Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

LEARNING OF RADAR SYSTEM FOR TARGET DETECTION

by
Wei Jiang

In this dissertation, the problem of data-driven joint design of transmitted waveform and detector in a radar system is addressed. Two novel learning-based approaches to waveform and detector design are proposed based on end-to-end training of the radar system. The first approach consists of alternating supervised training of the detector for a fixed waveform and reinforcement learning of the transmitter for a fixed detector. In the second approach, the transmitter and detector are trained simultaneously. Various operational waveform constraints, such as peak-to-average-power ratio (PAR) and spectral compatibility, are incorporated into the design. Unlike traditional radar design methods that rely on rigid mathematical models, it is shown that radar learning can be robustified to uncertainties about environment by training the detector with synthetic data generated from multiple statistical models of the environment. Theoretical considerations and results show that the proposed methods are capable of adapting the transmitted waveform to environmental conditions while satisfying design constraints.

LEARNING OF RADAR SYSTEM FOR TARGET DETECTION

by
Wei Jiang

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Helen and John C. Hartmann Department of
Electrical and Computer Engineering

August 2021

APPROVAL PAGE

LEARNING OF RADAR SYSTEM FOR TARGET DETECTION

Wei Jiang

Dr. Alexander M. Haimovich, Dissertation Advisor — Date
Distinguished Professor of Electrical and Computer Engineering, NJIT

Dr. Osvaldo Simeone, Committee Member — Date
Professor of Engineering, King's College London, U.K.

Dr. Ali Abdi, Committee Member — Date
Professor of Electrical and Computer Engineering, NJIT

Dr. Joerg Kliewer, Committee Member — Date
Professor of Electrical and Computer Engineering, NJIT

Dr. Zoi-Heleni Michalopoulou, Committee Member — Date
Professor of Mathematical Sciences, NJIT

BIOGRAPHICAL SKETCH

Author: Wei Jiang

Degree: Doctor of Philosophy

Date: August 2021

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2021
- Master of Science in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2015
- Bachelor of Science in Electrical Engineering,
  Harbin University of Science and Technology, Harbin, China, 2012

Major: Electrical Engineering

Presentations and Publications:

W. Jiang, A. M. Haimovich, and O. Simeone, "Joint design of radar waveform and detector via end-to-end learning with waveform constraints," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, 2021.

W. Jiang, A. M. Haimovich, M. Govoni, T. Garner, and O. Simeone, "Fast data-driven adaptation of radar detection via meta-learning," submitted to *IEEE 55th Asilomar Conference on Signals, Systems, and Computers*, 2021.

W. Jiang, A. M. Haimovich, and O. Simeone, "End-to-end learning of waveform generation and detection for radar systems," in *IEEE 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 1672-1676, 2019.

W. Jiang and A. M. Haimovich, "Waveform optimization in cloud radar with spectral congestion constraints," in *IEEE Radar Conference*, pp. 1-6, 2019.

W. Jiang and A. M. Haimovich, "Joint optimization of waveform and quantization in spectral congestion conditions," in *IEEE 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 1894-1898, 2018.

W. Jiang and A. M. Haimovich, "Cramer-Rao bound for non-coherent direction of arrival estimation in the presence of sensor location errors," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp 1303-1307, 2017.

W. Jiang, A. M. Haimovich, and Y. C. Eldar, "Direction of arrival by non-coherent arrays," in *IEEE Radar Conference*, pp. 1-6, 2016.

W. Jiang and A. M. Haimovich, "Cramer-Rao bound and approximate maximum likelihood estimation for non-coherent direction of arrival problem," in *IEEE Annual Conference on Information Science and Systems*, pp. 536-540, 2016.

Dedicated to my family.

ACKNOWLEDGMENT

First and foremost, I would like to express my deep sense of gratitude to my advisor, Dr. Alexander M. Haimovich, for providing me with the opportunity to join his lab, funding my research, and guiding me throughout my graduate studies. I have been fortunate to be supervised by him. His insight and immense knowledge steered me through this research. He has been supportive since the days I began working with him. During some difficult times in my life, he has given me the support and encouragement I needed to continue.

I would also like to show gratitude to the rest of my dissertation committee: Dr. Ali Abdi, Dr. Joerg Kliewer, Dr. Osvaldo Simeone, and Dr. Zoi-Heleni Michalopoulou for their helpful suggestions and insightful comments. Special thanks to Dr. Osvaldo Simeone for his invaluable advice, insights and feedback throughout our collaboration.

