Chapter 16

Adaptive polarization design for target detection and tracking

Martin Hurtado\textsuperscript{1,*}, Sandeep Gogineni\textsuperscript{2} and Arye Nehorai\textsuperscript{2}

Abstract

Transmitting waveforms with different polarizations in radar systems provide more complete information about the target and its environment, ensuring a significant enhancement of the radar’s performance. Conventional polarimetric radars transmit waveforms with a fixed polarization pattern, independent of the target and clutter characteristics. In this chapter, we explore the adaptive design of radar polarization waveforms. We focus on a closed-loop system that sequentially estimates the target and clutter scattering parameters and then uses these estimates to select the polarization of the subsequent waveforms. We demonstrate that the radar system performance is significantly improved when the polarization of the transmitted signal is optimally and adaptively selected to match the polarimetric aspects of the target and the environment. In particular, we include an overview of our recent results in polarimetric design for radar detection and tracking.

\textit{Keywords}: Polarimetric radar; mono-static radar; MIMO radar; detection; tracking; adaptive design.

16.1 Introduction

Radar systems transmit electromagnetic (EM) waves, collect their returns and process the recorded data to acquire information of a remote target or scene. The orientation of the oscillations of the electric and magnetic fields in the plane perpendicular to direction of travel is called \textit{polarization}. Multiple polarization states

\textsuperscript{1} Department of Electrical Engineering, National University of La Plata, Argentina
\textsuperscript{2} Department of Electrical and Systems Engineering, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA
\textsuperscript{*} Previously, Martin Hurtado was with the Department of Electrical and Systems Engineering, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA
Waveform design and diversity for advanced radar systems

of an EM signal enable it to capture multiple-copy information of a target, which results in so-called polarization diversity. Polarimetric diversity has become an important tool for detecting and tracking targets of small radar cross-sections. Unlike conventional radar systems, which operate with identically polarized antennas for transmission and reception, polarimetric radar transmits and receives waveforms with different polarizations, leading to the compatibility of acquiring complete polarimetric information of the target and the environment. Polarization provides more complete information about the target/environment features, such as the geometry, material and orientation. Exploiting polarimetric information can greatly enhance radar capabilities, particularly when the usual signal descriptors, such as time, frequency and bearing, are not sufficient for discriminating the target from the clutter/environment.

The effort of applying polarization diversity to enhance the radar performance can be traced back to the 1950s (see References 1, 2 and the references therein). In Reference 3, Sinclair formulated a model to characterize an antenna when transmitting a polarized wave and calculated the voltage at the sensor output when receiving waves of any arbitrary polarization. In Reference 4, Kennaugh demonstrated that there exist signal polarization states for which the radar receives maximum power. This idea of optimal polarization was later extended by Huynen [5]. In Reference 6, Ioannidis and Hammers proposed a method for selecting the optimum antenna polarizations for discriminating targets in the presence of clutter. In more recent work, Novak et al. [7,8] derived the optimal polarimetric detector. Moreover, they extended the use of product models to the full polarimetric case to account for the effects of non-homogeneous clutter. Several authors have demonstrated that polarization can enhance the radar resolution when jointly processed with other signal features, such as bearing, frequency or code [9–13]. The problem of polarimetric waveform design for improving the target detection and identification is addressed in Reference 14. Most of the existing literature about polarization diversity explores the performance of radar systems that transmit waveforms with a fixed polarization pattern (e.g. alternating between $H$ and $V$ polarized signals).

In this chapter, we demonstrate that the detection and tracking performances of radar systems are significantly improved when the polarization of the transmitted signal is optimally and adaptively selected to match the polarimetric aspects of the target and the environment. We provide an overview of our recent results showing that the adaptive design of the radar signal polarization enables achieving optimal performance in several operating modes [15–18]. In particular, we discuss three problems related to polarimetric waveform design.

We first discuss the problem of polarized waveform design for optimal target detection. We present a detection test statistic with a closed-form expression that incorporates information about the estimated polarimetric aspects of the target and the clutter. The analysis of the detection performance is used to adaptively schedule the next transmission polarization to enhance the target detection. We select the signal polarization that maximizes the non-centrality parameter of the detection statistic distribution under the assumption that the target is present [15].

Next, we examine the problem of optimal design for target detection using polarimetric multiple-input multiple-output (MIMO) radar systems with widely separated
antennas. These systems exploit spatial diversity in addition to the polarization diversity provided by conventional single-input single-output (SISO) polarimetric radar systems. Each transmitter is capable of adaptively choosing the polarization of its transmitted waveform based on the knowledge of the environment [16]. We analyse the performance of the detector by deriving approximate expressions for the probabilities of detection and false alarm. Using these expressions, we choose the optimal transmit waveform polarizations. We demonstrate significant improvement in performance due to optimal polarimetric design.

When the detection statistic exceeds the threshold, indicating the presence of a target, the tracking system is initiated in order to sequentially estimate the target parameters. Hence, we consider the problem of adaptive polarized waveform design for tracking targets in the presence of clutter under a framework of sequential Bayesian inference. We implement the tracking algorithm using a sequential Monte Carlo method that is suitable for non-linear and non-Gaussian state and measurement models. We discuss a criterion for selecting the optimal waveform polarization one step ahead by computing a recursive form of the posterior CRB (PCRB) [17].

16.2 Target detection in heavy inhomogeneous clutter

The detection of static or slowly moving targets in heavy-clutter environments is considered a challenging problem, mainly because it is not possible to discriminate the target from the clutter using the Doppler effect. Polarization diversity provides additional information that enhances the detection of targets, particularly under the conditions described above. Detection performance could be further improved if the polarization of the transmitted signal were optimally selected to match the target polarimetric aspects. In this section, we present a polarimetric detector that is robust against heavy inhomogeneous clutter, i.e. the detector false-alarm rate is insensitive to changes in the clutter, while still maintaining a good probability of detection. The test statistic derived from the detector has a well-known distribution that depends on the transmitted waveform parameters. Finally, we present an approach to select the signal polarization that will maximize the target probability of detection.

16.2.1 Polarimetric radar model

We consider a mono-static radar capable of transmitting waveforms with any arbitrary polarization on a pulse-by-pulse basis. The recorded data consist not only of the target echoes but also of the undesired reflections from the target environment (see Figure 16.1). We note that in order to fully identify the polarimetric aspects of the target and clutter, the radar dwell must consist of diversely polarized pulses. The output of a diversely polarized array of $Q$ sensors receiving the echoes from a single range-cell under test can be expressed as

$$y(t) = B(S^t + S^c)\xi(t) + e(t), \quad t = 1, \ldots, N$$

(16.1)
Waveform design and diversity for advanced radar systems

Figure 16.1 Geometry of the problem: object of interest (target) in an environment (clutter) producing undesired echoes

where

- The $Q \times 1$ vector $y(t)$ is the complex envelope of the measurements.
- The $Q \times 2$ matrix $B$ is the response to the diversely polarized sensor array. If the receiver array is a vector sensor [11], the array response is given by

\[
B = \begin{bmatrix}
-\sin \phi & -\cos \phi \sin \psi \\
\cos \phi & -\sin \phi \sin \psi \\
0 & \cos \psi \\
-\cos \phi \sin \psi & \sin \phi \\
-\sin \phi \sin \psi & -\cos \phi \\
\cos \psi & 0
\end{bmatrix} \tag{16.2}
\]

where $\phi$ and $\psi$ are the azimuth and elevation angles of the cell under test, respectively. If the array is a tripole antenna [9], then

\[
B = \begin{bmatrix}
-\sin \phi & -\cos \phi \sin \psi \\
\cos \phi & -\sin \phi \sin \psi \\
0 & \cos \psi
\end{bmatrix} \tag{16.3}
\]

For a conventional polarized radar measuring the horizontal and vertical components of the electric field and assuming these two sensors are orthogonal to the direction that points towards the cell under test, the array response matrix is $B = I_2$.

- The complex scattering matrix $S$ represents the polarization change of the transmitted signal upon its reflection on the target or clutter:

\[
S = \begin{bmatrix}
s_{11} & s_{12} \\
s_{21} & s_{22}
\end{bmatrix} \tag{16.4}
\]

where for a specific polarization basis, the variables $s_{11}$ and $s_{22}$ are co-polar scattering coefficients and $s_{12}$ and $s_{21}$ are cross-polar coefficients. For the monostatic radar case, $s_{12} = s_{21}$. Frequently, the polarization basis are the horizontal and vertical linearly polarized components; however, other polarization basis less
Adaptive polarization design for target detection and tracking

commonly used are left and right circular polarization, and left and right slant polarization. The superscripts \( t \) and \( c \) refer to the target and clutter.

- The vector \( \xi(t) \) is the narrowband transmitted signal, which can be represented by
  \[
  \xi(t) = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}
  \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}
  \begin{bmatrix} \cos \beta \\ j \sin \beta \end{bmatrix} s(t)
  \]  
  (16.5)

  where \( \xi_1 \) and \( \xi_2 \) are the signal components on the polarization basis of the transmitter, \( \alpha \) and \( \beta \) are the orientation and ellipticity angles, respectively, and \( s(t) \) is the complex envelope of the transmitted signal.

- The vector \( e(t) \) represents the thermal noise corrupting the radar measurements.

- \( N \) denotes the number of samples per pulse.

Equation (16.1) can be written as a linear equation in terms of the scattering coefficients:

\[
y(t) = s(t)B\tilde{\xi}(\mu + x) + e(t)
\]

(16.6)

where the scattering coefficient vectors of the target and clutter are \( \mu = [s_{t11}, s_{t22}, s_{t12}]^T \) and \( x = [s_{c11}, s_{c22}, s_{c12}]^T \), respectively, which have dimension \( P = 3 \). The polarization matrix \( \tilde{\xi} \) is

\[
\tilde{\xi} = \begin{bmatrix}
\xi_1 & 0 & \xi_2 \\
0 & \xi_2 & \xi_1
\end{bmatrix}
\]

(16.7)

The time samples can be stacked in one vector of dimension \( NQ \times 1 \):

\[
y = (s \otimes B\tilde{\xi})(\mu + x) + e
\]

(16.8)

where \( s = [s(1), \ldots, s(N)]^T \) and \( \otimes \) is the Kronecker product. Piling together the data models corresponding to a train of \( K \) pulses with different polarization yields to a single snapshot of the range cell under test

\[
y = A\mu + Ax + e
\]

(16.9)

where

\[
A = \begin{bmatrix}
s \otimes B\tilde{\xi}_1 \\
\vdots \\
s \otimes B\tilde{\xi}_K
\end{bmatrix}
\]

(16.10)

and \( \tilde{\xi}_k \) is the polarization matrix of each diversely polarized pulse \( (k = 1, \ldots, K) \); this matrix has dimension \( M \times P \), with \( M = KNQ \).

By observing the second term of expression (16.9), we note that the target recorded data are being corrupted by the clutter reflections. Since the latter depend also on the transmitted signal (which is included in the matrix \( A \)), this problem can be classified as a signal-dependent noise problem [19]. Assume that the target is a small man-made object. Hence, \( \mu \) is a deterministic vector. On the other hand, the clutter in the range cell under test can be considered as a large collection of point scatterers producing incoherent reflections of the radar signal. Then, \( x \) is a zero-mean complex
Waveform design and diversity for advanced radar systems

Gaussian random vector with covariance matrix $\Sigma$ [20, Chapter 15]. The noise $e$ is a zero-mean complex Gaussian random vector with covariance matrix $\sigma I_M$, where $I_M$ is the $M \times M$ identity matrix, since we consider that the thermal noise measurements are independent from sensor to sensor and each has the same power. In addition, assume that the clutter reflections and the thermal noise are statistically independent.

Frequently, the radar dwell consists of a sequence of snapshots of the range cell under test. If the duration of the pulses that form the snapshots is short with respect to the dynamic of the target and its environment, it is reasonable to assume that their scattering coefficients are constant during each pulse. However, from pulse to pulse, we consider the clutter scattering coefficients as independent realizations of the same random process (16.1). Then, the distribution of each snapshot is

$$y_d \sim \mathcal{CN}(A\mu, A\Sigma A^H + \sigma I_M), \quad d = 1, \ldots, D$$

where $\mathcal{CN}$ denotes a complex normal (Gaussian) distribution and $D$ is the total number of snapshots in the radar dwell such that $D > M$.

Model (16.9) can be rewritten by merging the two first terms and defining $x \sim \mathcal{CN}(\mu, \Sigma)$, without modifying the statistical model of the data given by (16.11). However, (16.9) has more intuitive insight, since it explicitly shows that $\mu$ and $x$ represent different objects: the target and the clutter, respectively.

The main difference between active and passive sensing systems is that for the former, the waveform and the direction in which it has been transmitted are known. Moreover, it is reasonable to assume that the receiver antenna array has been calibrated. Hence, the system response matrix $A$ is known. Assume that the power of the thermal noise $\sigma$ is known, since it can be easily estimated from the recorded data when no signal has been transmitted. We have no prior knowledge about the target and the clutter. Hence, the vector $\mu$ and the matrix $\Sigma$ are the unknown parameters of the statistical data model (16.11).

16.2.2 Detection test

The problem of interest is to decide whether a target is present or not in the range cell under test, based on the recorded data [21]. More formally, the decision problem consists of choosing between two possible hypotheses: the null hypothesis $\mathcal{H}_0$ (target-free hypothesis) or the alternative hypothesis $\mathcal{H}_1$ (target-present hypothesis). It can be stated as a parameter test

$$\begin{align*}
\mathcal{H}_0: & \quad \mu = 0, \Sigma \\
\mathcal{H}_1: & \quad \mu \neq 0, \Sigma
\end{align*}$$

(16.12)

where the matrix $\Sigma$ is considered as a nuisance parameter.

