eigenvalues, and fourth lifting the resulting vectors back up into the original unwhitened space. The fourth step may also be thought of as premultiplying by the interference and noise subspaces so that when the product $\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}$ is formed, the steering vector is projected onto these subspaces.

The effect of nulling becomes apparent if one views Eq. (115) for the special white noise case $\boldsymbol{R}_{\mathrm{n}} = \boldsymbol{I}$ (noise at 0 dB). Each of the interferernce eigenvalues represents an INNR = INR + 1, where INR is the interference-to-noise ratio, and Eq. (115) may be written

$$\boldsymbol{w} = \sum_{\substack{\text{interference} \\ \text{eigenvectors}}} \frac{1}{1+\mathrm{INR}_i}\boldsymbol{u}_i(\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{v}) + \sum_{\substack{\text{noise} \\ \text{eigenvectors}}} \boldsymbol{u}_i(\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{v}) \tag{111}$$

Adaptive nulling is achieved because the interference eigenvalues INR_i are larger than the noise; therefore, dividing any component of the steering vector in the interference subspace by the relatively large INR reduces the interference allowed through the adaptive filter. Note that unless the INR is infinite, the adaptive weights are not simply an orthogonal projection of the steering vector onto the noise subspace. The adaptive weights balance the amounts of interference and noise power to optimize the SINR (or probability of detection) at the adaptive filter's output.

Adaptive Cancellation

Though adaptive cancellation is synonymous with adaptive filtering, it is best identified with the adaptive filter of Eq. (95) expressed in the following way. As discussed in the subsection "Noise, Jammer, and Clutter Covariance Matrices" above, the interference-plus-noise covariance matrix may be expressed in the form

$$\boldsymbol{R} = \boldsymbol{R}_{\mathrm{n}} + \boldsymbol{C}\boldsymbol{C}^{\mathrm{H}} \tag{112}$$

where $\boldsymbol{C}$ is the Cholesky decomposition of the interference covariance. By the Sherman–Morrison–Woodbury matrix inversion lemma (26),

$$\boldsymbol{R}^{-1} = \boldsymbol{R}_{\mathrm{n}}^{-1} - \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C}(\boldsymbol{I} + \boldsymbol{C}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C})^{-1}\boldsymbol{C}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1} \tag{113}$$

Applying Eq. (118) to the adaptive weight equation (95),

$$\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v} = \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{v} - \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C}((\boldsymbol{I} + \boldsymbol{C}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C})^{-1}(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{v})) \tag{114}$$

The two terms in this equation provide the interpretation of an adaptive canceler—a mainbeam $\boldsymbol{R}^{-1}{}_{\mathrm{n}}\boldsymbol{v}$ is formed on the target, and any interference that appears in the several sidelobe beams $\boldsymbol{R}^{-1}{}_{\mathrm{n}}\boldsymbol{C}$ is subtracted out.

Equation (119) implies that the optimum weight vector lies in the subspace spanned by the quiescent adaptive weight vector $\boldsymbol{R}^{-1}{}_{\mathrm{n}}\boldsymbol{v}$ and the subspace $\boldsymbol{R}^{-1}{}_{\mathrm{n}}\boldsymbol{C}$. Explaining the appearance of the term $\boldsymbol{R}^{-1}{}_{\mathrm{n}}$ in Eq. (119) is instructive. The adaptive filter output

$$\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v} = \tilde{\boldsymbol{w}}^{\mathrm{H}}\tilde{\boldsymbol{v}} \tag{115}$$

may be expressed using either the unwhitened steering vector $\boldsymbol{v}$ and adaptive weight vector $\boldsymbol{w}$, or the whitened steering vector $\tilde{v} = \boldsymbol{R}^{-1/2}{}_{\mathrm{n}}\boldsymbol{v}$ and whitened adaptive weight vector $\tilde{w} = \boldsymbol{R}^{1/2}{}_{\mathrm{n}}\boldsymbol{w}$. This duality encountered between steering vectors and weight vectors arises frequently in mathematics, physics, and engineering, where the opposite conventions for transforming coordinates (whitening) are called covariant and contravariant. The appearance of $\boldsymbol{R}^{-1}{}_{\mathrm{n}}$ is now explained by the sequence of whitening the steering vector $\boldsymbol{v}$, then unwhitening the adaptive weight vector $\tilde{w} = \tilde{v}$:

$$\boldsymbol{w} = \boldsymbol{R}_{\mathrm{n}}^{1/2}\tilde{\boldsymbol{w}} = \boldsymbol{R}_{\mathrm{n}}^{1/2}\tilde{\boldsymbol{v}} = \boldsymbol{R}_{\mathrm{n}}^{1/2}(\boldsymbol{R}_{\mathrm{n}}^{1/2}\boldsymbol{v}) = \boldsymbol{R}_{\mathrm{n}}^{1/2}\boldsymbol{v}$$

Preprocessing for Adaptive Filtering

The adaptive weight vector $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v}$ is derived above using direct measurements from radar elements and/or radar pulses according to the model Eq. (81). However, it is oftentimes desirable to perform adaptivity on a subset of the elements or pulses, or the outputs of radar's beams and/or Doppler filters. There are several important reasons for preprocessing before adaptive filtering (26a–c) [parker (66)]:

- It is desirable in general to eliminate as much interference as possible nonadaptively because of the extra system complexity required for adaptive processing. Beamformers and Doppler filters are designed to mitigate sidelobe interference with low sidelobes; therefore, it is reasonable to consider using them before adaptive nulling. The engineering wisdom here is "do not eliminate adaptively what you can eliminate nonadaptively."
- There is a loss associated with estimating the unknown interference-plus-noise covariance matrix. This loss, quantified explicitly in the subsection "SINR Loss Factors" below, is essentially a ratio between the number of adaptive degrees of freedom and the number of samples used to estimate the covariance—the more samples, the smaller the loss. For many radar environments, the sample support for covariance matrix estimation is limited; therefore, the only way to decrease estimation loss is to reduce the dimension of the adaptive weight vector. This reduction can be accomplished by adapting on a limited number of antenna elements, pulses, beams, or Doppler bins.
- Radars must provide target information in real time for all rangegates, beams, and Doppler bins. Because the steering vector $\boldsymbol{v}$ changes for each beam and Doppler filter (and possibly each rangegate), adaptive radars must solve many, many matrix inverse problems in real time to computer all the different adaptive weight vectors. To avoid a potentially large and infeasible computational expense, quantified in the section "Adaptive Weight Computation" below, the dimensionality of the adaptive weight vector may need to be reduced.
- Because the number of adaptive degrees of freedom for STAP algorithms is the number of antenna elements multiplied by the number of radar pulses (potentially a very large number), reduced-dimension algorithms are a necessity for STAP. Fortunately, as discussed in the next section, such suboptimum STAP algorithms can achieve near-optimum performance due to the special structure of jamming and clutter interference.

Given an N-dimensional vector of radar data $\boldsymbol{z}$, a preprocessor for adaptive nulling is defined by an N-by-d matrix $\boldsymbol{T}$, and the d-dimensional vector $\tilde{z}$ of preprocessor outputs is determined by the linear transformation

$$\tilde{\boldsymbol{z}} = \boldsymbol{T}^{\mathrm{H}}\boldsymbol{z} \tag{116}$$

The columns of $\boldsymbol{T}$ may represent beamformers, Doppler filters, unit vectors of the form $(0, \ldots, 1, 0, \ldots, 0)^{\mathrm{T}}$ that select certain elements or pulses, or a space–time combination of two of these. Given that a preprocessor $\boldsymbol{T}$ is to be used, the adaptive weight vector $\boldsymbol{w}$ must be of the form

$$\boldsymbol{w} = \boldsymbol{T}\tilde{\boldsymbol{w}} \tag{117}$$

for some vector $\tilde{w}$ i.e., the weight vector must be constrained to lie in the subspace spanned by the columns of $\boldsymbol{T}$. From Eq. (101), the constrained SINR is

$$\mathrm{SINR}(\mathrm{constrained}) = |\alpha|^2 \frac{|\tilde{\boldsymbol{w}}^{\mathrm{H}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{v}|^2}{\tilde{\boldsymbol{w}}^{\mathrm{H}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T}\tilde{\boldsymbol{w}}} \tag{118}$$

As in the unconstrained case, the SINR is optimized by the vector

$$\tilde{\boldsymbol{w}} = (\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T})^{-1}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{v} \tag{119}$$

Equations (123) and (125) imply that the optimum constrained adaptive weight vector is

$$\boldsymbol{w} = \boldsymbol{T}(\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T})^{-1}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{v} \tag{120}$$

The computation of this weight vector is interpreted as consisting of the following three steps:

1. $\tilde{\boldsymbol{v}} = \boldsymbol{T}^{\mathrm{H}}\boldsymbol{v}$ (121)

2. $\tilde{\boldsymbol{w}} = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{v}} \qquad (\tilde{\boldsymbol{R}} = E[\tilde{\boldsymbol{n}}\tilde{\boldsymbol{n}}^{\mathrm{H}}] = \boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T}, \quad \tilde{\boldsymbol{n}} = \boldsymbol{T}^{\mathrm{H}}\boldsymbol{n})$ (122)

3. $\boldsymbol{w} = \boldsymbol{T}\tilde{\boldsymbol{w}}$ (123)

whereby a preprocessing steering vector is formed, an adaptive weight vector is formed in the preprocessor space, and a constrained adaptive weight vector is formed in the full space.

Note that the constrained weight vector is in general suboptimum compared with the fully adaptive weight vector $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v}$. Nevertheless, one important consequence of Eq. (126) is that any preprocessor $\boldsymbol{T}$ satisfying the property

$$\boldsymbol{R}^{-1}\boldsymbol{v} \in \operatorname{span}\boldsymbol{T} \tag{124}$$

yields a reduced dimension adaptive weight vector achieving the same performance as the fully optimum adaptive weight vector $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v}$ (e.g., take $\boldsymbol{T} = \boldsymbol{R}^{-1}\boldsymbol{v}$). In particular, Eq. (119) imples that

$$\boldsymbol{T}^{-1}\boldsymbol{v} \in \operatorname{span}(\boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{v}, \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C}) \tag{125}$$

yielding the important fact that fully optimum performance is achieved using the reduced-dimension preprocessor

$$\boldsymbol{T} = (\boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{v}, \boldsymbol{R}_{\mathrm{n}}^{-1}\boldsymbol{C}) \tag{126}$$

Recall that $\boldsymbol{C}$ is any matrix whose columns span the interference subspace $\boldsymbol{U}_{\mathrm{i}}$. The engineering interpretation of these facts is that one should design a preprocessor whose columns point to as much of the steering-vector direction and interference directions as possible.

Adaptive Degrees of Freedom and Nulling Performance

The number of adaptive degrees of freedom, that is, the number N of elements in the adaptive weight vector $\boldsymbol{w} = (w_1, w_2, \ldots, w_N)^{\mathrm{T}}$, sets a fundamental limit on the performance of adaptive processing. Because the dimension of the space of vectors perpendicular to $\boldsymbol{w}$,

$$(\boldsymbol{w})^{\perp} = \{\boldsymbol{c} \in \mathbb{C}^N : \boldsymbol{w}^{\mathrm{H}}\boldsymbol{c} = 0\} \tag{127}$$

has dimension $N - 1$, a weight vector with N adaptive degrees of freedom may null completely at most $N - 1$ independent interference sources. Equivalently, if the interference rank is equal to N, there will be unnulled interference for any adaptive weight vector of dimension N.