My thanks also go to my colleagues at CWiP lab for all the great moments we shared. In particular, I would like to thank Haley Kim, Nil Garcia, Alireza Bagheri, Annan Dong, Phuoc Vu, Sarah Obead, Malihe Aliasgari and many others for the many moments we shared talking, hiking, and hanging out. Special thanks to Haley Kim for his patient guidance through my transition from a master's student to a Ph.D. student.

I would like to thank Ms. Kathy Bosco, Ms. Monteria Bass, Ms. Clarisa Gonzalez-Lenahan, Ms. Angela Retino, Dr. Durgamadhab Misra, and Dr. Sotirios G. Ziavras for their advice, help and support with administrative matters during my Ph.D. studies.

Finally, I would like to express my profound gratitude to my parents, Zhaolin Jiang and Baihua Sang, who taught me the meaning of love, trust, and hope. I

sincerely thank my beloved husband, Chizhong Wang, for his constant encouragement along the way. Without their love and support, I could not have gone this far.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

Design of radar waveforms and detectors has been a topic of great interest to the radar community (see e.g., [1–4]). For best performance, radar waveforms and detectors should be designed jointly [5], [6]. Traditional joint design of waveforms and detectors typically relies on mathematical models of the environment, including targets, clutter, and noise. In contrast, this dissertation proposes data-driven approaches based on end-to-end learning of radar systems, in which reliance on rigid mathematical models of targets, clutter and noise is relaxed.

1.1 Optimal Radar Detector

Optimal detection in the Neyman-Pearon (NP) sense guarantees highest probability of detection for a specified probability of false alarm [1]. The NP detection test relies on the likelihood (or log-likelihood) ratio, which is the ratio of probability density functions (PDF) of the received signal conditioned on the presence or absence of a target. Mathematical tractability of models of the radar environment plays an important role in determining the ease of implementation of an optimal detector. For some target, clutter and noise models, the structure of optimal detectors is well known [7–9]. For example, closed-form expressions of the NP test metric are available when the applicable models are Gaussian [9], and, in some cases, even for non-Gaussian models [10].

However, in most cases involving non-Gaussian models, the structure of optimal detectors generally involves intractable numerical integrations, making the implementation of such detectors computationally intensive [11], [12]. For instance, it is shown in [11] that the NP detector requires a numerical integration with respect to the texture variable of the K-distributed clutter, thus precluding a closed-form

solution. Furthermore, detectors designed based on a specific mathematical model of environment suffer performance degradation when the actual environment differs from the assumed model [13], [14]. Attempts to robustify performance by designing optimal detectors based on mixtures of random variables quickly run aground due to mathematical intractability.

1.2 Radar Waveform Design

Alongside optimal detectors, optimal radar waveforms may also be designed based on the NP criterion. Solutions are known for some simple target, clutter and noise models (see e.g., [2], [4]). However, in most cases, waveform design based on direct application of the NP criterion is intractable, leading to various suboptimal approaches. For example, mutual information, J-divergence and Bhattacharyya distance have been studied as objective functions for waveform design in multistatic settings [15–18].

In addition to target, clutter and noise models, waveform design may have to account for various operational constraints. For example, transmitter efficiency may be improved by constraining the peak-to-average-power ratio (PAR) [19–22]. A different constraint relates to the requirement of coexistence of radar and communication systems in overlapping spectral regions. The National Telecommunications and Information Administration (NTIA) and Federal Communication Commission (FCC) have allowed sharing of some of the radar frequency bands with commercial communication systems [23]. In order to protect the communication systems from radar interference, radar waveforms should be designed subject to specified compatibility constraints. The design of radar waveforms constrained to share the spectrum with communications systems has recently developed into an active area of research with a growing body of literature [24–28].