It is well known that the optimal detector is the likelihood ratio test [22], which provides maximum probability of detection ($P_D$) given a certain probability of false

---

1 This assumption may not be valid at high pulse-repetition frequency (PRF). However, the correlation can be reduced by decimating or sampling the snapshot data.
alarm \((P_{FA})\). Because of the lack of complete knowledge of the data distribution, likelihood ratio test cannot be applied to our problem. One possible alternative is the generalized likelihood ratio (GLR) test in which the unknown parameters of the data distribution are replaced by their maximum likelihood estimates (MLE) in the likelihood ratio test [22]. Although the GLR test does not have the optimality property described before, it appears to work well in practice.

(1) **GLR test**: The logarithmic GLR test decides \(\mathcal{H}_1\) if

\[
\ln L_{GLR} = \ln f_1(y_1, \ldots, y_D; \hat{\mu}_1, \hat{\Sigma}_1) - \ln f_0(y_1, \ldots, y_D; \hat{\Sigma}_0) > \gamma \tag{16.13}
\]

where \(f_0\) and \(f_1\) are the likelihood functions under \(\mathcal{H}_0\) and \(\mathcal{H}_1\), \(\hat{\Sigma}_0\) and \(\hat{\Sigma}_1\) are the MLEs of \(\Sigma\) under \(\mathcal{H}_0\) and \(\mathcal{H}_1\), \(\hat{\mu}_1\) is the MLE of \(\mu\) under \(\mathcal{H}_1\), and \(\gamma\) is the detection threshold. For simplicity of notation, we will omit references to the data in the arguments of the functions \(f_0\) and \(f_1\) in the rest of the chapter.

Under hypothesis \(\mathcal{H}_0\), it is assumed that \(\mu = 0\); then

\[
\ln f_0(\Sigma) = -D[M \ln \pi + \ln |C| + \text{tr}(C^{-1}S_0)] \tag{16.14}
\]

where \(|\cdot|\) denotes the determinant of the matrix, \(C = A\Sigma A^H + \sigma I_M\) is the theoretical covariance matrix of the data, defined in (16.11), and \(S_0\) is the sample covariance matrix

\[
S_0 = \frac{1}{D} \sum_{d=1}^{D} y_d y_d^H \tag{16.15}
\]

The MLE of \(\Sigma\) is (see Reference 23)

\[
\hat{\Sigma}_0 = A^+ S_0 A^+ - \sigma (A^H A)^{-1} \tag{16.16}
\]

where \(A^+ = (A^H A)^{-1} A^H\) is the pseudo-inverse matrix. The logarithmic likelihood function concentrated with respect to \(\Sigma\) is given by

\[
\ln f_0(\hat{\Sigma}_0) = -D[M \ln \pi + \ln |C| + \text{tr}(C^{-1}S_0) + \ln |A^+ S_0 A^+| + \sigma^{-1} \text{tr}(\Pi^\perp S_0)] \tag{16.17}
\]

where \(\Pi^\perp = I_M - A A^+\) is the orthogonal projection matrix that projects a vector onto the space orthogonal to the one spanned by the columns of \(A\). Under hypothesis \(\mathcal{H}_1\), the likelihood function is

\[
\ln f_1(\mu, \Sigma) = -D[M \ln \pi + \ln |C| + \text{tr}(C^{-1}\tilde{C})] \tag{16.18}
\]

where

\[
\tilde{C}_1 = \frac{1}{D} \sum_{d=1}^{D} (y_d - A\mu)(y_d - A\mu)^H \tag{16.19}
\]

The MLE of the unknown parameters are

\[
\hat{\mu}_1 = A^+ \tilde{y} \tag{16.20}
\]

\[
\hat{\Sigma}_1 = A^+ S_1 A^+ - \sigma (A^H A)^{-1} \tag{16.21}
\]
where $\bar{y}$ is the sample mean vector

$$\bar{y} = \frac{1}{D} \sum_{d=1}^{D} y_d$$

and $S_1$ is the sample covariance matrix

$$S_1 = \frac{1}{D} \sum_{d=1}^{D} (y_d - \bar{y})(y_d - \bar{y})^H$$

(16.23)

The logarithmic likelihood function concentrated with respect to $\mu$ and $\Sigma$ is

$$\ln f_1(\hat{\mu}_1, \hat{\Sigma}_1) = -D[P + M \ln \pi + (M - P) \ln \sigma + \ln |A^H A|$$

$$+ \sigma^{-1} \text{tr}(\Pi_1 S_0) + \ln |A^H S_1 A^H|]$$

(16.24)

In addition, consider the following equality, which is valid for any matrix $S$ of dimension $M \times M$:

$$\ln |A^H S_1 A| = \ln |A^H S_0 A| = 2 \ln |A^H S_0 A|$$

(16.25)

Then, substituting the concentrated likelihood functions (16.17) and (16.24) in (16.13), and using (16.25), the logarithmic of the GLR statistic is

$$\ln L_{GLR} = -D(\ln |A^H S_1 A| - \ln |A^H S_0 A|)$$

(16.26)

It is straightforward to show that

$$|A^H S_0 A| = |A^H S_1 A|^2 [1 + \bar{y}^H A(A^H S_1 A)^{-1} A^H \bar{y}]$$

(16.27)

Thus, by removing the logarithmic operator, (16.26) can be rewritten as

$$L_{GLR} = [1 + \bar{z}^H (A^H S_1 A)^{-1} A^H \bar{z}]^D$$

(16.28)

Since (16.28) is a monotonically increasing function of the second term inside the brackets, an equivalent detection test statistic can be defined as

$$T_{GLR} = \bar{z}^H S_1^{-1} \bar{z}$$

(16.29)

(2) Detection performance: Let $z_d = A^H y_d$ ($d = 1, \ldots, D$), then the test statistic (16.29) can be written as

$$T_{GLR} = \bar{z}^H S^{-1} \bar{z}$$

(16.30)

where $\bar{z}$ and $S_1$ are the sample mean and covariance formed from a random sample of size $D$ of the distribution $CN(A^H A \mu, A^H A \Sigma A^H A + \sigma A^H A)$:

$$\bar{z} = \frac{1}{D} \sum_{d=1}^{D} z_d$$

(16.31)

$$S_1 = \frac{1}{D} \sum_{d=1}^{D} (z_d - \bar{z})(z_d - \bar{z})^H$$

(16.32)
Applying Corollary 5.2.1 from Reference 24, it is straightforward to verify that the detection statistic is distributed as follows:

\[ T_{\text{GLR}} \frac{D - P}{P} \sim \begin{cases} 
\mathcal{F}_{2P,2(D-P)} & \text{under } \mathcal{H}_0 \\
\mathcal{F}_{2P,2(D-P)}(\lambda) & \text{under } \mathcal{H}_1 
\end{cases} \]  

(16.33)

where \( \mathcal{F}_{v_1,v_2}(\lambda) \) denotes a non-central \( F \) distribution with \( v_1 \) and \( v_2 \) degrees of freedom and non-centrality parameter \( \lambda \). The non-centrality parameter is given by

\[
\lambda = 2\mu^H A^H (A\Sigma A^H + \sigma I_M) A^{-1} A^H \mu
\]

\[
= 2D\mu^H [A^+(A\Sigma A^H + \sigma I_M)A^+]^+ \mu
\]

\[
= 2D\mu^H [\Sigma + \sigma (A^H A)^{-1}]^{-1} \mu
\]  

(16.34)

The last term of this equality is found by using the fact that \( A^+ A = I_P \). Thus, the detection performance becomes

\[ P_{\text{FA}} = Q_{F_{2P,2(D-P)}}(\gamma) \]

\[ P_D = Q_{F_{2P,2(D-P)}(\lambda)}(\gamma) \]  

(16.35)

where \( Q \) is the right-tail probability function [22, Chapter 2] and \( \gamma \) is the detection threshold for the required probability of false alarm. In particular, note that the expression for \( P_{\text{FA}} \) does not depend on the covariance of clutter and thermal noise, nor on the transmitted signal; thus (16.29) is a CFAR test.

The corollary mentioned above has been stated for real random variables. However, the results for the complex case are similar except that there is a factor of 2 in the non-centrality parameter and the degrees of freedom of the \( F \) distribution, since the complex case has twice the number of real parameters compared with the real case. The reader can refer to Reference 25 for further information on the \( F \) distribution derived from complex normal variables.

### 16.2.3 Target detection optimization

We aim at improving target detection by optimizing the design of our system. We have shown that the target probability of detection depends on the system characteristics through the non-centrality parameter \( \lambda \), which in turn depends on the system response \( A \). We recall that matrix \( A \) carries the information of the transmitted waveform and the receiver sensor array. Our optimization approach consists of designing the matrix \( A \) in order to maximize the parameter \( \lambda \) and consequently the probability of detection. To find the value of the matrix \( A \) that maximizes the parameter \( \lambda \), we rewrite (16.34) as

\[
\frac{\lambda}{2D} = \frac{1}{\sigma} \mu^H \left[ (A^H A)^{-1} + \frac{\Sigma}{\sigma} \right]^{-1} \mu
\]

\[
= \mu^H \Sigma^{-1} \mu - \mu^H \left( \Sigma + \frac{\Sigma A^H A \Sigma}{\sigma} \right)^{-1} \mu
\]  

(16.36)
Maximizing $\lambda$, given $\mu$ and $\Sigma_1$, is equivalent to minimizing the second term of (16.36). We denote $\eta$ the vector of the waveform parameters whose entries can be features of the applied signals, e.g. bandwidth, pulse duration and polarization, or an index that corresponds to a certain signal in a waveform library. Then, the system response matrix is parameterized as $A = A(\eta)$. To improve target detection, we seek

$$\hat{\eta} = \arg \min_{\eta} \left\{ \mu^H \left[ \Sigma + \frac{\Sigma A^H(\eta)A(\eta)\Sigma}{\sigma} \right]^{-1} \mu \right\}$$

(16.37)

We mention here that in a real application the true values of $\mu$ and $\Sigma$ are not known. Instead, their estimates $\hat{\mu}_1$ and $\hat{\Sigma}_1$ should be used to obtain the optimal waveform parameters for the next transmission based on the current recorded data. Nevertheless, solving (16.37) with the true target and clutter values provides an upper bound of the detection improvement. More details and simulation results are provided in Reference 15.

### 16.3 Polarimetric MIMO radar with distributed antennas for target detection

In conventional single-antenna radar systems, the transmitter sends a signal in order to detect a target that reflects the signal towards the receiver. The attenuation experienced by the signal depends on the properties of the target. In a realistic scenario, it is highly likely that the attenuation experienced will be a function of the angle of view of the target. If the angles of view of the target are sufficiently distinct from one another, then it is highly likely that the attenuation coefficients will have very little correlation. Therefore, even if some of the attenuation coefficients are extremely small, it is highly probable that they will be compensated by the others. MIMO radar with widely separated (distributed) antennas exploits this property by obtaining different views of the target [26,27]. It employs multiple antennas to capture information from different angles, thereby exploiting the spatial diversity. We will present a radar system that combines the advantages of distributed-antenna MIMO systems with the advantages offered by optimally choosing the transmit waveform polarizations. We examine the problem of target detection for point targets.

#### 16.3.1 Signal model

Before we give the mathematical model, we describe the target and the radar system. We assume that the target is stationary and is present in the illuminated space. The target is further assumed to be point-like with a scattering matrix that depends on the angle of view. We consider a radar system that has $M$ transmit antennas and $N$ receive antennas with all the antennas widely spaced as shown in Figure 16.2. Each of the receive antennas employs a two-dimensional vector sensor that measures both the horizontal and vertical components of the received polarized signal separately. Polarimetric models exist for describing the signals received in single-antenna systems [1]. We extend these models to distributed antenna systems.
We begin by describing the signals on the transmitter side. Define the polarization vector for the \( i \)th transmitter to be \( \mathbf{t}_i = [t_{ih}, t_{iv}]^T \), where each of the entries of the polarization vectors is a complex number and \([ \cdot ]^T\) represents the transpose of \([ \cdot ]\). We further assume that \( \| \mathbf{t}_i \| = 1, \forall i = 1, \ldots, M \). The complex pulse wave shape transmitted from the \( i \)th transmit antenna is defined as \( w_i(t) \). We assume that all these transmit waveforms are orthonormal to each other for all mutual delays between them [26,27]. In other words, we assume that the cross-correlation among these different waveforms is negligible for different lags. At the receiver side, this condition helps us differentiate between the signals transmitted from different transmit antennas. In Reference 28, we studied MIMO detection problem when the signal cross-correlations are non-zero.