Constrained Adaptive Weights

In addition to the adaptive weights constrained to lie in a given subspace, considered in the subsection "Preprocessing for Adaptive Filtering" above, it is also often desirable to either constrain the gain of $\boldsymbol{w}$ on a particular steering vector $\boldsymbol{t}$ (52), i.e.,

$$\boldsymbol{w}^{\mathrm{H}}\boldsymbol{t} = 1 \tag{128}$$

or to constrain the output interference-plus-noise power, i.e.,

$$\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w} = \text{constant} \tag{129}$$

or to constrain the output noise power of the weight vector, i.e.,

$$\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}\boldsymbol{w} = \text{constant} \tag{130}$$

Satisfying the gain constraint of Eq. (134) is trivial: for any weight vector $\boldsymbol{w}$, use instead the weight vector

$$\boldsymbol{w} \mapsto \frac{1}{\boldsymbol{t}^{\mathrm{H}}\boldsymbol{w}}\boldsymbol{w} \tag{131}$$

The weight vector of Eq. (137) with $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{t}$ is often called the minimum-variance distortionless response (MVDR) beamformer because it provides unit ("distortionless") gain on the target steering vector and minimizes the variance of its output.

Constraining the output interference-plus-noise power to unity is similarly trivial; for any weight vector $\boldsymbol{w}$, use instead the weight vector

$$\boldsymbol{w} \mapsto \frac{1}{(\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w})^{1/2}}\boldsymbol{w} \tag{132}$$

A weight vector optimizing the SINR and satisfying the output noise power constraint of Eq. (136) may be obtained by applying the method of Lagrange multipliers to the problem

$$\max_{\boldsymbol{w}} |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}|^2 \qquad \text{such that} \quad \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w} = c_1 \text{ and } \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}\boldsymbol{w} = c_2 \tag{133}$$

The augmented function to be optimized may be expressed as

$$F(\boldsymbol{w}, \lambda, \delta) = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}\boldsymbol{v}^{\mathrm{H}}\boldsymbol{w} - \lambda[(\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{w} - c_1) - \delta(\boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}\boldsymbol{w} - c_2)] \tag{134}$$

where λ and $\lambda\delta$ are Lagrange multipliers and an equivalent expression for $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{v}|^2$ has been used. Differentiating F with respect to $\boldsymbol{w}^{\mathrm{H}}$ yields the generalized eigenvalue problem

$$\boldsymbol{v}\boldsymbol{v}^{\mathrm{H}}\boldsymbol{w} = \lambda(\boldsymbol{R} + \delta\boldsymbol{R}_{\mathrm{n}})\boldsymbol{w} \tag{135}$$

The only nontrivial solution to Eq. (141) is

$$\boldsymbol{w} = (\boldsymbol{R} + \delta\boldsymbol{R}_{\mathrm{n}})^{-1}\boldsymbol{v} \tag{136}$$

To satisfy both the interference-plus-noise and the noise output power constraint, this weight vector must be scaled similarly to Eq. (138) and an appropriate value for δ must be found.

Diagonal Loading

In practice the sample covariance matrix $\hat{\boldsymbol{R}}$ is used to form the adaptive weight vector as in Eq. (74). The eignevalue

decomposition of $\hat{\boldsymbol{R}}$ is

$$\hat{\boldsymbol{R}} = \boldsymbol{R}_{\mathrm{n}}^{1/2}\hat{\boldsymbol{U}}\,\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{1/2} \tag{137}$$

[cf. Eq. (114)], where it assumed that the noise covariance $\boldsymbol{R}_{\mathrm{n}}$ is known and that the eigenvalues are sorted so that $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_N$. In terms of these estimated eigenvalues $\hat{\lambda}_i$ and estimated eigenvectors $\hat{\boldsymbol{u}}_i$, the adaptive weight vector $\hat{\boldsymbol{w}} = \hat{\boldsymbol{R}}^{-1}\boldsymbol{v}$ is

$$\hat{\boldsymbol{w}} = \sum_{i=1}^{N} \frac{1}{\hat{\lambda}_i}\boldsymbol{R}_{\mathrm{n}}^{-1}\hat{\boldsymbol{u}}_i(\hat{\boldsymbol{u}}_i^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1/2}\boldsymbol{v}) \tag{138}$$

It is a fact that for sample covariance matrices formed with K Gaussian random vectors [such matrices are said to be Wishart-distributed (27)], the smallest noise eigenvalue is approximated by the value (28)

$$\hat{\lambda}_N \approx [1 - (N/K)^{1/2}]^2 \tag{139}$$

for large N and $K \geq N$. Equation (145) indicates how a small sample support may yield poor adaptive performance. Ideally, the noise eigenvalues are all unity as in Eq. (113); however, for $K = 2N$, say, the smallest noise eigenvalue will be approximately $\frac{3}{2} - \sqrt{2} \approx -11$ dB. When this noise eigenvalue is inverted in Eq. (195), the component of the steering vector in the direction of the corresponding noise eigenvector is erroneously scaled by 11 dB, instead of the correct scaling of 0 dB. Similarly, the component of the steering vector in the entire noise subspace is scaled incorrectly. The result is an adaptive weight vector possessing very high sidelobes.

To correct this problem and the high sidelobes it causes, an adaptive weight vector with *diagonal loading*

$$\hat{\boldsymbol{w}} = (\hat{\boldsymbol{R}} + \delta\boldsymbol{R}_{\mathrm{n}})^{-1}\boldsymbol{v} \tag{140}$$

is formed. This procedure is so called because the diagonal of the whitened sample covariance matrix $\boldsymbol{R}^{-1/2}{}_{\mathrm{n}}\hat{\boldsymbol{R}}\mathrm{R}^{-1/2}{}_{\mathrm{n}}$ is "loaded" by the term $\delta\boldsymbol{I}$. Diagonal loading adds the constant δ to all the eigenvalues, so that the diagonally loaded adaptive weight vector becomes

$$\hat{\boldsymbol{w}} = \sum_{i=1}^{N} \frac{1}{\delta + \hat{\lambda}_i}\boldsymbol{R}_{\mathrm{n}}^{-1/2}\hat{\boldsymbol{u}}_i(\hat{\boldsymbol{u}}_i^{\mathrm{H}}\boldsymbol{R}_{\mathrm{n}}^{-1/2}\boldsymbol{v}) \tag{141}$$

If the typical δ = 5 dB to 10 dB of diagonal loading is used (29), the small (less than 0 dB) noise eigenvalues become unimportant, and the component of the steering vector in the noise subspace is scaled almost uniformly by $1/\delta$. In effect, diagonal loading raises the noise floor by the factor $1 + \delta$. While an adaptive weight vector formed using diagonal loading will yield good performance against strong interference (large interference eigenvalues), its performance against weaker interference (small interference eigenvalues) will be degraded (30) [Li 2003].

Diagonal loading (*DL*) may also be viewed in the following ways:

- DL incorporates prior knowledge about the noise covariance to improve the quiescent (interference-free) adaptive weight response.
- DL provides a tradeoff between nonadaptive steering vectors (δ large) and fully adaptive weight vectors (δ small).
- DL constrains the length of the adaptive weight vector as in Eq. (142) to counteract the effects of poorly estimated noise eigenvalues.

Figure 7 illustrates the clutter eigenvalues found in the PRI-staggered post-Doppler STAP algorithm described in the sub-subsection "Element-Space Post-Doppler STAP" below for the Mountaintop (31) scenario. The clutter subspace dimension is seen to be about 20, which is above the rank of about 16 predicted by Brennan's rule (Eqs. (56) and (151) for $N = 14, K_{\mathrm{t}} = 16$, and $\beta = 0.93$) because the array axis is not aligned with the velocity vector, violating a key assumption. Sample support of twice the number of degrees of freedom was used for the sample covariance matrix; the smallest eigenvalue is above the -11 dB predicted by Eq. (145) because the ideal covariance matrix is not white.

SPACE–TIME ADAPTIVE PROCESSING

Introduction

Space–time adaptive processing (STAP) (3,4,11–13), [Guerci 2003], [Steinhardt 2000] is used to mitigate clutter and jammer interference in airborne radar. Because ground clutter has a spatiotemporal dependence—clutter to the front of the aircraft has a positive Doppler shift and clutter to the rear has a negative shift—a two-dimensional filter utilizing both these dimensions is required. Such a filter is obtained from the adaptive weight vector $\boldsymbol{w} = \boldsymbol{R}{-}\boldsymbol{1}\boldsymbol{v}$ of Eq. (72) if $\boldsymbol{v}$ is a space–time steering vector of the form described above in the sub-subsection on that subject and $\boldsymbol{R}$ is a space–time covariance matrix of the form described in the subsection "Noise, Jammer, and Clutter Covariance Matrices" above. However, this full space–time dimension approach is impractical for several reasons:

- Inverting the (estimated) covariance matrix requires on the order of $(MN)^3$ real-time floating-point operations, where M is the number of radar pulses and N is the number of independent antenna elements. This is prohibitively expensive for real-time operations even for modestly sized radars that use tens of pulses and tens of elements.
- Estimating the covariance matrix introduces an SINR loss proportional to the ratio of the number of adaptive degrees of freedom to the sample support (see the next section). The large number of degrees of freedom found in full-dimension STAP requires a correspondingly large number of samples to achieve low estimation losses, and the sample support is typically limited.
- As will be seen in this section, the special structure of clutter interference allows for a family of reduced-dimension STAP algorithms that achieve near-optimum performance using a small fraction of the full-dimension degrees of freedom. Reducing the

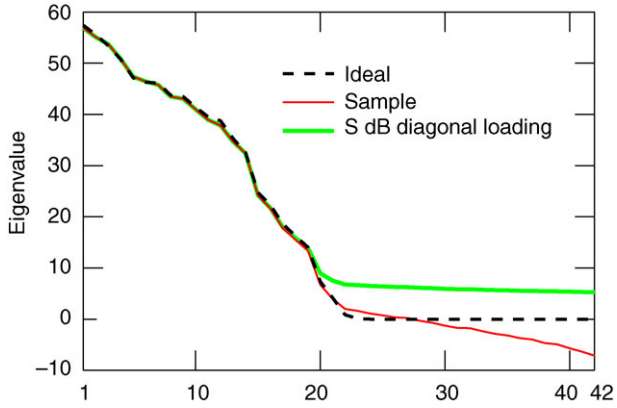

Figure 7. Clutter eigenvalues of the ideal, sample, and diagonally loaded covariance matrices from the PRI-staggered post-Doppler STAP algorithm described in the sub-subsection "Element-Space Post-Doppler STAP" for the Mountaintop (31) scenario.

dimension of adaptive algorithms reduces the cost and complexity of the processor used to implement them.

- Preprocessing with Doppler filters or beamformers eliminates some interference nonadaptively, which may simplify the system complexity required for the adaptive processor.

Practical STAP algorithms are based upon the preprocessor approach described in the subsection "Preprocessing for Adaptive Filtering" above for a specific choice of a space–time preprocessor matrix $\boldsymbol{T}$. In all cases, the choice of this matrix is based upon a generalization of the straightforward physical principle that the angle- and Doppler-dependent clutter may be canceled by taking the difference between clutter measurements at two separate points in space and time. The block diagram for the STAP radar signal-processing chain is illustrated in Fig. 8. Real-time implementation of STAP methods demands high performance front-end signal processing and computing architectures [Martinez (64)], [Martinez (63)], [Martinez (62)].