1.3 Machine Learning Applied in Radar

Machine learning has been successfully applied to solve problems for which mathematical models are unavailable or too complex to yield optimal solutions, in domains such as computer vision [29], [30] and natural language processing [31], [32]. Recently, a machine learning approach has been proposed for implementing the physical layer of communication systems. Notably, in [33], it is proposed to jointly design the transmitter and receiver of communication systems via end-to-end learning. Reference [34] proposes an end-to-end learning-based approach for jointly minimizing PAR and bit error rate in orthogonal frequency division multiplexing systems. This approach requires the availability of a known channel model. For the case of an unknown channel model, reference [35] proposes an alternating training approach, whereby the transmitter is trained via reinforcement learning (RL) on the basis of noiseless feedback from the receiver, while the receiver is trained by supervised learning. In [36], the authors apply simultaneous perturbation stochastic optimization for approximating the gradient of a transmitter's loss function. A detailed review of the state of the art can be found in [37] (see also [38–40] for recent work).

In the radar field, learning machines trained in a supervised manner based on a suitable loss function have been shown to approximate the performance of the NP detector [41], [42]. As a representative example, in [42], a neural network trained in a supervised manner using data that includes Gaussian interference, has been shown to approximate the performance of the NP detector. Note that design of the NP detector requires express knowledge of the Gaussian nature of the interference, while the neural network is trained with data that happens to be Gaussian, but the machine has no prior knowledge of the statistical nature of the data.

1.4 Organization and Contributions

In this dissertation, we introduce two learning-based approaches for the joint design of waveform and detector in a radar system. Inspired by [35], end-to-end learning of a radar system is implemented by alternating supervised learning of the detector for a fixed waveform, and RL-based learning of the transmitter for a fixed detector. In the second approach, the learning of the detector and waveform are executed simultaneously, potentially speeding up training in terms of required radar transmissions to yield the training samples compared alternating training. In addition, we extend the problem formulation to include training of waveforms with PAR or spectral compatibility constraints.

The organization and main contributions of this dissertation are outlined.

Chapter 2: In this chapter, we formulate a radar system architecture based on the training of the detector and the transmitted waveform, both implemented as feedforward multi-layer neural networks.

Chapter 3: This chapter presents two end-to-end learning algorithms for detection and waveform generation. In the first learning algorithm, detector and transmitted waveform are trained alternately: For a fixed waveform, the detector is trained using supervised learning so as to approximate the NP detector; and for a fixed detector, the transmitted waveform is trained via policy gradient-based RL. In the second algorithm, the detector and transmitter are trained simultaneously.

Chapter 4: In this chapter, we extend learning algorithms to incorporate waveform constraints, specifically PAR and spectral compatibility constraints.

Chapter 5: This chapter provides theoretical results that relate alternating and simultaneous training by computing the gradients of the loss functions optimized by both methods. This chapter also provides theoretical results that justify the use of RL-based transmitter training by comparing the gradient used by this procedure with the gradient of the ideal model-based likelihood function.

Chapter 6: This chapter presents numerical results that the proposed end-to-end learning approaches are able to obtain a more robust radar performance in clutter and colored noise of arbitrary probability density functions as compared to conventional methods. This chapter also shows that the proposed methods are capable of adapting the transmitted waveform to environmental conditions while satisfying design constraints.

Throughout the dissertation bold lowercase and uppercase letters represent vector and matrix, respectively. The conjugate, the transpose, and the conjugate transpose operator are denoted by the symbols $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, respectively. The notations $\mathbb{C}^K$ and $\mathbb{R}^K$ represent sets of K-dimensional vectors of complex and real numbers, respectively. The notation $|\cdot|$ indicates modulus, $||\cdot||$ indicates the Euclidean norm, and $\mathbb{E}_{x\sim p_x}\{\cdot\}$ indicates the expectation of the argument with respect to the distribution of the random variable $x \sim p_x$, respectively. $\Re(\cdot)$ and $\Im(\cdot)$ stand for real-part and imaginary-part of the complex-valued argument, respectively. The letter j represents the imaginary unit, i.e., $j = \sqrt{-1}$. The gradient of a function f: $\mathbb{R}^n \to \mathbb{R}^m$ with respect to $\mathbf{x} \in \mathbb{R}^n$ is $\nabla_{\mathbf{x}} f(\mathbf{x}) \in \mathbb{R}^{n\times m}$.

CHAPTER 2

PROBLEM FORMULATION

Consider a pulse-compression radar system that uses the baseband transmit signal

$$x(t) = \sum_{k=1}^{K} y_k \zeta\big(t - [k-1]T_c\big), \tag{2.1}$$

where $\{y_k\}_{k=1}^{K}$ are complex deterministic coefficients, and $\zeta(t)$ is, for example, a square root Nyquist subpulse (referred to as "chip") with chip rate $1/T_c$, such that $\big\{\zeta\big(t - [k-1]T_c\big)\big\}_{k=1}^{K}$ are orthogonal. The vector $\mathbf{y} \triangleq [y_1, \ldots, y_K]^T$ is referred to as *coded waveform*, and is a subject of design.