After transmission, the polarized waveforms will travel in space and reflect off the surface of the target towards the receivers with altered polarimetric properties. We now consider the measurements on the receiver side. The polarized signal reaching the \( j \)th receive antenna is a combination of all the signals reflecting from the surface of the target towards the \( j \)th receiver. Let \( y_j(t) \) be the complex envelope of the signal received by the \( j \)th receive antenna. Note that \( y_j(t) \) is a two-dimensional column vector consisting of the horizontal and the vertical components of the received signal, and it is expressed using a formulation similar to that presented in References 15, 29, and 30:

\[
y_j(t) = \sum_{i=1}^{M} a_{ij} S_{ij} w_i(t - \tau_{ij}) + e_j(t) \quad (16.38)
\]

where \( e_j(t) \) is the two-dimensional additive noise, \( \tau_{ij} \) is the time delay because of propagation and the attenuation is divided into two factors \( a_{ij} \) and \( S_{ij} \). \( a_{ij} \) is that part of attenuation that depends on the properties of the medium, distance between the target and radar, and so on. We assume that the coefficients \( \{a_{ij}\} \) are known because the radar has an idea about the region which it is illuminating and the properties of the medium. \( S_{ij} \) represents the scattering matrix of the target, which completely describes the change in the polarimetric properties of the signal transmitted from
the $i$th transmit antenna to the $j$th receive antenna. This represents the unknown part of the attenuation. It has four complex components and is given as

$$S^{ij} = \begin{bmatrix} S_{hh}^{ij} & S_{hv}^{ij} \\ S_{vh}^{ij} & S_{vv}^{ij} \end{bmatrix}$$

(16.39)

To separate the signals coming from different transmit antennas, the received signal is processed using a series of $M$ matched filters at each receiver. At each receiver, the $i$th matched filter corresponds to a matching with the $i$th transmit waveform. We derive the mathematical model for the MIMO radar system by using an approach similar to that presented for the single-antenna system in Reference 29. The signals at the output of the matched filters are normalized by dividing by $a_{ij}$. Note that normalization changes the variances of the normalized noise term, and hence these variances need not be the same for all transmitter–receiver pairs. The normalized vector output of the $i$th matched filter at the $j$th receiver is expressed as

$$y_{ij} = s_{ij}^t + e_{ij}$$

(16.40)

where the column vector $y_{ij} = [y_{ij}^h, y_{ij}^v]^T$ consists of the horizontal and vertical components, respectively. We have now obtained the expressions for the measurements at each of the antennas on the receiver side. Next, we perform some simple operations to express all these measurements using a linear model.

Stacking the elements of the scattering matrix $S^{ij}$ into a vector, we define $s_{ij} = [s_{ij}^{hh}, s_{ij}^{hv}, s_{ij}^{vh}, s_{ij}^{vv}]^T$. There are $MN$ such vectors, and arranging them into a single vector gives us a $4MN \times 1$ dimensional column vector:

$$s = [(s^{11})^T, \ldots, (s^{1N})^T, \ldots, (s^{M1})^T, \ldots, (s^{MN})^T]^T$$

(16.41)

Similarly, stacking the normalized outputs of the matched filters and also the corresponding additive noise components into column vectors, we define

$$y = [(y^{11})^T, \ldots, (y^{1N})^T, \ldots, (y^{M1})^T, \ldots, (y^{MN})^T]^T$$

$$e = [(e^{11})^T, \ldots, (e^{1N})^T, \ldots, (e^{M1})^T, \ldots, (e^{MN})^T]^T$$

(16.42) (16.43)

Define a set of matrices

$$P^i = \begin{bmatrix} t^i_h & t^i_v & 0 & 0 \\ 0 & 0 & t^i_h & t^i_v \end{bmatrix}$$

(16.44)

$\forall i = 1, \ldots, M$, each corresponding to a particular transmit antenna.

Using the above definitions, we express the measurement vector $y$ using the following mathematical model:

$$y = Hs + e$$

(16.45)
where

\[
H = \begin{bmatrix}
P^1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & P^1 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & P^M & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & P^M
\end{bmatrix}
\]  

(16.46)

\(0\) is a zero matrix of dimensions \(2 \times 4\). Terms \(y\) and \(e\) are \(2MN \times 1\) dimensional observation and noise vectors, respectively. Thus, we have reduced our mathematical model to the well-known linear form. We now look at the statistical assumptions made on these terms.

We assume that the noise terms present in \(e\) are uncorrelated and that \(e\) follows proper complex Gaussian distribution. A complex random vector \(\varsigma = \varsigma_R + j\varsigma_I\) is said to be proper if \(\text{Cov}(\varsigma_R, \varsigma_R) = \text{Cov}(\varsigma_I, \varsigma_I)\) and \(\text{Cov}(\varsigma_R, \varsigma_I) = -\text{Cov}(\varsigma_I, \varsigma_R)\). Hence, the covariance matrix of \(e\) will be diagonal of the form \(\sigma^2 I\). This diagonal assumption states that the noise components at the outputs of the matched filters across the various widely separated receivers over both the polarizations are statistically independent for any given time snapshot. This assumption is reasonable, given the wide separation between the antennas [26]. Define this covariance matrix as \(\Sigma_e\) and assume that it is known. The matrix \(H\) is a \(2MN \times MN\) dimensional design matrix whose constituent elements depend on the transmit waveform polarizations. We assume that the vector \(s\), which contains elements from all the scattering matrices, is a random vector following proper complex Gaussian distribution with a \(4MN \times 4MN\) covariance matrix given by \(\Sigma_s\). We further assume that \(\Sigma_s\) is known. If the random matrices \(S^w\) are statistically independent, then \(\Sigma_s\) will have a block diagonal structure. However, we do not impose any such structural constraint on \(\Sigma_s\). Furthermore, we assume that \(s\) and \(e\) are independent.

### 16.3.2 Problem formulation

The above mathematical model gives an expression for the observation vector when the target is present in the illuminated space. When the target is absent, the observations will consist of only the receiver noise vector \(e\). Hence, the problem of detecting the target reduces to the following binary hypothesis testing problem:

\[
H_0: y = e
\]

(16.47)

\[
H_1: y = Hs + e
\]

(16.48)

Therefore, under the null hypothesis, \(y\) will have complex Gaussian distribution with zero mean and covariance matrix \(\Sigma_e\). Under the alternative hypothesis, the independence of \(s\) and \(e\) implies that \(y\) will follow complex Gaussian distribution with zero mean and covariance matrix given by \(C + \Sigma_e\), where \(C = HH^H\) denotes the covariance matrix of \(Hs\). This result in an application of the well-known properties
of Gaussian random vectors [31]. Next we describe the Neyman–Pearson detector for this problem.

16.3.3 Detector

1) Test statistic: Under the above-mentioned hypotheses, the probability density functions of the observation vector are given as

\[ f(y|H_0) \propto \frac{1}{|\Sigma_e|} e^{-\frac{1}{2}y^\top \Sigma_e^{-1}y} \]  
(16.49)

\[ f(y|H_1) \propto \frac{1}{|\Sigma_e + C|} e^{-\frac{1}{2}y^\top (\Sigma_e + C)^{-1}y} \]  
(16.50)

The Neyman–Pearson lemma states that the likelihood ratio test is the most powerful test for any given size [32]. The likelihood ratio is given as

\[ \frac{f(y|H_0)}{f(y|H_1)} = \frac{|\Sigma_e + C|}{|\Sigma_e|} e^{y^\top (\Sigma_e^{-1} - (\Sigma_e + C)^{-1})y} \]  
(16.51)

Computing the logarithm of the above expression and ignoring the known constants, we clearly see that \( y^\top (\Sigma_e^{-1} - (\Sigma_e + C)^{-1})y \) is our test statistic and we compare it with a threshold before selecting a hypothesis:

\[ y^\top (\Sigma_e^{-1} - (\Sigma_e + C)^{-1})y \geq k_{H_0} \]  
(16.52)

where the threshold \( k \) is chosen based on the size specified for the test.

2) Estimating covariance matrices: In practice, the covariance matrices needed for implementing the detector may not be known in advance. In such a scenario, the MLE of these matrices can be substituted to perform the test. Since the observations follow Gaussian distribution under both the hypotheses, the MLE of the covariance matrices are given by the corresponding sample covariance matrices [32,33]. The sample covariance matrices are easy to compute in practice. We assume sufficient number of samples to obtain accurate estimates of these covariance matrices. The variance of noise at each receiver is calculated before the detector starts functioning by evaluating the sample variance using a large set of training data. The covariance matrix under the alternative hypothesis is estimated by evaluating the sample covariance matrix using all the samples of observations in a particular window of time when the detector is in use. These two estimated matrices are sufficient for implementing the detector. If there is no target in the illuminated space, then these two estimated matrices will be close to each other, thereby causing the test statistic to fall below the threshold.

3) Performance analysis: To analyse the performance of the above-mentioned detector, we need to know the distribution of the test statistic under both hypotheses. The test statistic is a quadratic form of the complex Gaussian random vector \( y \). It is well known in statistics that a quadratic form \( z^\top Uz \) of a real Gaussian random vector \( z \) with covariance matrix \( B \) will follow Chi-square distribution if and only if the matrix \( UB \) is idempotent [34]. Using this result, we infer that our test statistic does not necessarily follow Chi-square distribution for all feasible choices of \( \Sigma_e \) and \( C \).
because we did not impose any constraint on $\Sigma_e$. Hence, it is difficult to find the exact probability density function (pdf) for it. To study the pdf of our test statistic, we first begin with an assumption that $C$ is diagonal. Later, we will extend this approach to the non-diagonal case by applying proper diagonalization.

Define the $l$th diagonal element of $C$ as $c_l$ and that of $\Sigma_1$ as $v_l$. Then, the test statistic reduces to

$$
\sum_{i=1}^{M} \sum_{j=1}^{N} \left( \frac{1}{y_{ij}^{2(l-1)N+2l-1}} - \frac{1}{y_{ij}^{2(l-1)N+2l-1} + \varepsilon^{2(l-1)N+2l-1}} \right) |y_{ij}^{v_l}|^2 
$$

$$
+ \sum_{i=1}^{M} \sum_{j=1}^{N} \left( \frac{1}{y_{ij}^{2(l-1)N+2l}} - \frac{1}{y_{ij}^{2(l-1)N+2l} + \varepsilon^{2(l-1)N+2l}} \right) |y_{ij}^{v_l}|^2 
$$

where $y_{ij}^{v_l}$ are always independent Gaussian random variables under both hypotheses for all transmitter–receiver pairs because of the diagonal assumption of $\Sigma_e$ and $C$. Therefore, the test statistic is a weighted sum of independent Chi-square random variables and it does not necessarily follow the Chi-square distribution. Its actual distribution depends on the weights. The pdf of a sum of independent random variables is obtained by performing multiple convolutions among the constituent pdfs. However, in this case, it is difficult to find the exact solution. Hence, we shall look for approximations to the actual pdf.

In Reference 35, the distribution of the weighted sum of Chi-squares is studied. If $\pi_q$ are real positive constants and $N_q$ are independent standard normal random variables $\forall q = 1, \ldots, K$, then the pdf of the Gamma approximation of $R = \sum_{q=1}^{K} \pi_q N_q^2$ is given as

$$
f_R(r, \alpha, \beta) = r^{\alpha-1} e^{-\frac{r}{\beta}} \frac{1}{\beta^\alpha \Gamma(\alpha)}
$$

(16.53)

where the parameters $\alpha$ and $\beta$ are given as

$$
\alpha = \frac{1}{2} \left( \frac{\sum_{q=1}^{K} \pi_q}{\sum_{q=1}^{K} \pi_q^2} \right)^2
$$

(16.54)

$$
\beta = \frac{1}{2} \left( \frac{\sum_{q=1}^{K} \pi_q}{\sum_{q=1}^{K} \pi_q^2} \right)^{-1}
$$

(16.55)

$\Gamma$ is the gamma function defined as $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$.

Under the null hypothesis, $y_{ij}^{v_l}$ and $y_{ij}^{v_l}$ have zero mean and variances $\varepsilon^{2(l-1)N+2l-1}$ and $\varepsilon^{2(l-1)N+2l}$, respectively. Hence, applying the above approximation with appropriate weights, the parameters of the Gamma distribution are

$$
\alpha_{H_0} = \left( \frac{\sum_{l=1}^{2MN} \frac{c_l}{\varepsilon^{l+2l}}}{\sum_{l=1}^{2MN} \left( \frac{c_l}{\varepsilon^{l+2l}} \right)^2} \right)^2
$$

(16.56)
468 Waveform design and diversity for advanced radar systems

\[ \beta_{H_0} = \left( \frac{\sum_{l=1}^{2MN} \frac{c_l^2}{\sigma^2_l}}{\sum_{l=1}^{2MN} \left( \frac{c_l^2}{\sigma^2_l} \right)^2} \right)^{-1} \]  

(16.57)

Under the alternative hypothesis, \( y_{ij}^0 \) and \( y_{ij}^1 \) have zero mean and variances \( e^{2(i-1)(N+2j-1)} + c^{2(i-1)(N+2j-1)} \) and \( e^{2(i-1)(N+2j)} + c^{2(i-1)(N+2j)} \), respectively.