STAP algorithms are compared by examining the SINR loss they achieve (see the sub-subsection "SINR and SINR Loss" above). Because pulse–Doppler radars operate by transmitting a beam of energy in a given direction, then examining the radar echoes for the presence of targets at any speed, the SINR loss is plotted as a function of target Doppler frequency at a fixed direction of interest. The minimum detectable velocity is then derived from this plot.

Sample support for covariance matrix estimation is obtained from the radar's rangegates. The operating assumption is that the interference-plus-noise statistics are independent and identically distributed over these ranges, so that the corresponding sample covariance matrix accurately represents the ideal covariance. Of course, ground clutter is not so ideally homogeneous, and modified training strategies are necessary to solve the problems introduced by this fact (32, 33). The data available to the STAP processor are conveniently represented as a cube (Fig. 9) whose dimensions represent the independent antenna element outputs, the radar pulses, and the radar range gates to be used for covariance matrix estimation.

STAP Algorithm Taxonomy

STAP algorithms may be classified using Ward's taxonomy (12) by the dimension in which adaptivity occurs. The natural choices for adaptive degrees of freedom in a pulse Doppler radar are pre-Doppler versus post-Doppler (adapting on pulses versus adapting on Doppler bin outputs) and element space versus beam space (adapting on elements versus adapting on beamformer outputs). All four combinations of these two choices are possible (Fig. 10).

Each STAP algorithm is defined by choosing one of these quadrants and the subset of elements or beams and pulses or Doppler filter banks used for adaptivity (i.e., the choice of a space–time preprocessor matrix $\boldsymbol{T}$), as well as choice of beamformers and Doppler filters. This design flexibility allows for many different STAP algorithms, each of whose performance is dependent upon specific radar system parameters. Oftentimes, the radar's entire spatial aperture is used for adaptive degrees of freedom, and temporal adaptivity is introduced using time-varying sub-CPIs. We shall describe one representative algorithm per quadrant; results are shown in Figs. 11 and 12. Ward (12) provides a more complete compilation.

Element-Space Pre-Doppler STAP. Element-space pre-Doppler STAP adapts over a sequence of time-dependent sub-CPIs. Given an M-pulse CPI, space–time adaptive filter outputs are computed for all elements over a sliding window of sub-CPIs with length-K_t pulses, resulting in $M' = M - K_t + 1$ adaptive filters. These M' adaptive filter outputs are then coherently combined using an M'-pulse Doppler filter. Adaptation over sub-CPIs can mitigate interference that varies faster than the CPI length.

Expressed using matrices, the adaptive weight vector for this STAP algorithm is

$$\boldsymbol{w} = \boldsymbol{W}\boldsymbol{v}_{t'}(f_d) \tag{142}$$

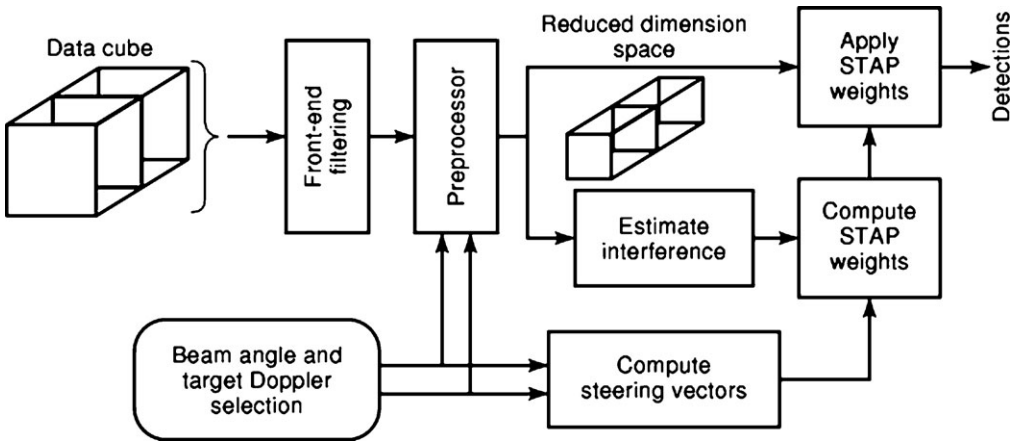

Figure 8. STAP radar signal-processing chain.

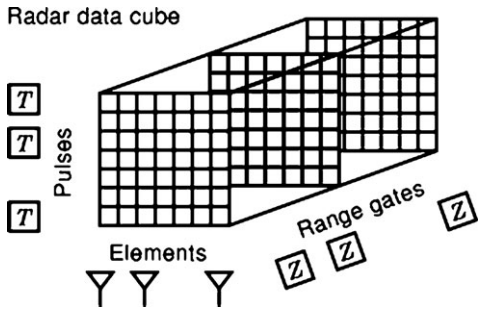

Figure 9. STAP radar data cube.

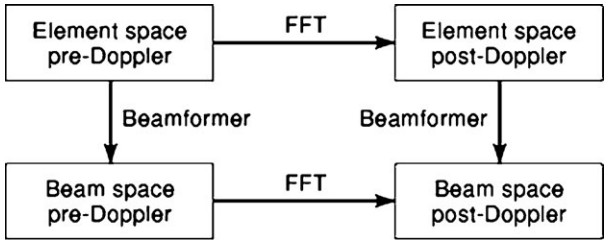

Figure 10. STAP algorithm taxonomy (12).

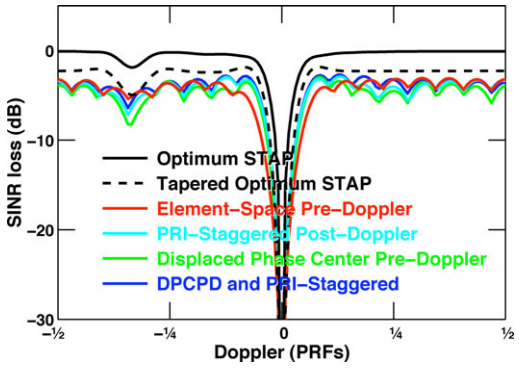

Figure 11. SINR loss for the STAP algorithms considered.

where

$$\boldsymbol{W} = \begin{pmatrix} \tilde{\boldsymbol{R}}_1^{-1}\tilde{\boldsymbol{v}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{R}}_2^{-1}\tilde{\boldsymbol{v}} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & & \vdots \\ \vdots & \vdots & & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \tilde{\boldsymbol{R}}_{M'}^{-1}\tilde{\boldsymbol{v}} \end{pmatrix} \tag{143}$$

$$\tilde{\boldsymbol{R}}_k = \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{R} \boldsymbol{T}_k, \qquad k = 1, \ldots, M' \tag{144}$$

$$\boldsymbol{T}_k = (\boldsymbol{e}_k \quad \boldsymbol{e}_{k+1} \quad \cdots \quad \boldsymbol{e}_{k+K_t-1}) \otimes \boldsymbol{I}_N, \qquad k = 1, \ldots, M' \tag{145}$$

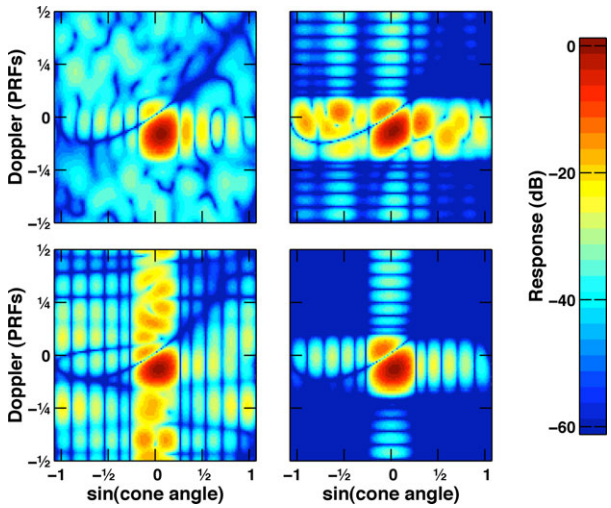

Figure 12. Adaptive weight response for the STAP algorithms considered. Compare these response patterns with those shown in Figs. 1 and 5.

$$\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}_{\mathrm{t}} \otimes \boldsymbol{v}_{\mathrm{s}}(\phi, \theta) \tag{146}$$

$\boldsymbol{v}_{t'}(f_{\mathrm{d}})$ is a tapered Doppler filter of the form of Eq. (26), $\boldsymbol{v}_{\mathrm{s}}(\phi, \theta)$ is a beamformer of the form of Eq. (21) (both may be tapered for sidelobe control as described in the subsection "Tapered Steering Vectors" above), and $\boldsymbol{e}_k$ represents the kth canonical unit vector $(0, \ldots, 0, 1, 0, \ldots, 0)^{\mathrm{T}}$ (unity in the kth element, zeros elsewhere), $\boldsymbol{I}_N$ represents an N-by-N identify matrix, and each boldface zero represents an N-by-1 block of zeros. The sub-CPI steering vector $\tilde{\boldsymbol{v}}_{\mathrm{t}}$ is given by the equation

$$\tilde{\boldsymbol{v}}_{\mathrm{t}} = [\boldsymbol{I}_{K_{\mathrm{t}}} \otimes \boldsymbol{v}_{\mathrm{s}}(\phi, \theta)]^{\mathrm{H}} \boldsymbol{V} \boldsymbol{v}^{*}_{\mathrm{t}'}(f_{\mathrm{d}}), \qquad \boldsymbol{V} = (\boldsymbol{T}_1^{\mathrm{H}}\boldsymbol{t} \quad \boldsymbol{T}_2^{\mathrm{H}}\boldsymbol{t} \quad \cdots \quad \boldsymbol{T}_{M'}^{\mathrm{H}}\boldsymbol{t}) \tag{147}$$

where $\boldsymbol{t} = \boldsymbol{t}_{\mathrm{t}}(f_{\mathrm{d}}) \otimes \boldsymbol{t}_{\mathrm{s}}(\phi, \theta)$ is the true target steering vector, which assures that in the absence of interference, the signal power $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{t}|^2$ is maximized over all possible sub-CPI steering vectors. There are several other choices for $\tilde{\mathrm{v}}_{\mathrm{t}}$: a binomial taper (12) [e.g., $\boldsymbol{t} = (1\ -2\ 1)^{\mathrm{T}}$ for $K_{\mathrm{t}} = 3$] that is independent of Doppler frequency, or a Doppler-dependent optimized adaptive taper (34). Experience indicates that diagonal loading must be used with this algorithm when the sample covariance matrix $\hat{\boldsymbol{R}}$ is used, to avoid a problem associated with coherently combining the M' adaptive filters whose null positions are not colocated.

Note that this pre-Doppler STAP algorithm does not take the form of the optimum constrained adaptive weight vector in Eq. (126). Its adaptive weight vector is a coherent combination of the outputs of many suboptimum adaptive weight vectors. These outputs are not statistically independent, because they are formed using sub-CPI data with overlapping pulses; therefore, this pre-Doppler STAP algorithm does not satisfy the assumptions of the subsection "Adaptive Radar Hypothesis (Unknown Covariance)" above and the next section on adaptive detection.