The backscattered baseband signal from a stationary point-like target is given by

$$z(t) = \alpha x(t - \tau) + c(t) + n(t), \tag{2.2}$$

where α is the target complex-valued gain, accounting for target backscattering and channel propagation effects; τ represents the target delay, which is assumed to satisfy the target detectability condition $\tau \gg KT_c$; $c(t)$ is the clutter component; and $n(t)$ denotes signal-independent noise comprising an aggregate of thermal noise, interference, and jamming. The clutter component $c(t)$ associated with a detection test performed at $\tau = 0$ may be expressed

$$c(t) = \sum_{g=-K+1}^{K-1} \gamma_g x\big(t - gT_c\big), \tag{2.3}$$

where γ_g is the complex clutter scattering coefficient at time delay $\tau = 0$ associated with the gth range cell relative to the cell under test. Following chip matched filtering with $\zeta^*(-t)$, and sampling at T_c-spaced time instants $t = \tau + [k-1]T_c$ for $k \in \{1, \ldots, K\}$, the $K \times 1$ discrete-time received signal $\mathbf{z} = [z(\tau), z(\tau + T_c), \ldots, z(\tau + [K-1]T_c)]^T$ for the range cell under test containing a point target with complex amplitude α, clutter and noise can be written as

$$\mathbf{z} = \alpha\mathbf{y} + \mathbf{c} + \mathbf{n}, \tag{2.4}$$

where $\mathbf{c}$ and $\mathbf{n}$ denote, respectively, the clutter and noise vectors.

Detection of the presence of a target in the range cell under test is formulated as the following binary hypothesis testing problem:

$$\begin{cases} \mathcal{H}_0 : \mathbf{z} = \mathbf{c} + \mathbf{n} \\ \mathcal{H}_1 : \mathbf{z} = \alpha\mathbf{y} + \mathbf{c} + \mathbf{n}. \end{cases} \tag{2.5}$$

In traditional radar design, the golden standard for detection is provided by the NP criterion of maximizing the probability of detection for a given probability of false alarm. Application of the NP criterion leads to the likelihood ratio test

$$\Lambda(\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{y}, \mathcal{H}_1)}{p(\mathbf{z}|\mathbf{y}, \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T_\Lambda, \tag{2.6}$$

where $\Lambda(\mathbf{z})$ is the likelihood ratio, and T_Λ is the detection threshold set based on the probability of false alarm constraint [5]. The NP criterion is also the golden standard

for designing a radar waveform that adapts to the given environment, although, as discussed earlier, a direct application of this design principle is often intractable.

The design of optimal detectors and/or waveforms under the NP criterion relies on channel models of the radar environment, namely, knowledge of the conditional probabilities $p(\mathbf{z}|\mathbf{y}, \mathcal{H}_i)$ for $i = \{0, 1\}$. The channel model $p(\mathbf{z}|\mathbf{y}, \mathcal{H}_i)$ is the likelihood of the observation $\mathbf{z}$ conditioned on the transmitted waveform $\mathbf{y}$ and hypothesis $\mathcal{H}_i$. In the following, we introduce an end-to-end radar system in which the detector and waveform are jointly learned in a data-driven fashion.

2.1 End-to-end Radar System

The end-to-end radar system illustrated in Figure 2.1 comprises a transmitter and a receiver that seek to detect the presence of a target. Transmitter and receiver are implemented as two separate parametric functions $f_{\boldsymbol{\theta}_T}(\cdot)$ and $f_{\boldsymbol{\theta}_R}(\cdot)$ with trainable parameter vectors $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_R$, respectively.