The parameters of the Gamma approximation are

\[ \alpha_{H_1} = \left( \frac{\sum_{l=1}^{2MN} \frac{c_l^2}{\sigma^2_l}}{\sum_{l=1}^{2MN} \left( \frac{c_l^2}{\sigma^2_l} \right)^2} \right) \]  

(16.58)

\[ \beta_{H_1} = \left( \frac{\sum_{l=1}^{2MN} \frac{c_l^2}{\sigma^2_l}}{\sum_{l=1}^{2MN} \left( \frac{c_l^2}{\sigma^2_l} \right)^2} \right)^{-1} \]  

(16.59)

Note that so far we have assumed a diagonal structure for matrix \( C \) in the aforementioned discussion. However, we still need to find expressions for the pdf of the test statistic when \( C \) is not diagonal. Diagonalization will be used to extend the analysis even for the case of non-diagonal matrices [36]. Since \( \Sigma_e \) and \( C \) are covariance matrices, \( (\Sigma_e^{-1} - (\Sigma_e + C)^{-1}) \) will be a Hermitian matrix, which therefore decomposes into \( D \Gamma \Lambda D^H \), where \( \Lambda \) is a diagonal matrix consisting of eigenvalues as the diagonal elements and \( D \) contains the corresponding orthonormal eigenvectors. The test statistic now becomes \( (Dy)^H \Lambda (Dy) \). If we show that \( Dy \) has a diagonal covariance matrix under both hypotheses, then our analysis extends to the case in which \( C \) is not diagonal also, with appropriate adjustments made to the parameters of the Gamma approximation. Under \( H_0 \), \( Dy \) is a complex Gaussian random vector with a covariance matrix \( \text{Cov}_{H_0}(Dy) = D \Sigma_e D^H \), which is diagonal \( \Sigma_e = \sigma^2 I \) and \( D \) has orthonormal vectors. Similarly, under \( H_1 \), \( Dy \) is a complex normal random vector with covariance matrix

\[ \text{Cov}_{H_1}(Dy) = D(\Sigma_e + C)D^H \]  

(16.60)

which is diagonal. Hence, under both hypotheses, the test statistic is a weighted sum of Chi-square random variables even when matrix \( C \) is not diagonal. The only difference is that the weights will now be different, and they are defined by the diagonalization process.

After approximating the pdf using the Gamma density, the probability of detection \( (P_D) \) and the probability of false alarm \( (P_{FA}) \) are defined as follows:

\[ P_D = \int_k^{\infty} t^{\alpha_{H_1}-1} e^{-\frac{t}{\beta_{H_1}}} \frac{1}{\beta_{H_1} \Gamma(\alpha_{H_1})} dt \]  

(16.64)
Adaptive polarization design for target detection and tracking

\[ P_{\text{FA}} = \int_k^{\infty} r_{H_0}^{-1} e^{-\frac{r_{H_0}}{\beta_{H_0}}} \Gamma(\alpha_{H_0}) \, dt \]  

(16.65)

where the parameters \( \alpha_{H_0}, \beta_{H_0}, \alpha_{H_1} \) and \( \beta_{H_1} \) are as mentioned earlier. For a given value of \( P_{\text{FA}} \), the value of the threshold \( k \) is calculated easily using the above expression because functions for evaluating the above expressions exist in MATLAB. After finding the threshold, \( P_D \) is calculated accordingly. Note that the value of the threshold and \( P_D \) depends on matrix \( C \), which in turn depends on the polarizations of the transmitted waveforms. Hence, the performance of the detector is related to the transmit waveform polarizations.

(4) Optimal design: To find the optimal design, we perform a grid search over the possible waveform polarizations across all the transmit antennas with the help of the above expressions for \( P_D \) and \( P_{\text{FA}} \). The optimal design corresponds to the transmit polarizations that give the maximum \( P_D \) for a given \( P_{\text{FA}} \). Later, we will plot the receiver operating characteristic (ROC) curves to visualize the improvement in performance because of the optimal design.

16.3.4 Scalar measurement model

Most of the conventional polarimetric radar systems combine the two received signals linearly and coherently at each receiver to give only a scalar measurement that depends on the receive polarization vector. For such systems, the output at each receive antenna is modeled as an inner product of the received signal and the receive antenna polarization [1,29]. This receive polarization vector is optimally chosen along with the transmit waveform polarizations in order to achieve improved performance. We now obtain the signal model for such systems. From now on, we refer to this model as the scalar measurement model.

Let \( r^j = [r_{h}^j, r_{v}^j]^T \) be the polarization vector of the \( j \)th receiver, where each of the entries is a complex number. We further assume that \( \|r^j\| = 1, \forall j = 1, \ldots, N \). The rest of the variables remain the same as defined earlier, except that the measurement and the noise at each receiver according to this model will be complex scalars. The scalar observation at the \( j \)th receiver \( y^j(t) \) is now expressed as follows [15–29]:

\[ y^j(t) = \sum_{i=1}^{M} a_{ij} r^j S^i t^i w^j(t - \tau_{ij}) + e^j(t) \]  

(16.66)

This signal is now passed through a series of matched filters whose outputs are appropriately normalized to move the effect of \( a_{ij} \) into the noise term. Finally, the normalized output of the \( i \)th matched filter at the \( j \)th receiver is given as

\[ y_{ij} = r^j S^i t^i + e^j \]  

(16.67)

Stacking all the observations and the noise components into column vectors, in a similar fashion to the approach used earlier, we obtain \( MN \times 1 \) dimensional vectors \( y \) and \( e \), respectively. Vector \( s \) remains the same as defined earlier. However, matrix \( H \)
changes and now contains the elements of the receive polarization vectors also. Let us define a set of vectors

$$\eta_{ij} = \left( \eta_{i1}^j, \eta_{i2}^j, \eta_{i3}^j, \eta_{i4}^j \right), \quad \forall i = 1, \ldots, M,$$

where each of which corresponds to a particular transmitter–receiver pair. Under this definition, the observation vector is expressed as

$$y = Hs + e$$

(16.69)

where \( H \) is a \( MN \times 4MN \) dimensional matrix given by

$$H = \begin{bmatrix}
\eta_{11} & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & \eta_{1N} & \cdots & 0 & \cdots & 0 \\
\vdots & \cdots & 0 & \cdots & \eta_{M1} & \cdots & 0 \\
0 & \cdots & 0 & \cdots & \eta_{M2} & \cdots & \eta_{MN}
\end{bmatrix}$$

(16.70)

Therefore, we obtain a similar linear model even for the systems with scalar measurements. The only difference lies in the dimensionality of some of the vectors in the model and also the constituent elements of the matrix \( H \). The optimal design for such a system will not only include optimization over the transmit polarizations \( t_i \) but will also include the optimal selection of the receive polarization vectors \( r_j \). The problem formulation and analysis of the detector remains the same as for the earlier model because the basic structure of the model is still the same.

### 16.3.5 Numerical results

We consider a system with two transmit antennas and two receive antennas under the same target detection scenario as described so far. Hence, there are 16 complex elements in the random vector \( s \). We choose the covariance matrix of this vector to be of the following form:

$$\Sigma_s = \begin{bmatrix}
\Sigma_{s1} & 0 & 0 & 0 \\
0 & \Sigma_{s2} & 0 & 0 \\
0 & 0 & \Sigma_{s3} & 0 \\
0 & 0 & 0 & \Sigma_{s4}
\end{bmatrix}$$

(16.71)

where \( \Sigma_s \) represents the covariance matrix of the random vector \( s \) and \( \Theta \) is a \( 4 \times 4 \) dimensional zero matrix. Each of these matrices was chosen as follows:
uncorrelated, with the variance of each equal to 

differ only at the third decimal point. 

values of 5 and 7.5, respectively. This shows that the values taken by these two curves 
presented in Reference 35. The sample cdf takes values 0.5827 and 0.9233, whereas 
approximation under the null hypothesis turn out to be 

turns out to be non-diagonal for this example. Hence, after performing the appro-
approximate Gamma distribution and comparing it with that formed by generating 
random samples from the constituent Chi-squares. This comparison assumes all the 

the above-mentioned parameters. In order to check if this is indeed a good approx-
in this scenario, we have the following information available: 

\[
\begin{align*}
\epsilon^1 &= [1, 0]^T \\
\epsilon^2 &= [1, 0]^T
\end{align*}
\]  

(16.76)  

(16.77) 

Therefore, the matrices \( P^1 \) and \( P^2 \) become \( P^1 = P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \). The matrix \( C \) 
turns out to be non-diagonal for this example. Hence, after performing the ap-
appropriate diagonalization and calculating the weights, the coefficients of the Gamma 
approximation under the null hypothesis turn out to be \( \alpha_{10} = 7.6833 \) and \( \beta_{10} = 
0.6283 \). Figure 16.3(b) shows the cdf of this approximated Gamma distribution with 
the above-mentioned parameters. In order to check if this is indeed a good approx-
imation, we generated random samples of the observation vector \( y \) under the null 

hypothesis. We evaluated the test statistic \( y^T (\Sigma^{-1} - (\Sigma_x + C)^{-1}) y \) for each of these 
random samples and generated the sample cdf, which is plotted in Figure 16.3(a). It is 
clear from both figures that the Gamma approximation we made is indeed very ac-
curate and close to the sample distribution. This finding is consistent with the results 
presented in Reference 35. The sample cdf takes values 0.5827 and 0.9233, whereas 
the cdf of the Gamma approximation takes values 0.5863 and 0.9242 for argument 
values of 5 and 7.5, respectively. This shows that the values taken by these two curves 
differ only at the third decimal point.

\[
\Sigma^{11} = \\
\begin{bmatrix}
0.3 & 0.1\epsilon & 0.1\epsilon & 0.1\epsilon \\
0.1\epsilon & 0.2 & 0.1\epsilon & 0.1\epsilon \\
0.1\epsilon & 0.1\epsilon & 0.4 & 0.1\epsilon \\
0.1\epsilon & 0.1\epsilon & 0.1\epsilon & 0.5 \\
\end{bmatrix}
\]  

(16.72) 

\[
\Sigma^{12} = \\
\begin{bmatrix}
0.5 & 0.05\epsilon & 0.05\epsilon & 0.05\epsilon \\
0.05\epsilon & 0.3 & 0.05\epsilon & 0.05\epsilon \\
0.05\epsilon & 0.05\epsilon & 0.4 & 0.05\epsilon \\
0.05\epsilon & 0.05\epsilon & 0.05\epsilon & 0.3 \\
\end{bmatrix}
\]  

(16.73) 

\[
\Sigma^{21} = \\
\begin{bmatrix}
0.4 & 0.1\epsilon & 0.1\epsilon & 0.1\epsilon \\
0.1\epsilon & 0.3 & 0.1\epsilon & 0.1\epsilon \\
0.1\epsilon & 0.1\epsilon & 0.2 & 0.1\epsilon \\
0.1\epsilon & 0.1\epsilon & 0.1\epsilon & 0.4 \\
\end{bmatrix}
\]  

(16.74) 

\[
\Sigma^{22} = \\
\begin{bmatrix}
0.4 & 0.05\epsilon & 0.05\epsilon & 0.05\epsilon \\
0.05\epsilon & 0.4 & 0.05\epsilon & 0.05\epsilon \\
0.05\epsilon & 0.05\epsilon & 0.2 & 0.05\epsilon \\
0.05\epsilon & 0.05\epsilon & 0.05\epsilon & 0.5 \\
\end{bmatrix}
\]  

(16.75)
Waveform design and diversity for advanced radar systems

Figure 16.3 Cumulative distribution function of the test statistic for the chosen example under the null hypothesis: (a) Sample cdf and (b) Gamma approximation

Now that we have a good enough approximation to the distribution of our test statistic, we look at how the optimal choice of polarizations improves the performance of the detector. We fix the complex noise variance to $\sigma^2 = 0.2$ and vary the value of $P_{FA}$. This method enables us to plot the optimal ROC curve by performing a grid search using the analytical results derived earlier. Next, we obtain the reference curves for our results by computing the ROC curves assuming that all the transmit antennas are horizontally or vertically polarized. These plots are presented in Figure 16.4, and a significant improvement in performance is clearly visible while using the optimal waveform polarizations.

So far, we have demonstrated that by optimally selecting the transmit polarizations, we get performance improvement over conventional MIMO systems with fixed polarizations. Now, we plot the ROC curves for SISO radar with optimal transmit polarizations to show the gain in performance because of the multiple widely separated antennas. For the SISO system, we consider only the first transmit and receive antennas in our above-mentioned example. Therefore, the covariance matrix of the scattering vector $s$ becomes $\Sigma_s = \Sigma_{11}$. To make a fair comparison, we transmit more power than the power transmitted per antenna while using MIMO radar. It is clear from Figure 16.5 that $2 \times 2$ polarimetric MIMO radar system significantly outperforms its SISO counterpart even when the SISO system uses four times the transmit power used by each antenna in the $2 \times 2$ system.

The complexity of the grid search for optimization using the vector measurement model does not increase much with the increase in the number of receivers, because the number of variables over which the optimization is performed depends only on the number of transmit antennas. However, with the scalar measurement
Adaptive polarization design for target detection and tracking

Figure 16.4 ROC curves demonstrating the improvement offered by the optimal choice of polarizations when $\sigma^2 = 0.2$

Figure 16.5 ROC curves demonstrating the improvement offered by employing multiple widely separated antennas compared with single-input single-output systems when $\sigma^2 = 0.2$
model, the addition of each extra receiver adds extra variables (receive polarization vectors) in the grid search and makes the calculations more complex. Therefore, in order to compare the performance of the vector measurement system with that of the scalar measurement system, we use the same numerical example as described so far; however, this time we stick to just two transmitters and one receiver to reduce the complexity of the optimization step. The \( \Sigma_s \) matrix now has the following form:

\[
\Sigma_s = \begin{bmatrix}
\Sigma_{s1}^{11} & 0 \\
0 & \Sigma_{s1}^{21}
\end{bmatrix}
\]  

where matrices \( \Sigma_{s1}^{11} \) and \( \Sigma_{s1}^{21} \) are chosen to be the same, as defined earlier in this section. The noise variance remains the same for both the systems because the receive polarization vectors are assumed to be unit norm. We assume the same noise variance \( \sigma^2 = 0.1 \) for both systems in order to make a fair comparison. Figure 16.6 compares the performance of both systems under the optimal choice of polarization vectors. It clearly shows that by retaining the 2D vector measurements, we get significantly improved results as compared with scalar measurement systems. Even though we perform joint optimization over both the transmit and the receive polarizations for the scalar measurement systems, we are still finding just the best linear combination of the two received measurements at each receiver. However, combining them linearly need not be the overall optimal solution and we might be losing some important information by doing so. This can be avoided by retaining the vector measurements, thereby giving better performance as demonstrated in Figure 16.6.