If $\boldsymbol{R}$ is a Toeplitz–block-Toeplitz clutter covariance matrix with rank $N + \beta(M - 1)$, then the reduced dimension covariance $\boldsymbol{T}^{\mathrm{H}}{}_k \boldsymbol{R} \boldsymbol{T}_k$ is also Toeplitz–block-Toeplitz and has rank

$$\operatorname{rank} \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{R} \boldsymbol{T}_k \approx N + \beta(K_t - 1) \tag{148}$$

Element-Space Post-Doppler STAP. Element-space post-Doppler STAP adapts over all elements and a subset of Doppler filter outputs. In this class of STAP algorithms, the order of "adapt, then filter" used in pre-Doppler algorithms is reversed. In the case of the PRI-staggered post-Doppler (or multiwindow post-Doppler) STAP algorithm, a time-dependent series of Doppler filter outputs are formed from a sequence of sub-CPIs, then space–time adaptivity is performed using these Doppler outputs at every antenna element. Because pulse Doppler radars generally have very good pulse-to-pulse stability, Doppler filter banks with very low sidelobes (typically 60 dB or better) may be constructed, allowing post-Doppler STAP approaches to nonadaptively eliminate clutter interference away from the Doppler cell of interest.

This "filter, then adapt" procedure may be formulated using the preprocessing approach described in the subsection "Preprocessing for Adaptive Filtering" above. The adaptive weight vector

$$\boldsymbol{w} = \boldsymbol{T}(\boldsymbol{T}^{\mathrm{H}} \boldsymbol{R} \boldsymbol{T})^{-1} \boldsymbol{T}^{\mathrm{H}} \boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta) \tag{149}$$

is determined by the space–time preprocessing matrix

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{v}_{\mathrm{t}'}(f_{\mathrm{d}}) & 0 & \cdots & 0 \\ 0 & \boldsymbol{v}_{\mathrm{t}'}(f_{\mathrm{d}}) & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & \vdots & & 0 \\ 0 & 0 & \cdots & \boldsymbol{v}_{\mathrm{t}'}(f_{\mathrm{d}}) \end{pmatrix} \otimes \boldsymbol{I}_N \tag{150}$$

The Doppler filter $\boldsymbol{v}_{\mathrm{t}'}$ used to construct the preprocessor $\boldsymbol{T}$ uses $M' = M - K_{\mathrm{t}} + 1$ PRI taps, where K_{t} is the number of PRI staggers; $\boldsymbol{T}$ is an M N-by-$K_{\mathrm{t}}N$ matrix. The vector $\boldsymbol{v}_{\mathrm{t}'}$ is a tapered temporal steering vector of the form of Eqs. (26) and (34). The space–time steering vector $\boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta)$ is

given by Eq. (31) and may also have a space–time taper for sidelobe control.

If $\boldsymbol{R}$ is a Toeplitz–block-Toeplitz clutter covariance matrix with rank $N + \beta(M-1)$, then the reduced dimension covariance $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T}$ is also Toeplitz–block-Toeplitz and has rank

$$\operatorname{rank} \boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T} \approx N + \beta(K_{\mathrm{t}} - 1) \tag{151}$$

Beam-Space Pre-Doppler STAP. Beam-space pre-Doppler STAP algorithms are a generalization of the earliest nonadaptive space–time technique used to mitigate clutter interference—an airborne MTI (*AMTI*) technique called displaced phase-center antenna, or *DPCA* (3). DPCA works by forming a *right* beam along the platform's velocity vector, then forming a *left* beam one PRI later, and subtracting the left beam from the right beam. If the PRF is chosen correctly, the phase centers of the two beams will coincide because the platform is moving from left to right, and the resulting filter places a space–time null along the clutter interference.

In the general adaptive case, a time sequence of multiple independent beams are focused on the clutter. For example, the displaced phase center pre-Doppler STAP algorithm uses a sliding subaperture of beamformers. As with the PRI-staggered post-Doppler algorithm, it may be formulated using the preprocessing approach of the subsection "Preprocessing for Adaptive Filtering" above. Its adaptive weight vector is given by Eq. (155) using the preprocessing matrix

$$\boldsymbol{T} = \boldsymbol{I}_M \otimes \begin{pmatrix} \boldsymbol{v}_{s'}(\phi,\theta) & 0 & \cdots & 0 \\ 0 & \boldsymbol{v}_{s'}(\phi,\theta) & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & \vdots & & 0 \\ 0 & 0 & \cdots & \boldsymbol{v}_{s'}(\phi,\theta) \end{pmatrix} \tag{152}$$

The beamformer $\boldsymbol{v}_{s'}$ used to construct the preprocessor $\boldsymbol{T}$ uses a subaperture of length $N' = N - K_{\mathrm{s}} + 1$, where K_{s} is the number of beams; $\boldsymbol{T}$ is an $M\,N$-by-$M\,K_{\mathrm{s}}$ matrix. The vector $\boldsymbol{v}_{s'}$ is in general a tapered spatial steering vector of the form of Eqs. (21) and (34). The space–time steering vector $\boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta)$ is given by Eq. (31) and may also be tapered for sidelobe control.

Note that the displaced phase-center pre-Doppler STAP algorithm is the spatial dual of the PRI staggered post-Doppler algorithm of the preceding sub-subsection. Unlike post-Doppler algorithms, which can exploit low Doppler sidelobes before adaptation, this pre-Doppler algorithm may have to contend with higher clutter levels because it is difficult to construct antennas whose beamformers' sidelobes are comparable to those of Doppler filter banks.

Finally, if $\boldsymbol{R}$ is a Toeplitz–block-Toeplitz clutter covariance matrix with rank $N + \beta(M-1)$, then the reduced-dimension covariance $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T}$ is also Toeplitz–block-Toeplitz and has rank

$$\operatorname{rank} \boldsymbol{T}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{T} \approx K_{\mathrm{s}} + \beta(M - 1) \tag{153}$$

Beam-Space Post-Doppler STAP. Beam-space post-Doppler STAP algorithms combine beamformer and Doppler filter outputs adaptively. Because adaptation is typically accomplished with just a few beams and Doppler filter banks, these algorithms have the potential to greatly reduce the adaptive degrees of freedom. One straightforward method for constructing a beam-space post-Doppler algorithm is to combine a beamspace pre-Doppler algorithm with a element-space post-Doppler algorithm. For example, if the displaced phase center algorithm of the preceding sub-subsection and the PRI-staggered post-Doppler algorithm of the sub-subsection before that are combined, the resulting preprocessor has the form

$$\boldsymbol{T} = \boldsymbol{T}_{\text{PRI-staggered post-Doppler}} \otimes \boldsymbol{T}_{\text{Displaced phase center pre-Doppler}} \tag{154}$$

$$= \begin{pmatrix} \boldsymbol{v}_{t'}(f_{\mathrm{d}}) & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & & \vdots \\ \vdots & & 0 \\ 0 & \cdots & \boldsymbol{v}_{t'}(f_{\mathrm{d}}) \end{pmatrix} \otimes \begin{pmatrix} \boldsymbol{v}_{s'}(\phi,\theta) & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & & \vdots \\ \vdots & & 0 \\ 0 & \cdots & \boldsymbol{v}_{s'}(\phi,\theta) \end{pmatrix} \tag{155}$$

[cf. Eqs. 156 and 158]. The adaptive weight vector is then given by Eq. (155).

If $\boldsymbol{R}$ is a Toeplitz–block-Toeplitz clutter covariance matrix with rank $N + \beta(M-1)$, then the reduced dimension covariance $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T}$ is also Toeplitz–block-Toeplitz and has rank

$$\operatorname{rank} \boldsymbol{T}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{T} \approx K_{\mathrm{s}} + \beta(K_{\mathrm{t}} - 1) \tag{156}$$

ADAPTIVE DETECTION

Radar detection performance is measured by the probability of detecting a target given a fixed probability of false alarm. In the subsection "Detection Optimality" above it is seen for the case of known covariance matrix $\boldsymbol{R}$ that maximizing the PD yields the likelihood ratio detection test of Eq. (91) involving the adaptive weight vector $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v}$. In practice, the covariance matrix is unknown and must be estimated. In Neyman–Pearson theory, the PD is maximized if the likelihood ratio test is used; however, as this test involves the unknown covariance matrix, we do not have an optimum detection criterion for practical scenarios. The difference between detection tests involving estimates of the unknown covariance matrix and the optimum test of Eq. (91) is viewed as an estimation loss due to the unknown covariance. This estimation loss is comparable to the CFAR loss encountered in nonadaptive radars (14–53). Because we do not have an optimality criterion to drive our choice of adaptive detector, the selection is done with the following goals in mind:

- Low Estimation Loss The estimation loss, compared to the optimum known covariance detection performance, should be small, say a few decibels or less.
- CFAR Property The detection performance should be independent of the interference and noise scenario, that is, it should be independent of the underlying

covariance matrix $\boldsymbol{R}$.

- Robustness to Model Mismatch The detection performance should be robust to modeling errors such as steering vector mismatches to the true target, inclusion of targets in the covariance matrix estimate, and non-Gaussian interference.

The estimation loss may be approximated independently of a specific adaptive detector by computing the average SINR obtained from using the sample covariance matrix; this approach is presented in the following subsection. The detection performance of two important adaptive detectors—the generalized likelihood ratio test (*GLRT*) and the adaptive matched filter (*AMF*)—is presented in the subsection after.

SINR Loss Factors

In their influential paper on adaptive filtering, Reed et al. (36) derived the distribution of the amount of SNR lost when an adaptive filter is constructed using a sample covariance matrix. Boroson (37) generalized this loss factor to include the effects of steering vector mismatch and targets in the training data. A related loss factor will also be seen to play an essential role in the GLRT and AMF adaptive detectors discussed below.

Mismatched Steering Vectors. Let $\boldsymbol{t}$ be the true target steering vector and $\boldsymbol{v}$ be the assumed steering vector. A mismatch can arise from modeling errors, insufficient information about the target's angle or velocity, or the use of tapered steering vectors for sidelobe control in the adaptive filter design. If a matched filter $\boldsymbol{w} = \boldsymbol{R}^{-1}\boldsymbol{v}$ is formed with the assumed steering vector $\boldsymbol{v}$ and the (given) covariance matrix $\boldsymbol{R}$, then SINR of the matched filter as a function of $\boldsymbol{v}$ and $\boldsymbol{t}$ is

$$\mathrm{SINR}(\boldsymbol{v}, \boldsymbol{t}) = |\alpha|^2 \frac{|\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{t}|^2}{\boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}} \tag{157}$$

Let t denote the maximum achievable SINR, i.e., SINR($\boldsymbol{t}$, $\boldsymbol{t}$). For a mismatched steering vector $\boldsymbol{v}$, t is decomposed via the Pythagorean theorem into the SINR in the direction of $\boldsymbol{v}$ and its orthogonal complement. That is,

$$t = x + y \tag{158}$$

$$x = \mathrm{SINR}(\boldsymbol{v}, \boldsymbol{t}) \tag{159}$$

$$y = \mathrm{SINR}(\boldsymbol{t}, \boldsymbol{t}) - \mathrm{SINR}(\boldsymbol{v}, \boldsymbol{t}) \tag{160}$$

The steering vector mismatch is measured by the angle

$$\theta = \arctan[(y/x)^{1/2}] \tag{161}$$

The SINRs x, y, and t will be important in the characterization of the adaptive detector's performance.