As shown in Figure 2.1, the input to the transmitter is a user-defined initialization waveform $\mathbf{s} \in \mathbb{C}^K$. The transmitter outputs a radar waveform obtained through a trainable mapping $\mathbf{y}_{\boldsymbol{\theta}_T} = f_{\boldsymbol{\theta}_T}(\mathbf{s}) \in \mathbb{C}^K$. The environment is modeled as a stochastic system that produces the vector $\mathbf{z} \in \mathbb{C}^K$ from a conditional PDF $p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$ parameterized by a binary variable $i \in \{0, 1\}$. The absence or presence of a target is indicated by the values $i = 0$ and $i = 1$ respectively, and hence i is referred to as the *target state indicator*. The receiver passes the received vector $\mathbf{z}$ through a trainable mapping $p = f_{\boldsymbol{\theta}_R}(\mathbf{z})$, which produces the scalar $p \in (0, 1)$. The final decision $\hat{i} \in \{0, 1\}$ is made by comparing the output of the receiver p to a hard threshold in the interval $(0, 1)$.

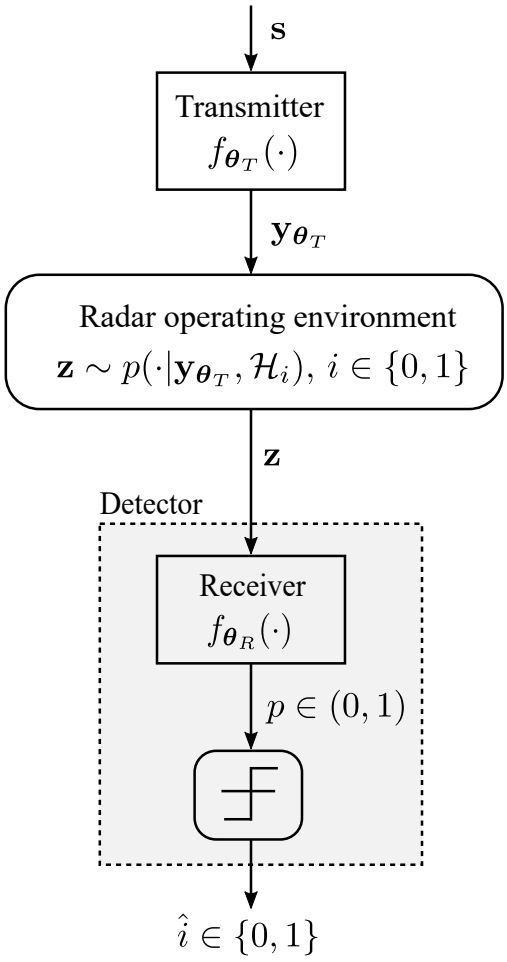

Figure 2.1 An end-to-end radar system operating over an unknown radar operating environment. Transmitter and receiver are implemented as two separate parametric functions $f_{\boldsymbol{\theta}_T}(\cdot)$ and $f_{\boldsymbol{\theta}_R}(\cdot)$ with trainable parameter vectors $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_R$, respectively.

2.2 Transmitter and Receiver Architectures

As discussed in Section 2.1, the transmitter and the receiver are implemented as two separate parametric functions $f_{\boldsymbol{\theta}_T}(\cdot)$ and $f_{\boldsymbol{\theta}_R}(\cdot)$. We now detail an implementation of the transmitter $f_{\boldsymbol{\theta}_T}(\cdot)$ and receiver $f_{\boldsymbol{\theta}_R}(\cdot)$ based on feedforward neural networks.

A feedforward neural network is a parametric function $\tilde{f}_{\boldsymbol{\theta}}(\cdot)$ that maps an input real-valued vector $\mathbf{u}_{\text{in}} \in \mathbb{R}^{N_{\text{in}}}$ to an output real-valued vector $\mathbf{u}_{\text{out}} \in \mathbb{R}^{N_{\text{out}}}$ via L successive layers, where N_{in} and N_{out} represent, respectively, the number of neurons of the input and output layers. Noting that the input to the lth layer is the output of the $(l-1)$th layer, the output of the lth layer is given by

$$\mathbf{u}_l = \tilde{f}_{\boldsymbol{\theta}^{[l]}}(\mathbf{u}_{l-1}) = \phi\big(\mathbf{W}^{[l]}\mathbf{u}_{l-1} + \mathbf{b}^{[l]}\big), \text{ for } l = 1, \ldots, L, \tag{2.7}$$

where $\phi(\cdot)$ is an element-wise activation function, and $\boldsymbol{\theta}^{[l]} = \{\mathbf{W}^{[l]}, \mathbf{b}^{[l]}\}$ contains the trainable parameter of the lth layer comprising the weight $\mathbf{W}^{[l]}$ and bias $\mathbf{b}^{[l]}$. Let N_{l-1} and N_l represent the number of neurons at the $(l-1)$th layer and the lth layer, respectively. The weight $\mathbf{W}^{[l]}$ at the lth layer is an $N_l \times N_{l-1}$ matrix, and the bias $\mathbf{b}^{[l]}$ at the lth layer is an $N_l \times 1$ vector. The vector of trainable parameters of the entire neural network comprises the parameters of all layers, i.e., $\boldsymbol{\theta} = \text{vec}\{\boldsymbol{\theta}^{[1]}, \cdots, \boldsymbol{\theta}^{[L]}\}$.