![Figure 16.6 Comparison of performance between systems with scalar measurements and those with 2D vector measurements as a function of the probability of false alarm when \( \sigma^2 = 0.1 \)]
Adaptive polarization design for target detection and tracking

16.4 Adaptive polarized waveform design for target tracking based on sequential Bayesian inference

We present a scheme for polarized waveform design for tracking targets in the presence of clutter. This scheme is a combination of sequential Bayesian filtering for parameter estimation and optimal transmitted waveform design in active sensing systems.

16.4.1 Sequential Bayesian framework for adaptive waveform design

This framework for adaptive waveform design includes four phases: (i) creation of a dynamic state model and a statistical measurement model, (ii) belief prediction and update, (iii) Bayesian state estimation and (iv) optimal waveform selection. They are described in details as follows:

1) Dynamic state model and measurement model: To formulate a sequential Bayesian estimation, we first consider a state sequence \( \{x_k, k \in \mathbb{N}\}, x_k \in \mathbb{R}^{n_x} \), which is assumed to be an unobserved (hidden) Markov process with initial distribution \( p(x_0) \). The evolution of the state sequence is given by

\[
x_k = f_k(x_{k-1}, v_{k-1})
\]

where \( f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_y} \) is a non-linear function of the state; \( \{v_k, k \in \mathbb{N}\} \) is a process noise sequence and \( n_x \) and \( n_v \) are the dimensions of the state and process noise vectors, respectively. This state model represents our prior knowledge about, e.g. the underlying dynamic movement of a target.

We also have a sequence of measurements \( \{y_k, k \in \mathbb{N}\}, y_k \in \mathbb{R}^{n_y} \). These measurements are related to the current state vector via the observation equation:

\[
y_k = h_k(x_k, e_k)
\]

where \( h_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_e} \rightarrow \mathbb{R}^{n_y} \) is a non-linear function; \( \{e_k, k \in \mathbb{N}\} \) is a measurement noise sequence and \( n_x \) and \( n_e \) are the dimensions of the measurement and noise vectors, respectively.

2) Belief prediction and update: We denote by \( x_{0:k} \triangleq \{x_0, \ldots, x_k\} \) and \( y_{1:k} \triangleq \{y_1, \ldots, y_k\} \), respectively, the state sequence and the observations up to \( k \). Under the Bayesian inference framework, all relevant information about \( x_{0:k} \) given observations \( y_{1:k} \) can be obtained from the posterior probability density (also called belief) \( p(x_{0:k} | y_{1:k}) \). Therefore, our aim is to estimate recursively in time the distribution \( p(x_{0:k} | y_{1:k}) \) and its associated features, including \( p(x_k | y_{1:k}) \).

To derive a recursive Bayesian inference process, we consider that the following conditional independent assumptions for a first-order hidden Markov process are satisfied.

A1: Conditioned on \( x_k \), the current measurements \( y_k \) are independent of the past states \( x_{0:k-1} \) and past measurement history \( y_{1:k-1} \), i.e.

\[
p(y_k | x_{0:k}, y_{1:k-1}) = p(y_k | x_k)
\]
A2: Conditioned on \( x_{k-1} \), the current state \( x_k \) is independent of the states \( x_{0:k-2} \) and past measurement history \( y_{1:k-1} \), i.e.

\[
p(x_k \mid x_{0:k-1}, y_{1:k-1}) = p(x_k \mid x_{k-1})
\]  

(16.82)

Based on the above assumptions, we obtain recursive formulas to calculate the new belief \( p(x_{0:k} \mid y_{1:k}) \) when the new measurements \( y_k \) are available, as follows:

\[
p(x_{0:k} \mid y_{1:k-1}) = p(x_k \mid x_{k-1})p(x_{0:k-1} \mid y_{1:k-1})
\]  

(16.83)

and

\[
p(x_{0:k} \mid y_{1:k}) = \frac{p(y_k \mid x_k)p(x_{0:k} \mid y_{1:k-1})}{p(y_k \mid y_{1:k-1})}
\]  

(16.84)

where

\[
p(y_k \mid y_{1:k-1}) = \int p(y_k \mid x_k)p(x_{0:k} \mid y_{1:k-1})dx_{0:k}
\]  

(16.85)

For linear and Gaussian state and measurement models, the above equations become Kalman filters.

Equations (16.83) and (16.84) form a procedure for belief prediction and update in a recursive belief propagation. In the prediction stage (16.83), we use the probabilistic model of the state transition \( p(x_k \mid x_{k-1}) \) and the measurement history \( y_{1:k-1} \) to predict the prior pdf of the state at the \( k \)th time step. In the update stage (16.84), the current measurement \( y_k \) (via the likelihood function \( p(y_k \mid x_k) \)) is used to modify the prior density \( p(x_k \mid y_{1:k-1}) \) to obtain the belief at the current time step.

3 Bayesian state estimation: At the \( k \)th time step, after obtaining the current belief \( p(x_k \mid y_{1:k}) \), we can obtain an optimal estimate of the current state \( x_k \). In target tracking, this estimate can be used to determine the current target states (e.g. position and velocity) and environment parameters. Under the Bayesian framework, the estimate is calculated by optimizing a utility function. For example, when we apply a minimum-mean-squared error (MMSE) criterion, the estimate is the mean of the belief \( p(x_k \mid y_{1:k}) \).

4 Optimal waveform selection: In optimal waveform selection, we use the information from the current belief \( p(x_k \mid y_{1:k}) \), together with the state transition distribution and measurement model, to optimally select the waveform one step ahead in response to the target state and the environmental situation. Hence, we can achieve the best possible sensing performance.

To derive a mathematical formulation for optimal waveform selection, we first create a utility function according to certain criteria that represent the sensing performance; then, we determine the parameters for the next transmitted waveform by optimizing (e.g. maximizing) this utility function. We denote by \( J(\cdot) \) the utility function, \( \theta_{k+1} \) the waveform parameters at the \((k+1)\)th time step, and \( y_{k+1}(\theta_{k+1}) \) the measurements at the \((k+1)\)th time step. At the current time step \( k \), we select the next transmitted waveform \( \theta_{k+1}^* \) to be

\[
\theta_{k+1}^* = \arg \max_{\theta_{k+1} \in \Theta} J[p(x_{k+1} \mid y_{1:k}, y_{k+1}(\theta_{k+1}))]
\]  

(16.86)
Adaptive polarization design for target detection and tracking

where Θ denotes the set of the allowed values for θ_{k+1} or a library of possible waveforms.

We note that the former utility function is related to the belief at the (k + 1)th time step. To determine this belief, we need the measurements y_{k+1}, which are not available at the current time step k. Therefore, we compute the utility function J(·) by marginalizing out the particular value of y_{k+1}. We observe that for any given y_{k+1}, we obtain a particular value for J(·) acting on the new belief p(x_{k+1} \mid y_{1:k}, y_{k+1}(θ_{k+1})).

Now for each waveform parameter θ_{k+1} we consider the set of all values of J(·) for different values of y_{k+1}. Possibilities for summarizing the set of values of J(·) by a single quantity include the average, the worst or the best case [37]. For example, if we use the average as a utility, the next transmitted waveform is selected by

\[ θ^*_{k+1} = \arg \max_{θ_{k+1} \in Θ} E_{y_{k+1} \mid y_{1:k}} \{ J[p(x_{k+1} \mid y_{1:k}, y_{k+1}(θ_{k+1}))] \} \]  

(16.87)

where \( E_{y_{k+1} \mid y_{1:k}} \{ \cdot \} \) represents the average over the set of new belief weighted by \( p(y_{k+1} \mid y_{1:k}) \).

We note that many tracking applications require fast real-time processing. The trade-off between performance and computation cost should be considered when choosing the utility function J(·).

16.4.2 Target dynamic state model and measurement model

We first create a dynamic state model for target tracking. Based on this model, we can track the target position, velocity and scattering coefficients. We then derive a measurement model that is the output of the receiver sensor array. This model provides a natural way of incorporating the polarimetric aspects of the target and clutter into the tracking filter.

1. Target dynamic state model: In our state model, we include the target scattering coefficients that are important, for example, for target identification and classification [1]. We denote by \( S_t \) the complex scattering matrix representing the polarization change of the transmit signal upon its reflection on the target:

\[ S_t = \begin{bmatrix} s_{hh} & s_{hv} \\ s_{vh} & s_{vv} \end{bmatrix} \]  

(16.88)

The scattering matrix of the target can be written in terms of the radar polarization basis as [38]

\[ S_t = R^T S_d R \]  

(16.89)

where

- \( R \) is a unitary transformation matrix from the target eigenbasis to the radar basis

\[ R = \begin{bmatrix} \cos ϑ & \sin ϑ \\ -\sin ϑ & \cos ϑ \end{bmatrix} \cdot \begin{bmatrix} \cos ε & j \sin ε \\ j \sin ε & \cos ε \end{bmatrix} \]  

(16.90)

where ϑ is the orientation angle of the target eigenbasis around the line of sight and relative to the radar (−90° ≤ ϑ ≤ 90°), and ε is the ellipticity of the target (−45° ≤ ε ≤ 45°).
Waveform design and diversity for advanced radar systems

- $S_d$ is a diagonal matrix representing the target scattering matrix in its eigenpolarization basis

$$S_d = me^{j\rho} \begin{bmatrix} e^{j2\nu} & 0 \\ 0 & \tan^2 \gamma e^{-j2\nu} \end{bmatrix}$$

(16.91)

where $m$ is the maximum target amplitude; $\rho$ is the absolute phase of the scattering matrix ($-180^\circ \leq \rho \leq 180^\circ$); $\nu$ is called the skip angle, which is associated with the depolarization of the reflected signal ($-45^\circ \leq \nu \leq 45^\circ$); and $\gamma$ is called the characteristic angle, representing the ability of the target to polarize an incident unpolarized field ($0^\circ \leq \gamma \leq 45^\circ$) [1]. These four parameters $\{m, \rho, \nu, \gamma\}$ do not change with the target orientation about the line of sight; hence, they are called invariant parameters. The decomposition of the scattering matrix for the non-reciprocal case (i.e. $s_{hv} \neq s_{vh}$) can be found in Reference 39.

Then, we represent the target state at the $k$th time step as

$$x_k = [\rho_k^T, s_k^T]^T$$

(16.92)

where $\rho_k = [x_k, y_k, z_k, \dot{x}_k, \dot{y}_k, \dot{z}_k]^T$ includes the target position and velocity at the $k$th time step in a Cartesian coordinate system, and $s_k = [\varrho_k, \epsilon_k, m_k, \vartheta_k, \nu_k, \gamma_k]^T$ represents the target scattering parameters.

We assume that (i) the target movement is characterized by a constant velocity and random acceleration, (ii) the target scattering parameters are nearly constant and have random rate of change and (iii) the position and velocity are statistically independent of the scattering coefficients. Then, we obtain a linear target dynamic state model given by

$$x_k = F x_{k-1} + v_{k-1} = \begin{bmatrix} F_\rho & 0 \\ 0 & F_s \end{bmatrix} x_{k-1} + v_{k-1}$$

(16.93)

where

- $F_\rho$ is the transition matrix for states $\rho$ as

$$F_\rho = \begin{bmatrix} I_3 & T_{PRI} I_3 \\ 0 & I_3 \end{bmatrix}$$

(16.94)

where $I_n$ denotes the identity matrix of size $n$, and $T_{PRI}$ is the pulse repetition interval (PRI). $F_s = I_6$ is the transition matrix for state $s$.

- $v_k$ is the independent process noise, representing the uncertainty about the state model and is assumed to be zero-mean Gaussian distributed with covariance matrix $Q$:

$$Q = \begin{bmatrix} Q_\rho & 0 \\ 0 & Q_s \end{bmatrix}$$

(16.95)
where $Q_{\rho}$ and $Q_{s}$ denote the covariance matrices for the target acceleration and rate of change of the scattering parameters [40]:

\[
Q_{\rho} = q_{\rho} \begin{bmatrix}
T_{\text{PRI}}^4/4 & 0 & 0 & T_{\text{PRI}}^3/2 & 0 & 0 \\
0 & T_{\text{PRI}}^4/4 & 0 & 0 & T_{\text{PRI}}^3/2 & 0 \\
0 & 0 & T_{\text{PRI}}^4/4 & 0 & 0 & T_{\text{PRI}}^3/2 \\
T_{\text{PRI}}^3/2 & 0 & 0 & T_{\text{PRI}}^2 & 0 & 0 \\
0 & T_{\text{PRI}}^3/2 & 0 & 0 & T_{\text{PRI}}^2 & 0 \\
0 & 0 & T_{\text{PRI}}^3/2 & 0 & 0 & T_{\text{PRI}}^2 \\
\end{bmatrix}
\]

(16.96)

\[
Q_{s} = q_{s} T_{\text{PRI}}^2 I_6
\]

and $q_{\rho}$ and $q_{s}$ are constants.