The RMB–Boroson Loss Factor. The SINR of an adaptive weight vector $\hat{\boldsymbol{w}} = \hat{\boldsymbol{R}}^{-1}\boldsymbol{v}$ is [by Eq. (101)]

$$(\mathrm{SINR\ of}\ \hat{\boldsymbol{w}}) = |\alpha|^2 \frac{|\boldsymbol{v}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{t}|^2}{\boldsymbol{v}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\hat{\boldsymbol{R}}^{-1}\boldsymbol{v}} \tag{162}$$

where $\hat{\boldsymbol{R}}$ is the sample covariance matrix of Eq. (73). Because the sample covariance matrix is assumed to be derived from random data, this SINR is also a random variable. If this random SINR is normalized by its maximum value $|\alpha|^2\boldsymbol{t}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{t}$ [obtained from the Cauchy–Schwartz inequality; cf. Eq. (102)], we obtain the random variable

$$\rho = \frac{|\boldsymbol{v}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{t}|^2}{(\boldsymbol{v}^{\mathrm{H}}\hat{\boldsymbol{R}}^{-1}\boldsymbol{R}\hat{\boldsymbol{R}}^{-1}\boldsymbol{v})(\boldsymbol{t}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{t})} \tag{163}$$

that takes values between zero and unity. The random variable ρ is called the SINR loss factor. It indicates the random SINR loss due to target mismatch and estimation. As the sample support K becomes large, $\hat{\boldsymbol{R}}$ approaches $\boldsymbol{R}$, and ρ approaches $\cos^2\theta = x/t$.

The estimation loss may be approximated by the average value of ρ (37, 38),

$$E\rho = \frac{(K - N + 1)\cos^2\theta + 1}{K + 1} \tag{164}$$

In the case of perfectly matched steering vectors ($\theta = 0$), the average loss factor is

$$E\rho = \frac{K - N + 2}{K + 1} \qquad \text{(no mismatch)} \tag{165}$$

This is the well-known *RMB* estimation loss (36) and yields the standard rules of thumb that a sample support of $K = 2N$ yields about a 3 dB loss (Fig. 13), and a sample support of $K = 5N$ yields about a 1 dB loss. In this matched case, ρ has the central beta distribution

$$f_{\mathrm{B}}(\beta;\, L + 1,\, N - 1) = \frac{K!}{L!(N-2)!}\beta^L(1 - \beta)^{N-2} \tag{166}$$

where

$$L \stackrel{\mathrm{def}}{=} K - N + 1 \tag{167}$$

is an integer parameter introduced for convenience. There is a comparable factor for the estimation loss using the sample covariance matrix [Smith (75)].

The Kelly Loss Factor and Its Distribution

In analyzing the performance of the GLRT detector discussed in the sub-subsection "GLRT Test" below, Kelly (39–41) introduced the loss factor

$$\varrho = (1 + \tilde{\boldsymbol{z}}^{\mathrm{H}}\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{z}})^{-1} \tag{168}$$

where $\tilde{z}$ and $\tilde{\boldsymbol{R}}$ represent the whitened primary data vector $\boldsymbol{z}$ and whitened sample covariance matrix $\hat{\boldsymbol{R}}$ projected into the space perpendicular to the whitened steering vector $\boldsymbol{v}$. The loss factor has a noncentral beta distribution, given by the expression

$$f_{\mathrm{P}}(\alpha) = f_{\mathrm{B}}(\alpha;\, L + 1,\, N - 1)e^{-y}{}_1F_1(K + 1;\, N - 1;\, (1 - \alpha)y) \tag{169}$$

where f_{B} is defined in Eq. (171), and

$${}_1F_1(a:\, c:\, z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(c)_k}\frac{z^k}{k!}, \qquad (a)_k = a(a+1)\cdots(a+k-1) \tag{170}$$

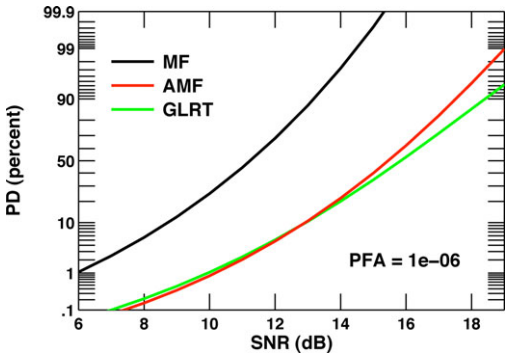

Figure 13. The probability of detection versus SNR for the MF, AMF, and GLRT adaptive detectors ($N = 42$, $K = 2N$, mismatch = 22°, PFA = 10^{-6}). The SINR loss of about 4 dB is approximately consistent with the SINR loss of 3.5 dB predicted by Eq. (169).

is the confluent hypergeometric function. Note that in the case of matched steering vectors (noncentrality parameter $y = 0$), Kelly's loss factor ϱ and the RMB loss-factor ρ are identically distributed.

Adaptive Detectors

Generalized Likelihood Ratio. Using the generalized likelihood ratio test (*GLRT*) approach (24) under a multivariate complex Gaussian assumption, Kelly (39, 40) derived the equivalent GLRTs

$$\Lambda = \frac{1 + \boldsymbol{z}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{z}}{1 + \boldsymbol{z}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{z} - \dfrac{|\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{z}|^2}{\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{v}}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda \tag{171}$$

$$\Gamma = \frac{|\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{z}|^2}{\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{v} (1 + \boldsymbol{z}^{\mathrm{H}} \hat{\boldsymbol{S}}^{-1} \boldsymbol{z})} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \gamma \tag{172}$$

where

$$\hat{\boldsymbol{S}} = K \hat{\boldsymbol{R}} = \sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \tag{173}$$

The random variable Λ lies in the range $[1, \infty)$, and Γ lies in the range $[0, 1)$. Whitening these equations shows that the GLRT is independent of the covariance matrix $\boldsymbol{R}$, i.e., it has a constant false alarm rate (CFAR property). The tests are equivalent—the random variables Λ and Γ are related by the equation

$$\Lambda = \frac{1}{1 - \Gamma} \tag{174}$$

it is the distributions of these random variables that determine the performance of the GLRT detector.

By a sequence of whitening, unitary, and other transformations, it can be shown (39–41) that the GLRT test of Eq. (176) reduces to the test

$$|\Upsilon|^2 \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} (\lambda - 1) T \tag{175}$$

where Υ is a complex Gaussian random variable conditioned on the loss factor ϱ discussed in the sub-subsection "The Kelly Loss Factor and Its Distribution" above, λ is the threshold used in Eq. (176), and T is a chi-squared random variable with $L = K - N + 1$ complex degrees of freedom. The conditional distribution of Υ is CN(0, 1) under H_0 and CN$((a/|a|)(x\varrho)^{1/2}, 1)$ under H_1, given the loss factor ϱ. Therefore, the GLRT test is equivalent to the test

$$\frac{|\Upsilon|^2 \sim \text{ noncentral chi-squared given } \varrho}{T \sim \text{ central chi-squared}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \text{ threshold}$$

which yields a noncentral F distribution (27) conditioned on the loss factor ϱ. This test is comparable to a classical CFAR detector (42).

The conditional density of the GLRT is given by the noncentral F distribution

$$f_{\Lambda}(1 + \alpha | \varrho) = \frac{L}{(1 + \alpha)^{L+1}} e^{-x\varrho} {}_1F_1\left(L + 1;\ 1;\ \frac{x\varrho\alpha}{1 + \alpha}\right) \tag{176}$$

where

$$\alpha = \lambda - 1 \tag{177}$$

The GLRT's probability of detection,

$$\mathrm{PD}_{\mathrm{GLRT}} = \int_{\alpha}^{\infty} \int_{0}^{1} f_{\Lambda}(1 + \alpha' | \varrho) f_{\mathrm{P}}(\varrho)\, d\varrho\, d\alpha' \tag{178}$$

may be obtained numerically by series methods (39, 40) or by moment methods (43). The GLRT's PFA, however, is given by the closed-form expression

$$\mathrm{PFA}_{\mathrm{GLRT}} = \frac{1}{(1 + \alpha)^L} \tag{179}$$

Adaptive Matched Filter. The AMF test (44) [Pulsone (68)], [Pulsone (67)] is obtained by the ad hoc procedure of replacing the known covariance matrix $\boldsymbol{R}$ with the sample covariance matrix $\hat{\boldsymbol{R}}$ in the Neyman–Pearson test of Eq. (91). The AMF test takes the form

$$A = \frac{|\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}|^2}{\boldsymbol{v}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{v}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \alpha \tag{180}$$

By a similar procedure to the GLRT test, the AMF test may be recast in the form

$$|\Upsilon|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \alpha \varrho T \tag{181}$$

where the random variables Υ, T, and the loss-factor ϱ are distributed identically to those in Eq. (180). Therefore, the AMF test is equivalent to the test

$$\frac{|\Upsilon|^2 \sim \text{ noncentral chi-squared given } \varrho}{T \sim \text{ central chi-squared}} \underset{H_0}{\overset{H_1}{\gtrless}} \text{threshold} \times \varrho$$

which yields a noncentral F distribution (27) conditioned on the loss factor ϱ. As with the GLRT, this test is comparable to a classical CFAR detector.

The conditional density of the AMF test is given by the noncentral F distribution

$$f_{\mathbf{A}}(\alpha|\varrho) = f_{\Lambda}(1 + \alpha\varrho|\varrho) \tag{182}$$

where $f_{\Lambda}(1 + \alpha|\varrho)$ is the GLRT's conditional density given in Eq. (182). The AMF's probability of detection

$$\mathrm{PD}_{\mathrm{AMF}} = \int_{\alpha}^{\infty} \int_{0}^{1} f_{\mathbf{A}}(\alpha'|\varrho) f_{\mathrm{P}}(\varrho)\, d\varrho\, d\alpha' \tag{183}$$

may be obtained numerically by integration methods (44) or by moment methods (43). The AMF's probability of false alarm, however, is given by the closed-form expression

$$\mathrm{PFA}_{\mathrm{AMF}} = {}_2F_1(L, L+1; K+1; -\alpha/K) \tag{184}$$

where

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!} \tag{185}$$

is the Gauss hypergeometric function and Pochammer's symbol $(a)_k$ is defined in Eq. (175).

The AMF test is seen in Fig. 12 to slightly outperform the GLRT for high SNRs. Furthermore, if there is a large sidelobe target present in the primary data vector $\boldsymbol{z}$ (which has not been modeled here), the term $\boldsymbol{z}^{\mathrm{H}}\hat{\boldsymbol{S}}^{-1}\boldsymbol{z}$ in the denominator of the GLRT test of Eq. (172) will also be large, desensitizing the GLRT and providing robustness against detecting strong sidelobe targets. The AMF test does not have this robustness. These detectors also retain near-optimality for certain classes of non-Gaussian interference, as well as other non-ideal assumptions [Richmond (71)], [Sangston (72)], [Kraut (61)].

ADAPTIVE WEIGHT COMPUTATION

Adaptive weight computation requires inverting a sample covariance matrix into a steering vector as shown in Eq. (74). Because radars must operate in real time while searching over a large combination of ranges, Doppler frequencies, and angles, and because many radar platforms can support only limited weight and power demands, the cost, accuracy, and stability the algorithm used for adaptive weight computation are all important issues, as well as the size of the processor used to implement the algorithm. In the past, the limited speeds and large weight and power requirements of digital processors forced implementors of real-time adaptive weight computation to use analog methods such as Widrow's LMS algorithm (4–6,8). Modern processor developments allow this computation to be performed digitally [Martinez (64)], [Martinez (63)] (45a–45g).

Voltage-Domain versus Power-Domain Methods

The adaptive weight computation can be performed using either the radar data themselves (the preferred choice) or the square of the radar data. The sample covariance matrix of Eq. (73) may be written in two different ways:

$$\hat{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} = \frac{1}{K}\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}} \tag{186}$$

where $\boldsymbol{Z} = (\boldsymbol{z}_1\ \boldsymbol{z}_2 \cdots \boldsymbol{z}_K)$ is an N-by-K matrix whose columns $\boldsymbol{z}_k$ $(k = 1, \ldots, K)$ are independent measurements of the interference in the absence of target returns. The direct approach for adaptive weight computation—forming the sample covariance matrix $\hat{\boldsymbol{R}}$ explicitly, then solving the linear system $\hat{\boldsymbol{R}}\hat{\boldsymbol{w}} = \boldsymbol{v}$—is undesirable for two reasons. First and most importantly, the radar data must be squared; therefore, the processor must accommodate computations involving twice the dynamic range of the data themselves. Second, the numerical error associated with this computation is proportional to the square of the radar data's dynamic range—typically a large number. The direct approach is said to be a power-domain algorithm because it is computed by squaring the radar data. The radar data themselves are said to exist in the voltage domain because they represent a sampled matched filter output.

Algorithms that use the radar data in the matrix $\boldsymbol{Z}$ directly are said to be voltage-domain algorithms. Voltage-domain approaches are highly preferred because the processor's dynamic range need only match that of the data, and the numerical error associated with these approaches is proportional to this dynamic range.

QR Decomposition

The QR decomposition is a voltage-domain method for solving many adaptive weight vectors. It relies on the fact that any N-by-K $(N \le K)$ matrix $\boldsymbol{Z}$ may be written

$$\boldsymbol{Z} = \boldsymbol{L}\boldsymbol{Q} \tag{187}$$

where $\boldsymbol{L}$ is an N-by-N lower triangular matrix and $\boldsymbol{Q}$ is an N-by-K matrix with orthonormal rows such that $\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{H}} = \boldsymbol{I}_N$. [The QR decomposition is named for its traditional form $A = QR$, where A is an arbitrary matrix, Q is a unitary matrix, and R is an upper triangular matrix. We use the lower triangular form of Eq. (194) for consistency with Eq. (193).] This decomposition may be computed using numerically stable methods such as Householder reflections or Givens rotations (26) and costs about $8N2K$ real floating-point adds and multiplies. It is amenable to parallel implementation on a multiprocessor computer architecture.

Expressed using the QR decomposition of Eq. (194), the sample covariance matrix of Eq. (193) is given by the equa-

tion

$$\hat{\boldsymbol{R}} = (1/K)\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}} \tag{188}$$

and the adaptive weight vector of Equation (74) is given by

$$\hat{\boldsymbol{w}} = (\boldsymbol{L}^{\mathrm{H}})^{-1}(\boldsymbol{Z}^{-1}\boldsymbol{v}) \tag{189}$$

Note that the $\boldsymbol{Q}$ part of the QR decomposition is not required; all information about the sample covariance matrix is stored in the matrix $\boldsymbol{L}$. Because $\boldsymbol{L}$ is lower triangular, each matrix inversion may be performed using forward substitution, and only costs about $4N^2$ real floating-point adds and multiplies. Furthermore, this low-cost operation may be repeated for many steering vectors $\boldsymbol{v}$ pointed at different angles or Doppler frequencies.

Updated *QR* Decomposition

Given the QR decomposition $\boldsymbol{Z} = \boldsymbol{L}\boldsymbol{Q}$ of the radar data matrix $\boldsymbol{Z} = (\boldsymbol{z}_1\ \boldsymbol{z}_2\ \cdots\ \boldsymbol{z}_K)$, it is sometimes necessary to update the matrix $\boldsymbol{L}$ if another measurement $\boldsymbol{z}_{K+1}$ is appended to $\boldsymbol{Z}$, or if the measurement $\boldsymbol{z}_1$ is deleted. This need arises when the trainig data are collected from a sliding range window around the rangegate of interest, perhaps with a few guard ranges deleted near this target range. As the radar steps through each rangegate looking for the presence of a target, radar data are added to the front of the training window and deleted from the rear. In this case efficient updating algorithms for the QR decomposition are used (26, 46) that cost on the order of $N2$ operations per addition or deletion.

Voltage-Domain Diagonal Loading

The benefits of diagonal loading (see the above subsection with that title) may be realized with a straightforward trick in the voltage domain. Instead of computing the QR decomposition of the radar data matrix $\boldsymbol{Z}$ as in Eq. (194), compute the QR decomposition of the matrix

$$\boldsymbol{Z}' = (K^{-1/2}\boldsymbol{Z},\ \delta^{1/2}\boldsymbol{C}) =: \boldsymbol{L}\boldsymbol{Q} \tag{190}$$

where δ is the diagonal loading level and $\boldsymbol{C}$ is any matrix such that $\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}} = \boldsymbol{R}_{\mathrm{n}}$, i.e., a Cholesky decomposition of the noise covariance matrix. Equation (197) implies that

$$\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}} = \hat{\boldsymbol{R}} + \hat{\delta}\boldsymbol{R}_{\mathrm{n}} \tag{191}$$

so that using this $\boldsymbol{L}$ in Eq. (196) yields the diagonally loaded weight vector of Eq. (146). Note that in many instances the noise covariance is white, so that $\boldsymbol{C} = \boldsymbol{I}$ may be used. In cases where the noise covariance is white and the preprocessing matrix $\boldsymbol{T}$ is used, as discussed above in the subsection "Preprocessing for Adaptive Filtering" and the section "Space–Time Adaptive Processing," in the matrix $\boldsymbol{C} = \boldsymbol{T}^{\mathrm{H}}$ may be used, or the $\boldsymbol{L}$ part of the QR decomposition of $\boldsymbol{T}^{\mathrm{H}}$ [as defined in Eq. (194)].

ADAPTIVE RADAR ESTIMATION

The final step of the adaptive radar signal-processing chain is the determination of a target's position and velocity. If the target SINR is high, these parameters may be estimated accurately; if it is low, then our estimates will be less accurate. The Cramér–Rao bound (47–49) [Van Trees (76)], [Dogandžić (56)], [Smith (74)]. provides the best accuracy achievable by any unbiased estimator of the signal parameters and therefore provides a fundamental limit on radar accuracy. At lower SINRs, resolution of closely-spaced targets affects estimation accuracy [Smith (74)], [Richmond (70)], [Richmond (69)].

Cramér–Rao Bounds for Estimation Accuracy

Given the probability distribution $f(\boldsymbol{z}|a, \boldsymbol{p})$ of the radar measurements $\boldsymbol{z}$ that depends upon an unknown complex amplitude a and the P-vector of real parameters $\boldsymbol{p} = (p_1, p_2, \ldots, p_P)^{\mathrm{T}}$, the Cramér–Rao bound of the parameter p_k is given by the inequality

$$(\text{variance of } \hat{p}_k) \geq (\boldsymbol{C}_p)_{kk} \qquad (k = 1, 2, \ldots, P) \tag{192}$$

where $\hat{p}_k$ is any unbiased estimator of p_k and $\boldsymbol{C}_p$ is the lower right P-by-P block from the inverse of the Fisher information matrix

$$\boldsymbol{G} = -E\frac{\partial^2 \log f(\boldsymbol{z}|a, \boldsymbol{p})}{\partial(a, a^*, \boldsymbol{p})^2} \tag{193}$$

corresponding to the unknown parameters in $\boldsymbol{p}$. The derivatives with respect to the complex amplitude a and its conjugate are defined below Eq. (105). Assume the known-covariance statistical model for adaptive radar of Eq. (82) for a space–time steering vector $\boldsymbol{v}(\boldsymbol{p})$ where $\boldsymbol{p} = (f_{\mathrm{d}}, \phi, \theta)^{\mathrm{T}}$ is a parameter vector containing the target parameters. Given this model, Equation (200) yields

$$\boldsymbol{G} = \begin{pmatrix} \begin{pmatrix} 0 & \xi \\ \xi & 0 \end{pmatrix} & \begin{pmatrix} \boldsymbol{x}^{\mathrm{T}} \\ \boldsymbol{x}^{*\mathrm{T}} \end{pmatrix} \\ (\boldsymbol{x} \quad \boldsymbol{x}^*) & 2\Delta \end{pmatrix} \tag{194}$$

[cf. Equation (110)] where

$$\Delta = |a|^2\,\mathrm{Re}(\boldsymbol{v}_{\boldsymbol{p}}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v}_{\boldsymbol{p}}) \tag{195}$$

$$\boldsymbol{x} = a^*\boldsymbol{v}_{\boldsymbol{p}}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} \tag{196}$$

$$\xi = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{v} \tag{197}$$

and $\boldsymbol{v}_p$ is the N-by-3 matrix of derivatives

$$\boldsymbol{v}_{\boldsymbol{p}} = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} = \left(\frac{\partial \boldsymbol{v}}{\partial f_d}\ \frac{\partial \boldsymbol{v}}{\partial \phi}\ \frac{\partial \boldsymbol{v}}{\partial \theta}\right) \tag{198}$$

Computing the block of $\boldsymbol{G}^{-1}$ corresponding to $\boldsymbol{p}$ yields

$$(\text{covariance of } \hat{\boldsymbol{p}}) \geq \boldsymbol{C}_{\boldsymbol{p}} = \tfrac{1}{2}(\Delta - \boldsymbol{M})^{-1} \tag{199}$$

where

$$\boldsymbol{M} = \xi^{-1}\,\mathrm{Re}(\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}) \tag{200}$$

Bounds on the variances of any unbiased estimators of f_{d}, ϕ, or θ are given by the diagonal elements of $\boldsymbol{C}_p$. In general any real target parameters may be used in Eq. (205)

without modification of the other formulae provided above. Furthermore, given the change of variables

$$\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{p}) = (q_1(\boldsymbol{p}), q_2(\boldsymbol{p}), \ldots, q_P(\boldsymbol{p}))^{\mathrm{T}} \tag{201}$$

Cramér–Rao bounds on the estimates of q_k are given by the diagonal of the matrix

$$\boldsymbol{C}_q = \boldsymbol{J}_q \boldsymbol{C}_p \boldsymbol{J}_q^{\mathrm{T}} \tag{202}$$

where

$$\boldsymbol{J}_q = \partial \boldsymbol{q} / \partial \boldsymbol{p} \tag{203}$$

is the Jacobian of the given change of variables. Typically, the one-dimensional case of Eq. (209) required for changing the units of an estimate is the most frequently encountered.

If the target amplitude a were known, then these variances would simply be bounded by the diagonal elements of the matrix Δ in Eq. (202). However, the fact that the target amplitude is unknown causes a loss in estimation accuracy represented by the matrix $\boldsymbol{M}$ in Eq. (207). For the standard case of a uniform linear array [Eq. (24)] in white noise with known Doppler and elevation angle, Eq. (206) reduces to the well-known result

$$\sigma_{\mathrm{u}}^2 \geq \frac{6}{\mathrm{SNR} \cdot (N^2 - 1)} \tag{204}$$

If one desires a bound on the angular standard deviation in beamwidths, σ_{u} must be divided by 2π, resulting in the rule-of-thumb that a 12 dB *SNR* is required for 10-to-1 beamsplitting. If one desires a bound on the angular standard deviation in degrees, Eq. (209) and the change of variable $u = \sin \phi'$ from Eq. (24) must be used, resulting in the relationship $\sigma_{\phi'} = (180^\circ/\pi)|\sec \phi'|\sigma_{\mathrm{u}}$.

These bounds must be evaluated for a particular choice of amplitude a whose magnitude is determined by a desired signal-to-noise ratio

$$\mathrm{SNR} = |a|^2 \boldsymbol{v}^{\mathrm{H}} \boldsymbol{R}_{\mathrm{n}}^{-1} \boldsymbol{v} \tag{205}$$

and the parameter vector $\boldsymbol{p}$ set at some nominal values of interest. The SNR is specified instead of SINR $= |a|^2 \boldsymbol{v}^{\mathrm{H}} \boldsymbol{R}^{-1} \boldsymbol{v}$ because we wish to compute bounds that represent noise-limited performance. If, on the other hand, the SINR were specified (at, say, a detection threshold of 12.5 dB), and the signals were very close to an interference source, then we would be forced to scale $|a|$ by a large amount to achieve the desired SINR, resulting in an unrealistically high SNR and unrealistic bounds.