In this state model, the assumption that the target scattering coefficients vary slowly is suitable for a situation in which the target is far away from the sensor array and the target position change during the tracking period is not large compared with the distance between the target and the sensor array.

In general, the dynamic model for the scattering coefficients is a non-linear function with respect to other states; hence, the target dynamic state model will be non-linear. In some cases, it is difficult even to determine a closed-form dynamic transition model for the scattering coefficients. One solution is to assume the state transition density $p(s_{k+1} \mid s_{k})$ to be a uniform distribution centred at $s_{k}$ with a radius equal to the possible maximum value of the change of the scattering coefficients during $T_{\text{PRI}}$. That is, we do not provide any prior information about the change of $s_{k}$ except that $s_{k+1}$ will be within a certain range.

(2) Statistical measurement model: We consider a target characterized by azimuth $\phi$, elevation $\psi$, range $r$, Doppler shift $\omega_D$ and scattering matrix $S$. These parameters are related to the states $x$ in (16.92). To uniquely identify the polarimetric aspects of a target, polarization diversity is required and the complete EM information of the signal reflected from the target has to be processed [41]. To provide these measurements, we employ an array of EM vector sensors [11] as the receiver, where each sensor measures the six components of the EM field (three electric and three magnetic components of the received signal).

Consider an array of $M$ vector sensors receiving the signal returns from a target. The complex envelope of the measurements can be expressed as

\[
y(t) = A(\phi, \psi) S \xi(t - \tau)e^{j\omega_D t} + e(t), \quad t = t_1, \ldots, t_N
\]

(16.97)

where

- The matrix $A(\phi, \psi) = p(\phi, \psi) \otimes V(\phi, \psi)$ is the array response, where $\otimes$ is the Kronecker product; $[\phi, \psi]^T$ is the bearing angle vector; $p(\phi, \psi) = [e^{j2\pi u^T r_1/\lambda}, \ldots, e^{j2\pi u^T r_M/\lambda}]^T$ represents the phase of the plane wave arriving from the direction given by the vector $u = [\cos \phi \cos \psi, \sin \phi \cos \psi, \sin \psi]^T$ at the
position $r_m$ of the $m$th sensor ($m = 1, \ldots, M$); $\lambda$ is the signal wavelength; and
$V(\phi, \psi)$ is the response of a single vector sensor given by Reference 11:

$$V(\phi, \psi) = \begin{bmatrix} -\sin \phi & -\cos \phi \sin \psi \\ \cos \phi & -\sin \phi \sin \psi \\ 0 & \cos \psi \\ -\sin \phi \sin \psi & \sin \phi \\ -\sin \phi \sin \psi & -\cos \phi \\ \cos \psi & 0 \end{bmatrix}$$  \hspace{1cm} (16.98)$$

• The polarized transmit wave $\xi(t)$ is a narrowband signal that can be represented by a complex vector $[1,11]$:

$$\xi(t) = \begin{bmatrix} \xi_h(t) \\ \xi_v(t) \end{bmatrix} = g(t)Q(\alpha)w(\beta)$$  \hspace{1cm} (16.99)$$

where

$$Q(\alpha) = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}, \quad w(\beta) = \begin{bmatrix} \cos \beta \\ \sin \beta \end{bmatrix}$$  \hspace{1cm} (16.100)$$

Angles $\alpha$ and $\beta$ are the orientation and ellipticity of the polarization ellipse. The function $g(t)$ represents the scalar complex envelope of the transmitted pulse. The time delay $\tau = 2r/c$, where $r$ is the distance from the target to the sensor array and $c$ is the wave propagation velocity.

• The vector $e(t)$ is the additive noise corrupting the radar measurements; it represents the thermal noise at the sensors and the reflections from the clutter (target environment).

• $N$ denotes the number of samples during the pulse repetition interval $T_{PRI}$.

Since $\xi(t)$ is the transmitted signal, the waveform design problem consists of selecting the envelope $g(t)$ and the polarization angles $\alpha$ and $\beta$ in (16.99). We denote these waveform parameters by $\theta$.

It can be verified that the relationship between the target parameters $[\phi, \psi, r, \omega_D, S_t]$ and the states $x = [\rho^T, s^T]^T$ is given by

$$\phi = \arctan \left( \frac{y}{x} \right)$$  \hspace{1cm} (16.101a)$$

$$\psi = \arctan \left( \frac{z}{\sqrt{x^2 + y^2}} \right)$$  \hspace{1cm} (16.101b)$$

$$r = \sqrt{x^2 + y^2 + z^2}$$  \hspace{1cm} (16.101c)$$

$$\omega_D = \frac{2\omega_c}{c} \frac{\dot{x}x + \dot{y}y + \dot{z}z}{\sqrt{x^2 + y^2 + z^2}}$$  \hspace{1cm} (16.101d)$$

$$S_t = S_t(s)$$  \hspace{1cm} (16.101e)$$

where $\omega_c$ is the carrier frequency, and the relation between $S_t$ and $s$ is given in (16.89)–(16.91). When we insert (16.101) into the measurement model (16.97), we
observe a non-linear relationship between measurements $y(t)$ and state $x$. We write this non-linear relationship at the $k$th time step as

$$y_k(t) = h(t, x_k; \theta_k) + e_k(t)$$

(16.102)

where

$$\tilde{h}(t, x; \theta) = A(\phi, \psi) S_c \xi (t - \tau) e^{j \omega_D t}$$

(16.103)

When we lump $\{y_k(t), t = t_1, \ldots, t_N\}$ together into a vector, we obtain the following as measurement model:

$$y_k = \begin{bmatrix} y_k(t_1) \\ \vdots \\ y_k(t_N) \end{bmatrix} = \begin{bmatrix} \tilde{h}(t_1, x_k; \theta_k) \\ \vdots \\ \tilde{h}(t_N, x_k; \theta_k) \end{bmatrix} + \begin{bmatrix} e_k(t_1) \\ \vdots \\ e_k(t_N) \end{bmatrix}$$

(16.104)

$$(3) \text{ Polarimetric clutter model:}$$ The measurement noise $e(t)$ represents not only the thermal noise at the sensors of the receiver but also the reflections from the environment surrounding or behind the target. We aim to represent by this model the clutter reflections, for example, in the case for which a target flies above a sea or land surface.

It is well known that the clutter response is highly dependent on the transmit signal polarization [1]. We propose a polarimetric clutter model that explicitly accounts for the polarization of the illuminating signal, and only the clutter scattering coefficients are represented by a random vector. For estimating the statistical parameters of this random vector, training data recorded with simple two different polarized pulses are required [41].

The transmit signal illuminates both the target and the clutter, and their reflections are recorded by the same receiver. Hence, we propose a noise model, similar to measurement model (16.97), as

$$e(t) = A(\phi_0, \psi_0) S_c \xi (t - t_0) + n(t), \quad t = t_1, \ldots, t_N$$

(16.105)

where $n(t)$ is the additive thermal noise and $S_c$ is the scattering matrix of the clutter. The angles $[\phi_0, \psi_0]$ are the direction in which the radar beam is been steered, which might be different from the target angles. The clutter delay $t_0$ is related to the average clutter position, and it may also differ from the target delay. For our cases of interest, we consider that the clutter does not introduce Doppler shift; i.e. the clutter velocity can be neglected when compared with the target velocity. The clutter scattering coefficients are random variables because they represent the reflections from many incoherent point scatterers constituting the clutter. Following the model in Reference 41, (16.105) can be rearranged to express the clutter scattering coefficients in a vector:

$$e(t) = A(\phi_0, \psi_0) \xi (t - t_0) S_c + n(t), \quad t = t_1, \ldots, t_N$$

(16.106)
482 Waveform design and diversity for advanced radar systems

where

\[
\xi(t) = \begin{bmatrix} \xi_h(t) & 0 & \xi_v(t) \\ 0 & \xi_v(t) & 0 \end{bmatrix}
\] (16.107)

and

\[
\tilde{S}_c = \begin{bmatrix} s_{h} \xi_v \\ s_{v} \xi_v \\ s_{v} \xi_h \end{bmatrix}
\] (16.108)

where the variables \( s \) are the scattering coefficients of the clutter.

We assume that the thermal noise and the clutter scattering coefficients can be modelled as

\[
n(t) \sim CN(0, \sigma^2 I_{6M}), \quad \tilde{S}_c \sim CN(0, \Sigma_c)
\] (16.109)

where \( \sigma^2 \) is the noise power, and the clutter covariance matrix can be parameterized as [41]:

\[
\Sigma_c = \begin{bmatrix} \sigma^2_s Q(\vartheta_c)w(\epsilon_c)w(\epsilon_c)^H + \sigma^2_u I_2 & 0 \\ 0 & p_x \end{bmatrix}
\] (16.110)

where \( \sigma^2_s \) and \( \sigma^2_u \) are the power of the polarized and unpolarized components of the clutter, \( \vartheta_c \) and \( \epsilon_c \) are the orientation and ellipticity angles of the clutter, matrix \( Q(\cdot) \) and vector \( w(\cdot) \) are defined as in (16.100), and \( p_x \) is the power of the cross-polarized component of the clutter.

(4) Polarized waveform structure: The design of the polarized waveform involves selecting the parameters of the signal envelope \( g(t) \) and its polarization in (16.99). Here, we consider as an example a linear frequency modulated (LFM) pulse with Gaussian envelope, which is defined as

\[
g(t) = (\pi \eta^2)^{-1/4} \exp \left[-\left(\frac{1}{2\eta^2} - jb\right) t^2\right]
\] (16.111)

where \( \eta \) is the pulse length and \( b \) is the frequency sweep rate. The signal bandwidth is \( BW = 7.4 \eta b \) [42]. Then, we propose to use the following scheme of polarized waveform [43]:

\[
\xi(t) = \sum_{l=0}^{L-1} g(t - l T_{EPL})Q(\alpha_l)w(\beta_l)
\] (16.112)

where \( L \) is the number of transmitted LFM pulses and \( T_{EPL} = 7.4 \eta \) is the effective pulse length [42]. Under this scheme, the waveform parameters are \( \Theta = [\eta, b, \alpha_0, \beta_0, \ldots, \alpha_{L-1}, \beta_{L-1}]^T \).

Note that if the scattering matrix is completely unknown, at least two pulses with different polarization should be transmitted, i.e. \( L > 0 \), to uniquely identify \( S_c \).

16.4.3 Target tracking using sequential Monte Carlo methods

In this section, we develop a target-tracking method based on the proposed dynamic state model (16.93) and the statistical measurement model (16.104). Since these models are non-linear, we propose a sequential Monte Carlo method (particle filter),
which is based on point mass representation of probability densities and is powerful for solving non-linear and non-Gaussian Bayesian inference problems.

In contrast to the ordinary sequential Monte Carlo methods, in our proposed approach, we adopt a Gibbs sampler to draw samples from an importance sampling function \([44]\) through which we can handle the potentially large dimension of a state vector. We first describe the ordinary sequential importance sampling (SIS) particle filter and then we discuss the use of other possible importance sampling functions.

1) Sequential importance sampling particle filter: The sequential Monte Carlo method is a technique for implementing a recursive Bayesian filter by Monte Carlo simulations \([44]\). The key idea is to represent the required posterior density function by a set of random samples with associated weights and to compute estimates based on these samples and weights.

Let \(\{x_{0:k}^{(i)}, w_{k}^{(i)}, i = 1, \ldots, N_s\}\) denote a random measure that characterizes the belief \(p(x_{0:k} \mid y_{1:k})\), where \(\{x_{0:k}^{(i)}, i = 1, \ldots, N_s\}\) is a set of support points with associated weights \(\{w_{k}^{(i)}, i = 1, \ldots, N_s\}\). Then, the belief at the \(k\)th time step can be approximated as

\[
p(x_{0:k} \mid y_{1:k}) \approx \sum_{i=1}^{N_s} w_k^{(i)} \delta(x_{0:k} - x_{0:k}^{(i)}) \tag{16.113}
\]

where the weights are chosen using the principle of importance sampling \([44]\). Let \(\{x_{0:k}^{(i)}, i = 1, \ldots, N_s\}\) be samples that are easily generated from a proposal importance density function \(q(x_{0:k} \mid y_{1:k})\). Then, the weights in (16.113) are given by Reference 45

\[
w_{k}^{(i)} \propto \frac{p(x_{0:k}^{(i)} \mid y_{1:k})}{q(x_{0:k}^{(i)} \mid y_{1:k})} \tag{16.114}
\]

For a sequential filtering case where only \(p(x_k \mid y_{1:k})\) is required at each time step, we can choose the importance density \(q(\cdot)\) such that we obtain a weight update equation \([46]\):

\[
w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(y_k \mid x_k^{(i)}) p(x_k^{(i)} \mid x_{k-1}^{(i)})}{q(x_k^{(i)} \mid x_{k-1}^{(i)}, y_k)} \tag{16.115}
\]

and the belief \(p(x_k \mid y_{1:k})\) can be approximated as

\[
p(x_k \mid y_{1:k}) \approx \sum_{i=1}^{N_s} w_k^{(i)} \delta(x_k - x_k^{(i)}) \tag{16.116}
\]

where \(\{x_{0:k}^{(i)}, i = 1, \ldots, N_s\}\) are sampled from the importance density \(q(x_k \mid x_{k-1}^{(i)}, y_k)\).