Joint Angle–Doppler Estimation

Nonadaptive radars typically estimate a target's angle and Doppler parameters using the monopulse method (50), which works by forming a *sum beam* [analogous to the standard steering vectors of Eq. (24) or (26)] and a *difference beam* [analogous to the derivatives of these steering vectors as in Eq. (205)]. The sum beam, which has full coherent gain in the hypothesized target direction, is sensitive to the target's presence but insensitive to its position. The difference beam, which places a null in the hypothesized target direction, is sensitive to the target's position but insensitive to its presence. Therefore, using both the sum and difference channels provides information about a target's presence and position.

In general, angle and Doppler estimation is accomplished by forming a dense grid of (adaptive) beams, then choosing the angle–Doppler beam combination with the largest output power. Estimation performance is determined by the log likelihood surface of Eq. (91) parametrized using f_{d}, ϕ, and θ. The maximum likelihood estimator is

$$(\hat{f}_{\mathrm{d}}, \hat{\phi}, \hat{\theta}) = \arg \max_{f_{\mathrm{d}}, \phi, \theta} \frac{|\boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta)^{\mathrm{H}} \boldsymbol{R}^{-1} \boldsymbol{z}|^2}{\boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta)^{\mathrm{H}} \boldsymbol{R}^{-1} \boldsymbol{v}(f_{\mathrm{d}}, \phi, \theta)} \tag{206}$$

For actual radar problems when the covariance $\boldsymbol{R}$ is unknown, the sample covariance matrix $\hat{\boldsymbol{R}}$ may be used, as in the AMF detector.

If the likelihoods of a target's angle and Doppler were independent of each other (Fig. 14, left-hand side), then these parameters could be estimated independently using a monopulse-like method. However, space–time adaptivity for clutter nulling yields distorted space–time likelihood surfaces for targets near the interference (Fig. 14, right-hand side). Angle and Doppler estimation performed independently of each other will yield inaccurate, biased estimates; therefore, joint estimation algorithms (51) must be considered. See also Ref. (54).

CONCLUSIONS

For effective surveillance against modern targets, radars must operate near their noise-limited performance and cannot be blinded by electromagnetic interference caused by jamming or clutter. Space–time adaptivity provides an effective method for robustly mitigating such interference. Incorporating adaptivity into the radar's signal processing chain affects all its traditional tasks of receiving signals, examining them for the presence of targets, and estimating the speeds and position of targets. Predictions of radar performance must also include adaptivity. These predictions may be made using simple physical and statistical assumptions about the radar and its environment. Based upon such assumptions, this article describes the performance of adaptive algorithms used to detect and estimate targets in the presence of jamming and radar ground clutter. Both the best possible performance and that achieved by practical algorithms is provided. The results provided are general and allow for the analysis of a very broad class of radar parameters.

With the goal of improving system performance in circumstances beyond the relatively simple scenarios assumed within this article, adaptive methods are being used to address increasingly challenging applications and environments. Current areas of research focus on the possibility of exploiting *a prior* knowledge, compiled databases, and a Bayesian framework when faced with the challenge of rapidly varying interference environments [Guerci (58)], [Haykin (60)]. In applications where there is multipath propagation, the spatial diversity of the multiple-input multiple-output (MIMO) channel may provide per-

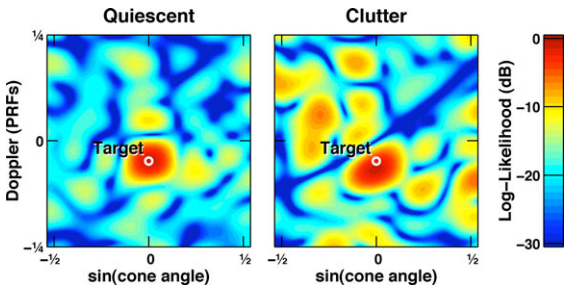

Figure 14. The log likelihood surface [Eq. (213)] for quiescent and clutter environments (SNR = 20 dB). Space–time adaptivity distorts the mainlobe seen on the right.

formance advantages in terms of additional sampling and resolution [Bliss (55)], [Fishler (57)].

NOTATION

1. $\mathbb{C}^N$
2. Re
3. Im
4. a^*
5. $\boldsymbol{x}^{\mathrm{T}}$
6. $\boldsymbol{z}^{\mathrm{H}}$
7. $\mathrm{tr}\,\boldsymbol{A}$
8. $\boldsymbol{x} \otimes \boldsymbol{y}$
9. $\boldsymbol{x} \odot \boldsymbol{y}$
10. $\boldsymbol{v}$
11. $\boldsymbol{t}$
12. $\boldsymbol{w}$
13. $\boldsymbol{R}$
14. N
15. M
16. f_{d}
17. PRI or T_{r}
18. PRF
19. CPI
20. SNR
21. CNR
22. JNR
23. SINR
24. PD
25. PFA
26. Complex N-space
27. Real part
28. Imaginary part
29. Complex conjugation of a
30. Transpose of $\boldsymbol{x}$
31. Hermitian transpose of $\boldsymbol{z}$
32. Trace of the square matrix $\boldsymbol{A}$
33. Tensor (Kronecker) product of $\boldsymbol{x}$ and $\boldsymbol{y}$
34. Schur–Hadamard product of $\boldsymbol{x}$ and $\boldsymbol{y}$
35. Steering vector
36. Target steering vector
37. Adaptive weight vector
38. Interference-plus-noise covariance matrix
39. Number of adaptive degrees of freedom or antenna elements
40. Number of radar pulses
41. Doppler frequency
42. Pulse repetition interval
43. Pulse repetition frequency
44. Coherent processing interval
45. Signal-to-noise ratio
46. Clutter-to-noise ratio
47. Jammer-to-noise ratio
48. Signal-to-interference-plus-noise ratio
49. Probability of detection
50. Probability of false alarm

BIBLIOGRAPHY

1. R. Buderi *The Invention that Changed the World*, New York: Simon and Schuster, 1996.
2. W. W. Shrader V. Gregers-Hansen MTI radar, inM. I. Skolnik (ed.), *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
3. F. M. Staudaher Airborne MTI, inM. I. Skolnik (ed.), *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
4. L. E. Brennan I. S. Reed Theory of adaptive radar, *IEEE Trans. Aerosp. Electron. Syst.*, **AES-9** (2): 237–252, 1973.
5. S. P. Applebaum Adaptive arrays, *IEEE Trans. Antennas Propag.*, **AP-24** (5): 585–598, 1976.
6. R. A. Monzingo T. W. Miller *Adaptive Arrays*, New York: Wiley, 1980.
7. R. T. Compton, Jr. *Adaptive Antennas*, Englewood Cliffs, NJ: Prentice-Hall, 1988.
8. S. Haykin *Adaptive Filter Theory*, 3rd ed., Upper Saddle River, NJ: Prentice-Hall, 1996.
9. A. Farina Electronic counter-countermeasures, inM. I. Skolnik (eds.), *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
10. A. Farina *Antenna-Based ECCM Techniques for Radar Systems*, Norwood, MA: Artech House, 1992.
11. L. E. Brennan J. D. Mallett I. S. Reed Adaptive arrays in airborne MTI radar, *IEEE Trans. Antennas Propag.*, **AP-24** (5): 607–615, 1976.

12. J. Ward Space–time adaptive processing for airborne radar, Technical Report 1015, DTIC No. ESC-TR-94-109, MIT Lincoln Laboratory, December 13, 1994.
13. J. K. Day Space–time adaptive processing from an airborne early warning radar perspective, in *Proc. 29th Asilomar Conf. Signals, Syst., Comput.*, 1995, pp. 1187–1192.
14. M. I. Skolnik *Introduction to Radar Systems*, New York: McGraw-Hill, 1980.
15. J. L. Doob *Stochastic Processes*, New York: Wiley, 1953.
16. D. R. Fuhrmann T. A. Barton Estimation of block Toeplitz covariance matrices, *Proc. 24th Asilomar Conf. Signals, Syst., Comput.*, 2: 779–783, 1990.
17. S. T. Smith Space–time clutter covariance matrix computation and interference subspace tracking, *Proc. 29th Asilomar Conf. Signals, Syst., Comput.*, 1995, pp. 1193–1197.
18. L. B. Wetzel Sea clutter, inM. I. Skolnik (ed.), *Radar Handbook*, 2nd ed., New York: McGraw-Hill, 1990.
19. J. B. Billingsley Exponential decay in windblown radar ground clutter Doppler spectra: Multifrequency measurements and model, Technical Report 997, DTIC No. ESC-TR-95-098, MIT Lincoln Laboratory, July 29, 1996.
20. C. D. Richmond A note on non-Gaussian adaptive array detection and signal parameter estimation, *IEEE Signal Process. Lett.*, **3**, 251–252, 1996.(a) F. Gini M. V. Greco L. Verrazzani Detection problem in mixed clutter environment as a Gaussian problem by adaptive pre-processing, *Elect. Lett.*, **31**: 1189–1190, 1995.(b) E. Conte M. Lops G. Ricci Adaptive matched filter detection in spherically invariant noise, *IEEE Signal Process. Lett.*, **3**: 248–250, 1996.(c) R. Raghavan N. Pulsone *A generalization of the AMF Receiver for array detection in a class of non-Gaussian interference*, in Proc. Adaptive Sensor Array Processing Workshop, MIT Lincoln Laboratory, ESC-TR-96-062, April 1996.(d) G. A. Tsihrintzis C. L. Nikias *STAP detection with sub-Gaussian distributions and fractional lower-order statistics for airborne radar*, in Proc. Adaptive Sensor Array Processing Workshop, MIT Lincoln Laboratory, ESC-TR-96-062, April 1996.(e) F. Gini *et al. Coherent adaptive radar detection in non-Gaussian sea clutter*, in Proc. 31st Asilomar Conference on Signals, Systems, and Computers, 1997.(f) C. Richmond *Adaptive array signal processing and performance analysis in non-Gaussian environment*, Ph.D. dissertation, Massachusetts Institute of Technology, 1996.
21. J. I. Marcum A statistical theory of target detection by pulsed radar: Mathematical appendix, Research Memorandum RM-753, RAND Corp., July 1, 1948; reprinted, *IRE Trans. Inf. Theory*, **IT-6** (2), 1960.
22. J. V. DiFranco W. L. Rubin *Radar Detection*, Dedham, MA: Artech House, 1980.
23. P. Swerling Probability of detection for fluctuating targets, Research Memorandum RM-1217, RAND Corp., 17 March 1954; reprinted in *IRE Trans. Inf. Theory*, **IT-6** (2), 1960.
24. H. L. Van Trees *Detection, Estimation, and Modulation Theory*, New York: Wiley, 1968.
25. P. E. Cantrell On the calculation of the generalized *Q*-function via Parl's method, *IEEE Trans. Inf. Theory*, **IT-32** (6): 817–824, 1986.
26. G. H. Golub C. F. Van Loan *Matrix Computations*, 3rd ed., Baltimore: Johns Hopkins Univ. Press, 1996.(a) A. M. Haimovich The eigencanceler: Adaptive radar by eigenanalysis methods, *IEEE Trans. Aerospace Electron. Syst.*, **32**: 532–542, 1996.(b) A. M. Haimovich and M. Berin, Eigenanalysis-based space-time adaptive radar: performance analysis, *IEEE Trans. Aerospace Electron. Syst.*, **33**: 1170–1179, 1997.(c) J. S. Goldstein I. S. Reed Theory of partially adaptive radar, *IEEE Trans. Aerospace and Electronic Syst.*, **33**: 1309–1324, 1997.
27. R. J. Muirhead *Aspects of Multivariate Statistical Theory*, New York: Wiley, 1982.
28. J. W. Silverstein The smallest eigenvalue of a large-dimensional Wishart matrix, *Ann. Prob.*, **13**: 1364–1368, 1985.
29. B. D. Carlson Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans. Aerosp. Electron. Syst.*, **24** (4): 397–401, 1988.
30. M. W. Ganz R. L. Moses S. L. Wilson Convergence of the SMI and the diagonally loaded SMI algorithms with weak interference, *IEEE Trans. Antennas Propag.*, **38** (3): 394–399, 1990.
31. G. W. Titi An overview of the ARPA/Navy Mountaintop Program, *Proc. IEEE Adaptive Antenna Syst. Symp.*, Nov. 1994.
32. G. K. Borsari A. O. Steinhardt Cost-efficient training strategies for space-time adaptive processing algorithms, *Proc. 29th Asilomar Conf. Signals, Systems, Comput.*, 1995, pp. 650–654.
33. D. J. Rabideau A. O. Steinhardt Improving the performance of adaptive arrays in nonstationary environments through data-adaptive training, *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, 1996, pp. 75–79.
34. E. J. Baranoski Improved pre-Doppler STAP algorithm for adaptive clutter nulling in airborne radars, *Proc. 29th Asilomar Conf. Signals, Systems, Comput.*, 1995, pp. 1173–1177.
35. J. W. Taylor, Jr. Receivers, inM. I. Skolnik (eds.), *Radar Handbook*, 2d ed., New York: McGraw-Hill, 1990.
36. I. S. Reed J. D. Mallett L. E. Brennan Rapid convergence rate in adaptive arrays, *IEEE Trans. Aerosp. Electron. Syst.*, **AES-10** (6): 853–863, 1974.
37. D. M. Boroson Sample size considerations for adaptive arrays, *IEEE Trans. Aerosp. Electron. Syst.*, **AES-16** (4): 446–451, 1980.
38. E. J. Kelly Adaptive detection in non-stationary interference, part 3, Technical Report 761, DITC No. ESD-TR-86-090, MIT Lincoln Laboratory, 24 August 1987.
39. E. J. Kelly Adaptive detection in non-stationary interference, parts 1 and 2, Technical Report 724, DTIC No. ESD-TR-85-197, MIT Lincoln Laboratory, 25 June 1985.
40. E. J. Kelly An adaptive detection algorithm, *IEEE Trans. Aerosp. Electron. Syst.*, **AES-22** (1): 115–127, 1986.
41. E. J. Kelly Performance of an adaptive detection algorithm: Rejection of unwanted signals, *IEEE Trans. Aerosp. Electron. Syst.*, **AES-25** (2): 122–133, 1989.
42. A. O. Steinhardt Adaptive multisensor detection and estimation, inS. Haykin andA. O. Steinhardt (eds.), *Adaptive Radar Detection and Estimation*, New York: Wiley, 1992.
43. S. T. Smith Adaptive detector statistics using moment-based approximations, *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, 1996, pp. 1118–1122.
44. F. C. Robey *et al.* A CFAR adaptive matched filter detector, *IEEE Trans. Aerosp. Electron. Syst.*, **28** (1): 208–216, 1992 and W. S. Chen and I. S. Read, A new CFAR detection test for radar, *Dig. Signal Proc.*, **4**: 198–214, 1991.
45. B. Widrow J. M. McCool A comparison of adaptive algorithms based on the methods of steepest descent and random search, *IEEE Trans. Antennas Propag.*, **AP-24** (5): 615–637, 1976.(a) C. R. Ward P. J. Hargrave J. G. McWhirter A novel algorithm and architecture for adaptive digital beamforming, *IEEE Trans. Antennas and Propagation*, **AP-34** (3): 338–346,

1986.(b) F. Ling D. Manolakis J. G. Proakis A recursive modified Gram-Schmidt algorithm for least-squares estimation, *IEEE Trans. on Acoust., Speech, Signal Process.*, **ASSP-34** (4): 829–836, 1986.(c) C. M. Rader Wafer-scale integration of a large systolic array for adaptive nulling, *Lincoln Lab. J.*, **4** (1): 3–30, 1991.(d) H. Liu A. Ghafoor P. H. Stockman Application of Gram-Schmidt algorithm to fully adaptive arrays, *IEEE Trans. Aerosp. Electron. Syst.*, **28** (2): 324–334, 1992.(e) L. Timmoneri *et al.* QRD-based MVDR algorithms for adaptive multipulse antenna array signal processing, *IEEE Proc. Radar Sonar Navig.*, **141**: 93–102, 1994.(f) P. Bollini *et al.* QR vs IQR algorithms for adaptive signal processing: performance evaluation for radar applications, *IEEE Proc. Radar Sonar Navig.*, **143**: 328–340, 1996.(g) J. O. McMahon Space-Time adaptive processing on the mesh synchronous processor, *Lincoln Lab. J.*, **9** (2): 131–152, 1996.

46. C. M. Rader A. O. Steinhardt Hyperbolic Householder transformations, *IEEE Trans. Acoust. Speech Signal Process.*, **ASSP-34** (6): 1589–1602, 1986.
47. C. R. Rao Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.*, **37**: 81–89, 1945.See also *Selected Papers of C. R. Rao*, S. Das Gupta (ed.), New York: Wiley, 1994.
48. P. Stoica *et al.* Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements, *IEEE Trans. Acoust. Speech Signal Process*, **37** (3): 378–392, 1989.
49. P. Stoica R. L. Moses *Introduction to Spectral Analysis*, Upper Saddle River, NJ: Prentice-Hall, 1997.
50. D. K. Barton H. R. Ward *Handbook of Radar Measurement*, Englewood Cliffs, NJ: Prentice-Hall, 1969.
51. J. Ward Maximum likelihood angle and velocity estimation with space–time adaptive processing radar, *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, 1996, pp. 1265–1267.
52. L. J. Griffiths C. W. Jim An alternative approach to linearly constrained optimum beamforming, *IEEE Trans. Antennas Propag.*, **AP-30**: 27–34, 1982.
53. C. G. Khatri C. R. Rao Effects of estimated noise covariance matrix in optimal signal detection, *IEEE Trans. Acoust. Speech Signal Process.*, **ASSP-35** (5): 671–679.
54. A. L. Swindlehurst P. Stoica Maximum likelihood methods in radar array signal processing, *Processing of the IEEE*, **86**: 421–441, 1998.

Additional References

Bliss, D.W.; Forsythe, K.W.,"Multiple-input multiple-output (MIMO) radar and imaging: degrees of freedom and resolution," *Signals, Systems and Computers, 2003*.Conference Record of the Thirty-Seventh Asilomar Conference on, vol. 1, no.pp. 54–59 Vol. **1**, 9–12 Nov. 2003.

A. Dogandžić and A. Nehorai," Cramér-Rao bounds for estimating range, velocity, and direction with an active sensor array," *IEEE Trans. on Signal Processing*, Vol. **SP-49**, pp. 1122–1138, June 2001.

E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, Jr., D. Chizhik, and R. A. Valenzuela, Spatial Diversity in Radars—Models and Detection Performance, *IEEE Transactions on Signal Processing*, Vol.**54**, No. 3,March 2006, pp. 823–838

J. R. Guerci and E. J. Baranoski, Knowledge-Aided Adaptive Radar at DARPA, *IEEE Signal Processing Magazine*, January 2006, pp. 41–50.

J. R. Guerci, *Space-Time Adaptive Processing for Radar*. Boston: Artech House, 2003.

S. Haykin, Cognitive Radar, *IEEE Signal Processing Magazine*, January 2006, pp. 30–40.

S. Kraut, L. L. Scharf, and L. T. McWhorter, Adaptive subspace detectors, *IEEE Trans. Signal Processing*, Vol.**49**, No. 1,January 2001, pp. 1–16.

D. R. Martinez, T. J. Moeller and K. Teitelbaum, Application of Reconfigurable Computing to a High Performance Front-End Radar Signal Processor, *Journal of VLSI Signal Processing Systems*, Volume**28**, Numbers 1/2,pp. 63–83,May/June 2001.

D. R. Martinez, F. Lee, and M. D. Davis, Space-Time Adaptive Technology Applied to Airborne Early Warning Radars,41st Annual Tri-Service Radar Symposium,27–29 June 1995.

D. R. Martinez and J. V. MacPhee, Real-Time Testbed for Space-Time Adaptive Techniques, *IEEE Proceedings of the 1994 Adaptive Antenna Systems Symposium*, 7–8 November, 1994, pp. 135–142.

J. Li, P. Stoica, and Z. Wang, On Robust Capon Beamforming and Diagonal Loading, *IEEE Trans. on Signal Processing*, Vol.**51**, No. 7,JULY 2003, pp. 1702–1715.

P. Parker and A. Swindlehurst. Space-Time autoregressive filtering for matched subspace STAP, *IEEE Trans. Aerospace and Electronic Systems*, Vol.**39**, No. 2,April 2003, pp. 510–520.

N. B. Pulsone and C. M. Rader. Adaptive beamformer orthogonal rejection test, *IEEE Trans. on Signal Processing*, Vol.**49**, No. 3,March 2001, pp. 521–529.

N. B. Pulsone and M. A. Zatman. A computationally efficient two-step implementation of the GLRT, *IEEE Trans. on Signal Processing*, Vol.**48**, No. 3,March 2000, pp. 609–616.

C. D. Richmond, Mean-Squared Error and Threshold SNR Prediction of Maximum-Likelihood Signal Parameter Estimation With Estimated Colored Noise Covariances. *IEEE Trans. Information Theory*, Vol.**52**, No. 5,May 2006, pp. 2146–2164.

C. D. Richmond. Capon Algorithm Mean Squared Error Threshold SNR Prediction and Probability of Resolution, *IEEE Trans. Signal Processing*, Vol.**53**, No. 8,August 2005, pp. 2748–2764.

C. D. Richmond. A Note on Non-Gaussian Adaptive Array Detection and Signal Parameter Estimation. *IEEE Signal Processing Letters*, Vol.**3**, No. 8,August 1996, pp. 251–252.

K. J. Sangston, F. Gini, M.V. Greco, and A. Farina, Structures for Radar Detection in Compound Gaussian Clutter, *IEEE Trans. on Aerospace and Electronic Systems*, Vol.**35**, No. 2,pp. 445–458,April 1999.

A. O. Steinhardt and N. B. Pulsone. Subband STAP processing, the fifth generation. *Proc. 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop*, 16–17 March 2000, pp. 1–6.

S. T. Smith. Statistical Resolution Limits and the Complexified Cramér-Rao Bound, *IEEE Trans. Signal Processing*, Vol.**53**, No. 5,May 2005, pp. 1597–1609.

S. T. Smith. Covariance, Subspace, and Intrinsic Cramér-Rao Bounds, *IEEE Trans. Signal Processing*, Vol.**53**, No. 5,May 2005, pp. 1610–1630.

H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, New York: Wiley-Interscience, 2002.

STEVEN T. SMITH
Lincoln Laboratory,
Massachusetts Institute of
Technology, Lexington, MA