2) Gibbs sampling–based particle filter: Considering our target tracking problem, from the dynamic state model (16.93) we observe that if we want to track the target position, velocity and scattering coefficients simultaneously, the dimension of the state space is large. Drawing samples directly from the importance density \(q(x_k \mid x_{k-1}^{(i)}, y_k)\) is typically inefficient. Hence, we apply a Markov chain Monte Carlo (MCMC) method, a class of iterative simulation-based methods, to sample from the
importance density. MCMC methods are a set of procedures that enable the successful solution of simulation problems for more complex models [47]. The basic idea of MCMC methods is to simulate an ergodic Markov chain whose samples are asymptotically distributed according to a desired density function. In our work, we adopt a classical MCMC algorithm – the Gibbs sampler. Given state \( \theta \), the Gibbs sampler consists of first defining a partition of the components of \( \theta \) as \( \theta_1, \ldots, \theta_p \) \((p \leq \dim(\theta))\), and then sampling successively from the full conditional distributions \( p(\theta_i | \theta_{-i}) \), where \( \theta_{-i} \triangleq (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p) \).

In our developed particle filter, we choose the importance density \( p(x_k | x_{k-1}^{(i)}) \), \( i = 1, \ldots, N_s \). We adopt the above Gibbs sampling and propose the following method to draw samples from \( p(x_k | x_{k-1}^{(i)}) \). According to the state model (16.93), we partition the components of \( x_i \) as \( x_i = [\rho_k^T, s_k^T]^T \), where \( \rho_k \) includes the target position and velocity and \( s_k \) includes the target scattering parameters. Then, we derive a Gibbs sampling algorithm to draw samples \( x_k^{(i)} \sim p(x_k | x_{k-1}^{(i)}) \) at the \( k \)th time step in a particle filter. Such a Gibbs sampling is described as follows:

1. Initialization, \( j = 0 \). Set randomly or deterministically:
   \[
   x_k^{(j,0)} = \left[ (\rho_k^{(0)}), (s_k^{(0)}) \right]^T
   \]

2. Iteration \( j, j = 1, \ldots, M \), where \( M \) is a large number.
   - Sample \( \rho_k^{(j)} \sim p(\rho_k | s_k^{(j-1)}, x_k^{(i-1)}) \)
   - Sample \( s_k^{(j)} \sim p(s_k | \rho_k^{(j)}, x_k^{(i-1)}) \)
   - Installation of \( \rho_k^{(M)} \) and \( s_k^{(M)} \) into \( x_k^{(i)} \):
     \[
     x_k^{(i)} = \left[ (\rho_k^{(M)}), (s_k^{(M)}) \right]^T
     \]

Then, the obtained \( x_k^{(i)} \) is a sample from \( p(x_k | x_{k-1}^{(i)}) \).

In a special case where the partitions \( \rho \) and \( s \) are statistical independent of each other, the Gibbs sampling can be simplified as

- Sample \( \rho_k^{(i)} \sim p(\rho_k | \rho_{k-1}^{(i)}) \)
- Sample \( s_k^{(i)} \sim p(s_k | s_{k-1}^{(i)}) \)

Then, we obtain \( x_k^{(i)} = \left[ (\rho_k^{(i)}), (s_k^{(i)}) \right]^T \)

(3) **Discussion:** In the above proposed Gibbs sampling–based particle filter, we use the simplest importance density function \( p(x_k | x_{k-1}^{(i)}) \). However, this importance function does not take into account the current measurements \( y_k \), and the state space is explored without any knowledge of the observations. Therefore, the filter can be inefficient and it is sensitive to outliers. A natural strategy to overcome this disadvantage is to use an optimal importance function that minimizes the variance of the importance weights conditional upon the states \( x_{0:k-1}^{(i)} \) and the measurements \( y_{1:k} \). Such an optimal importance function is given as [46]

\[
q(x_k | x_{k-1}^{(i)}, y_k) = p(x_k | x_{k-1}^{(i)}, y_k)
\] (16.117)
and the importance weight in (16.115) becomes
\[ w_k^{(i)} \propto w_{k-1}^{(i)} p(y_k \mid x_k^{(i)}) \] (16.118)

However, this optimal importance function suffers from two drawbacks: it requires the ability to sample from \( p(x_k \mid x_{k-1}^{(i)}, y_k) \), which is not easy, and it requires the evaluation of \( p(y_k \mid x_k^{(i)}) = \int p(y_k \mid x_k) p(x_k \mid x_{k-1}^{(i)}) \, dx_k \). This integral has no analytical form in general cases. A practical method to overcome this drawback is to use a Gaussian density to approximate the optimal importance function, which allows us to easily draw samples. The parameters of this Gaussian importance function are evaluated using a local linearization of the original optimal importance function [46,48]. This method can be extended to use a sum of Gaussian densities to approximate the optimal importance function, which can provide a more accurate approximation when the optimal importance function is multi-modal.

### 16.4.4 Optimal waveform design based on posterior Cramér–Rao bounds

Now, we propose a new optimal waveform design method for target tracking. This method is based on the proposed dynamic state model and the statistical measurement model in (16.93) and (16.104), respectively. It is combined with the aforementioned target-tracking algorithms and forms an adaptive waveform design scheme.

To pursue the optimization at the \( k \)th time step, we develop an algorithm that predicts the tracking performance at the \((k+1)\)th time step when employing specific waveform parameters. Then, we select the waveform parameters that optimize a certain criterion. Since the target tracking methods are derived under a sequential Bayesian inference framework, we design the waveform selection criterion based on a posterior Cramér–Rao bound (CRB).

(1) Posterior Cramér–Rao bounds: For random parameters, as in our sequential Bayesian filter for target tracking, a lower bound that is analogous to the CRB in a non-random parameter estimation exists and is derived in References 49 and 50. This lower bound is usually referred to as a posterior CRB (PCRB) or a Bayesian CRB. We denote by \( y \) a vector of measurements and by \( x \) a vector of random parameters to be estimated. Let \( p(y, x) \) be the joint pdf of the pair \((y, x)\), and \( \hat{x} = g(y) \) be an estimate of \( x \). Then, the PCRB on the mean-square estimation error satisfies

\[ \Sigma = \mathbb{E}_{y, x} [ (g(y) - x)(g(y) - x)^T ] \geq J^{-1} \] (16.119)

where \( J \) is the Bayesian information matrix (BIM), \( J^{-1} \) is the PCRB, \( \mathbb{E}_{y, x}[\cdot] \) denotes expectation with respect to \( p(y, x) \) and the inequality in the equation means that the difference \( \Sigma - J^{-1} \) is a non-negative definite matrix. Let \( \Delta^\psi_\eta \) be the \( m \times n \) matrix of second-order partial derivatives with respect to the \( m \)-dimensional parameter \( \psi \) and \( n \)-dimensional parameter vector \( \eta \), i.e.

\[ \Delta^\psi_\eta = \begin{bmatrix} \frac{\partial^2}{\partial \psi_1 \partial \eta_1} & \cdots & \frac{\partial^2}{\partial \psi_1 \partial \eta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial \psi_m \partial \eta_1} & \cdots & \frac{\partial^2}{\partial \psi_m \partial \eta_n} \end{bmatrix} \] (16.120)
Waveform design and diversity for advanced radar systems

Using this notation, the BIM for \( x \) is defined as \([50]\)

\[
J = E_{x,y}[-\Delta^x y \log p(y, x)]
\]  (16.121)

From this property, we observe that the PCRB is a lower bound on the error covariance matrix \( \Sigma \), and it is related only to the state and measurement models and independent of the specific estimation methods. Hence, we can use the PCRB as a precise measure of the tracking system performance.

(2) Criterion for optimal waveform selection: Consider our target tracking problem: at the \( k \)th time step, we want to estimate the state \( x_k \) using the measurements \( y_{1:k} \). We denote by \( X_k = [x_0^T, \ldots, x_k^T]^T \) the sequence of states up to time \( k \). Then, the BIM of the target states, whose inverse is the PCRB, is defined as

\[
\tilde{J}_k \triangleq E_{x_{1:k}, x_{0:k}}[-\Delta^x_{x_{1:k}} \log p(y_{1:k}, x_{0:k})]
\]  (16.122)

This BIM and the corresponding PCRB \( \tilde{J}_k^{-1} \) are \((k + 1)n_c \times (k + 1)n_c\) matrices. The lower right \( n_c \times n_c \) block of \( \tilde{J}_k^{-1} \) is the PCRB for estimating \( x_k \), and its inverse is the BIM for estimating \( x_k \), denoted by \( J_k \). According to this definition, in our optimal waveform selection algorithm, at the \( k \)th time step we design a criterion based on the BIM \( J_{k+1} \) to select the waveform to be transmitted at the \((k + 1)\)th time step.

To derive the optimal waveform selection criterion, we adopt the recursive equation in Reference 50 to update BIM \( J_{k+1} \). For the particular case of a linear state model with additive Gaussian noise, this recursive BIM can be written as (see Reference 51)

\[
J_{k+1}(\theta_{k+1}) = [Q + FJ_k(\theta_k)^{-1}F^T]^{-1} + \Gamma_{k+1}(\theta_{k+1})
\]  (16.123)

where \( \theta_k \) and \( \theta_{k+1} \) are the waveform parameters at time step \( k \) and \( k + 1 \), respectively, the matrices \( F \) and \( Q \) are defined in (16.93) and (16.95), and

\[
\Gamma_{k+1}(\theta_{k+1}) = E_{y_{k+1}, x_{k+1}}[-\Delta^y_{y_{k+1}} \log p(y_{k+1}, x_{k+1}) | x_{k+1}]
\]  (16.124)

In our sequential waveform design algorithm, we attempt to minimize the error on the estimation of the target state using the information provided by the state and measurement models and the measurement history \( y_{1:k} \). Hence, we modify the matrix \( \Gamma_{k+1} \) to include the measurement history and design a criterion based on a new matrix \( \tilde{\Gamma}_{k+1} \):

\[
\tilde{\Gamma}_{k+1}(\theta_{k+1}) = E_{y_{k+1}, x_{k+1}, y_{1:k}}[-\Delta^y_{y_{k+1}} \log p(y_{k+1}, \theta_{k+1}) | x_{k+1}]
\]  (16.125)

By replacing \( \Gamma_{k+1} \) with \( \tilde{\Gamma}_{k+1} \), we use the information from the measurement history \( y_{1:k} \) to improve our prior knowledge on the state \( x_{k+1} \). Mathematically, we replace the prior density \( p(x_{k+1}) \) with \( p(x_{k+1} | y_{1:k}) \) when calculating \( \Gamma_{k+1} \) (see (16.128) for a further understanding). Hence, \( \Gamma_{k+1} \) provides more information on state \( x_{k+1} \) than \( \Gamma_{k+1} \), and the waveform selection criterion based on \( \tilde{\Gamma}_{k+1} \) has the potential to provide better processing performance. Note that \( \tilde{\Gamma}_{k+1} \) is calculated by averaging over all the
Then, we can rewrite $x$ state vector possible values of $y$ from the belief $p$ of the inverse of (16.123), replacing $\Gamma_k + 1$ by $\Gamma_k + 1$:

$$\theta^*_k = \arg \min_{\theta_k \in \Theta} \text{Tr}[\Pi_J^{-1}(\theta_k + 1)]$$  \hspace{1cm} (16.126)

where $\Theta$ denotes a set of the allowed values for $\theta_k + 1$ or a library of all possible waveforms, $\Pi$ is the weighting matrix used to equalize the magnitude of the different parameters in the state vector (see also Reference 43), and $J_k + 1$ is defined as in (16.123) replacing $\Gamma_k + 1$ by $\Gamma_k + 1$. 

(3) **Computation of the criterion function:** The proposed criterion function depends not only on the information provided by the state model $F$ but also on the measurement model and history, through the term $\Gamma_k$. To compute the former matrix, in general, the expectation in (16.125) has no closed-form analytical solution and must be solved numerically. We propose to use Monte Carlo integration to calculate this expectation and merge this numerical procedure into the sequential Monte Carlo method for tracking the target.

To compute the numerical result for $\Gamma_k + 1$, we define the matrix function

$$\Lambda(y_{k+1}, x_{k+1}) = -\Delta x_{k+1} \log p(y_{k+1} \mid x_{k+1})$$  \hspace{1cm} (16.127)

Then, we can rewrite $\Gamma_k + 1$ as

$$\Gamma_k + 1 = \int_{x_{k+1}} \left[ \int \Lambda(y_{k+1}, x_{k+1}) p(y_{k+1} \mid x_{k+1}) dy_{k+1} \right] \times p(x_{k+1} \mid y_{1:k}) dx_{k+1}$$  \hspace{1cm} (16.128)

According to this equation, the expectation to calculate $\Gamma_k + 1$ can first be taken with respect to the conditional density function $p(y_{k+1} \mid x_{k+1})$ and then with respect to the density $p(x_{k+1} \mid y_{1:k})$, i.e.

$$\Gamma_k + 1 = \mathbb{E}_{x_{k+1} \mid y_{1:k}} [\Sigma_k + 1]$$  \hspace{1cm} (16.129)

$$\Sigma_k + 1 = \mathbb{E}_{x_{k+1} \mid y_{1:k}} [-\Delta x_{k+1} \log p(y_{k+1} \mid x_{k+1})]$$  \hspace{1cm} (16.130)

Note that $\Sigma_k + 1$ is the standard Fisher information matrix (FIM) for estimating the state vector $x_{k+1}$ based on the observations $y_{k+1}$.

To calculate (16.129), we need samples of the predicted target state $x_{k+1}$. We can apply sequential Monte Carlo methods to draw these samples. For a sequential Monte Carlo method, we obtain $N_s$ samples and its associated weights at the $k$th time step from the belief $p(x_k \mid y_{1:k})$ as $\{x_k^{(i)}, w_k^{(i)}; i = 1, \ldots, N_s\}$. Then, the corresponding samples and weights of the predicted state are $\{x_{k+1}^{(i)}, w_{k+1}^{(i)}; i = 1, \ldots, N_s\}$, where
\[ x_{k+1} \sim p(x_{k+1} | x_k). \] Therefore, the expectation in (16.129) can be computed by the following two steps:

- For \( i = 1, \ldots, N_s \), draw samples \( x_{k+1}^{(i)} \sim p(x_{k+1} | x_k) \)
- Approximate the matrix \( \tilde{\Gamma}_{k+1} \) as

\[
\tilde{\Gamma}_{k+1} \approx \frac{1}{N_s} \sum_{i=1}^{N_s} w_k^{(i)} \Xi_{k+1}^{(i)}(x_{k+1}^{(i)})
\] (16.131)

To calculate (16.130), for each \( x_{k+1}^{(i)} \), we draw \( N_y \) identically independently distributed (i.i.d.) samples \( \{y_{k+1}^{(j)}; j = 1, \ldots, N_y\} \) from the likelihood function \( p(y_{k+1} | x_{k+1}^{(i)}) \). Then, we approximate the FIM \( \Xi_{k+1}(x_{k+1}^{(i)}) \) as

\[
\Xi_{k+1}^{(i)}(x_{k+1}^{(i)}) \approx \frac{1}{N_y} \sum_{j=1}^{N_y} \Lambda(y_{k+1}^{(j)}, x_{k+1}^{(i)})
\] (16.132)

Therefore, we approximate \( \tilde{\Gamma}_{k+1} \) using the Monte Carlo method as

\[
\tilde{\Gamma}_{k+1} \approx \frac{1}{N_s} \sum_{i=1}^{N_s} \sum_{j=1}^{N_y} w_k^{(i)} \Lambda(y_{k+1}^{(j)}, x_{k+1}^{(i)})
\] (16.133)

(4) Computation under Gaussian measurement noise: The Monte Carlo integration for computing \( \tilde{\Gamma}_{k+1} \) given by (16.133) is suitable for any statistical measurement model. However, if the additive noise \( e_k \) in the measurement model (16.104) has Gaussian distribution, we can obtain an analytical form for the FIM \( \Xi_{k+1} \); thus, the cost of computing \( \tilde{\Gamma}_{k+1} \) using (16.129) can be significantly reduced.

Assuming that the measurement noise \( e_{k+1} \) follows a complex Gaussian distribution, the measurement \( y_{k+1} \) given \( x_{k+1} \) is distributed as

\[
y_{k+1} | x_{k+1} \sim \mathcal{CN}(h(x_{k+1}), \Sigma_{k+1})
\] (16.134)

where \( h(\cdot) \) is defined in (16.104). We also assume that the measurement noise values \( \{e_{k+1}(t), t = t_1, \ldots, t_N\} \) are independent at different sample times. Then, the covariance matrix in (16.134) can be written as a block diagonal matrix:

\[
\Sigma_{k+1} = \text{diag} \{\Sigma_{k+1}(t_1), \ldots, \Sigma_{k+1}(t_N)\}
\] (16.135)

where, if the measurement noise \( e_{k+1} \) follows the model described in section 16.4.2,

\[
\Sigma_{k+1}(t) = A(t)\Sigma \Sigma^H(t - \tau_0)A^H(t - \tau_0)d + \sigma^2 I_M
\] (16.136)

Therefore, according to the results in Chapter 15.7 in Reference 20 the FIM in (16.130) is

\[
[\Xi_{k+1}(x_{k+1})]_{ij} = 2 \sum_{t=1}^{t_N} \text{Re} \left\{ \left[ \frac{\partial \hat{h}(t, x_{k+1})}{\partial x_{k+1,j}} \right]^H \Sigma_{k+1}^{-1}(t) \left[ \frac{\partial \hat{h}(t, x_{k+1})}{\partial x_{k+1,j}} \right] \right\}
\] (16.137)

where \( \hat{h}(\cdot) \) is defined in (16.103).
(5) **Suboptimal criterion function:** Computing $\hat{\Gamma}_{k+1}$ using the Monte Carlo integration is intensive and time demanding because the FIM $\Sigma_{k+1}$ must be evaluated for every particle. Therefore, we propose a suboptimal criterion function in which the matrix $\hat{\Gamma}_{k+1}$ is replaced by $\Sigma_{k+1}$ evaluated at the expected predicted state. Therefore, the suboptimal criterion can be computed by the following steps:

- For $i = 1, \ldots, N_s$, draw samples $x_{k+1}^{(i)} \sim p(x_{k+1} \mid x_{k}^{(i)})$.
- The expectation of the predicted state is approximated as

$$\hat{x}_{k+1} \approx \sum_{i=1}^{N_s} w_k^{(i)} x_{k+1}^{(i)}$$

(16.138)

- Replace $\hat{\Gamma}_{k+1}$ by $\Sigma_{k+1}(\hat{x}_{k+1})$ in (16.123).

This suboptimal criterion significantly reduces computation time at the expense of accuracy in computing the integral; hence, the selected waveform may not be optimal.

### 16.4.5 Numerical examples

We use numerical examples to study the performance of the proposed adaptive waveform design method for tracking targets in the presence of clutter. Through these examples, we demonstrate the advantages of the adaptive waveform design scheme compared with the fixed transmitted waveform scheme. First, we provide a description of the simulation set-up considered for the target and tracking system, and then we discuss different numerical examples. The results reported in this section correspond to the average over 100 Monte Carlo simulations.

**Target and clutter.** The numerical examples consist of a single target that moves parallel to the horizontal plane at a velocity of 200 m/s. The target trajectory is a section of a circle of radius 1.5 km, which starts at the position $r_0 = [10, 10]$ km, as shown in Figure 16.7. We assume that the scattering parameters of the target are partially known and have the following values: $m = 1$, $\epsilon = 15^\circ$, $\vartheta = 0^\circ$ and $\gamma = 20^\circ$; however, its orientation angle $\vartheta$ can change as the target moves. In addition, we consider that the clutter covariance parameters have been estimated using training data and that they have the following values: $\vartheta_c = 85^\circ$, $\epsilon_c = 5^\circ$, $\sigma_p^2 = 0.4$, $\sigma_u^2 = 0.4$ and $p_x = 0.2$. The covariance of the clutter will be scaled to fulfill the required target-to-clutter ratio (TCR). We define the TCR following the work by Novak et al. in Reference 8:

$$\text{TCR} = \frac{||[s_{hh}, s_{vv}, s_{hv}]||^2}{\text{tr}(\Sigma_c)}$$

(16.139)

where $s_{hh}, s_{vv}, s_{hv}$ are the target scattering coefficients defined in (16.88) and $|| \cdot ||$ is the norm of the vector.

We define SNR as

$$\text{SNR} = \frac{\int_{t_i}^{t_f} |p(t)|^2 \, dt}{\int_{t_i}^{t_f} E[|n(t)|^2] \, dt} = \frac{L}{\sigma^2 (t_f - t_i)}$$

(16.140)
where \( L \) is the number of transmitted pulses, \( \sigma^2 \) is the power of the thermal noise process and \( t_i, t_f \) define the time-window during which the system is allowed to track the target. \( p(t) \) is the pulse envelope of the transmitted signal. For the simulation examples, these parameters were set up in a way that the system was able to follow a target in a radial distance between 10 and 25 km.

**Transmitted signal.** We consider a radar system that transmits one pulse \((L = 1)\) at intervals of \( T_{PRI} = 250 \text{ ms} \), with a carrier frequency \( f_c = 15 \text{ GHz} \) \((\lambda = 20 \text{ mm})\). The maximum signal bandwidth is \( \text{BW}_{\text{max}} = 500 \text{ kHz} \). The system is capable of transmitting LFM pulses that change length \( \eta \), frequency rate \( b \) and polarization angles \( \alpha \) and \( \beta \) on a pulse-to-pulse basis.

**Tracking system.** The receiver of the tracker consists of two vector sensors \((M = 2)\) located at \( r_1 = [-0.25\lambda, 0] \) and \( r_2 = [0.25\lambda, 0] \). The radar echoes are recorded at sampling frequency \( f_s = 1 \text{ MHz} \). The system tracks the position and velocity of the target, as well as its orientation angle; hence, the state vector is \( \mathbf{x} = [x, \dot{x}, y, \dot{y}, \vartheta]^T \). The particle filter is implemented using the transitional prior \( p(\mathbf{x}_k | \mathbf{x}^{(i)}_{k-1}) \) as the importance density function to draw \( N_s = 500 \) particles. The intensity of the process noise is given by \( q_p = 500 \) and \( q_s = 50 \). In addition, we assume the covariance of the initial state is \( J_0^{-1} = \text{diag}[500, 500, 200, 200, 0.5] \). The weighting matrix \( \Pi \) is a diagonal matrix whose main diagonal entries are a power of 10 intended to equalize the covariance of the different parameters.

**Example 1.** In this example, we compare the performance of the adaptive and fixed waveform system assuming that the orientation angle of the target is \( \vartheta = 0^\circ \) along the entire trajectory. For the adaptive system, the wave shape parameters are \( \eta = 100 \mu \text{s} \).
and \( b_{\text{max}} = \frac{\text{BW}_{\text{max}}}{7.4\eta} \) (maximum allowable frequency rate for the signal bandwidth), and the polarization aspects of the signal are selected from the following waveform library:

\[
\Theta = \{\theta_m = (\alpha_l, \beta_n, \eta, b); l = 0, \ldots, 36; n = 0, \ldots, 6\} \quad (16.141)
\]

where

\[
\alpha_l = -90^\circ + l \cdot 5^\circ, \quad \beta_n = -45^\circ + n \cdot 15^\circ \quad (16.142)
\]

For the fixed waveform, the transmitted signal corresponds to the waveform \( \theta_{03} \) (vertical polarization). Figure 16.7 shows the averaged tracking results of the moving target in an environment such that TCR = 10 dB and SNR = 10 dB. For the fixed waveform, the vertical polarization is unfavourable because it is close to the polarimetric response of the clutter. Hence, the received signal is highly corrupted by clutter reflections and the tracking filter is not capable of following the target. On the other hand, the adaptive waveform method, although it was also started with vertical polarization, immediately selects the waveform that matches the target polarimetric aspects increasing the energy of the signal reflected from the target and reducing the clutter reflections. Therefore, the tracking performance for the adaptive waveform selection scheme is significantly better than the fixed waveform scheme.

Using the same simulation set-up, the numerical example was repeated. However, this time, the waveform was selected by applying the suboptimal criterion function in order to reduce the computation cost of the adaptive waveform design algorithm. Figure 16.8 shows the square root of the averaged mean-square error (MSE) for the target position. As expected, the suboptimal algorithm generated estimates with

![Figure 16.8 Square root of the averaged mean-square error (MSE) for the target position](image-url)
larger error. However, since the loss of performance is small and the reduction of computation time is significant, we will apply this suboptimal method in the following example.

**Example 2.** We analyse the behaviour of the tracking filter when the state model does not match the target dynamics. In this case, we consider the set-up and waveform library as in the first example; however, the orientation angle of the target $\vartheta$ changes following the linear piecewise function depicted in Figure 16.9. The same figure shows the estimated target orientation angle and the waveform polarization angle $\alpha$ selected for transmission by the adaptive algorithm. The same simulation was solved for two scenarios: $TCR = SNR = 10$ dB and $TCR = SNR = 15$ dB.

In Figure 16.9, it can be observed that the filter tries to track the true orientation angle when it is changing linearly, even though this parameter is defined as constant in the state model. Clearly, the convergence of the estimated orientation angle is faster when the clutter and noise interference is lower. We note that the filter selects the waveform that best matches the estimated target polarization aspects, in order to increase the energy reflected by the target.

**Figure 16.9** Averaged orientation angles for two scenarios: $TCR = SNR = 10$ dB (solid line) and $TCR = SNR = 15$ dB (dotted line)

### 16.5 Conclusions

We addressed the problem of designing the optimal signal polarization for target detection and tracking. We reviewed some of our recent results on the adaptive selection
of the radar transmit waveform polarizations, where the goal is to explore polarization diversity through an adaptive design. We showed that the optimal selection of the polarization significantly improves the detection and tracking performances of the radar systems when compared with fixed polarization schemes. We demonstrated that radar systems supporting agile polarization greatly outperform conventional sensing systems. However, to further improve the radar performance, several research challenges need to be considered. From the statistical signal processing perspective, these include the development of more realistic models for the target and clutter scattering, appropriate performance measures for various critical scenarios and robust but efficient optimization algorithms. Moreover, the problem of signal polarization design can be extended to other radar applications, such as sequential detection and target classification.

Acknowledgements

This work was supported by the Department of Defence under the Air Force Office of Scientific Research MURI Grant FA9550-05-1-0443, ONR Grant N000140810849, and AFOSR Grant FA9550-11-1-0210.

References

Waveform design and diversity for advanced radar systems


Adaptive polarization design for target detection and tracking


496  Waveform design and diversity for advanced radar systems