The architecture of the end-to-end radar system with transmitter and receiver implemented based on feedforward neural networks is shown in Figure 2.2. The transmitter applies a complex initialization waveform $\mathbf{s}$ to the function $f_{\boldsymbol{\theta}_T}(\cdot)$. The complex-value input $\mathbf{s}$ is processed by a complex-to-real conversion layer. This is followed by a real-valued neural network $\tilde{f}_{\boldsymbol{\theta}_T}(\cdot)$. The output of the neural network

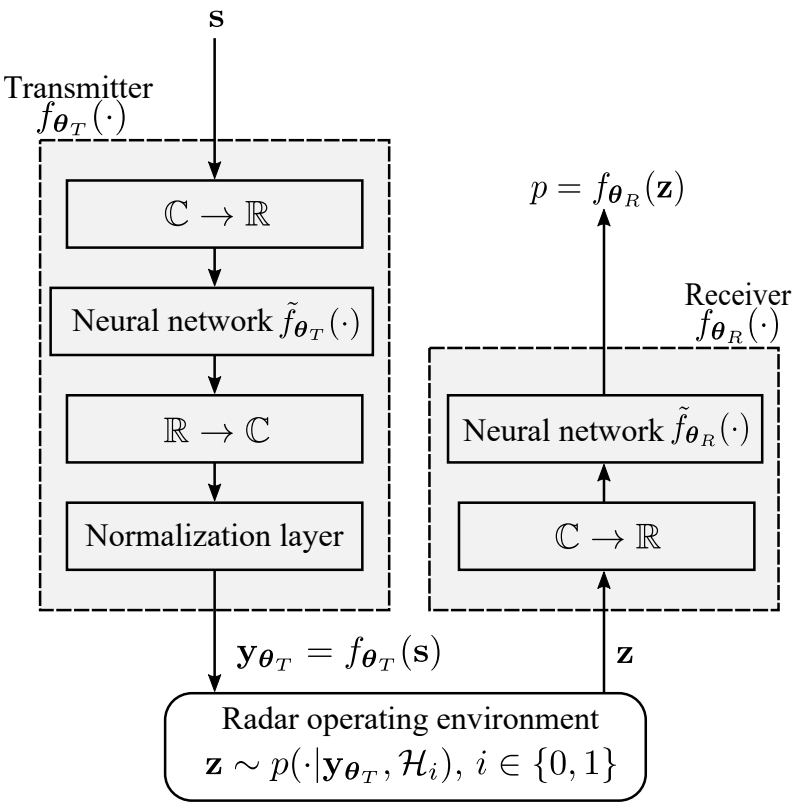

Figure 2.2 Transmitter and receiver architectures based on feedforward neural networks.

is converted back to complex-values, and an output layer normalizes the transmitted power. As a result, the transmitter generates the radar waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$.

The receiver applies the received signal $\mathbf{z}$ to the function $f_{\boldsymbol{\theta}_R}(\cdot)$. Similar to the transmitter, a first layer converts complex-valued to real-valued vectors. The neural network at the receiver is denoted $\tilde{f}_{\boldsymbol{\theta}_R}(\cdot)$. The task of the receiver is to generate a scalar $p \in (0, 1)$ that approximates the posterior probability of the presence of a target conditioned on the received vector $\mathbf{z}$. To this end, the last layer of the neural network $\tilde{f}_{\boldsymbol{\theta}_R}(\cdot)$ is selected as a logistic regression layer consisting of operating over a linear combination of outputs from the previous layer. The presence or absence of the target is determined based on the output of the receiver and a threshold set according to a false alarm constraint.

CHAPTER 3

TRAINING OF END-TO-END RADAR SYSTEMS

This chapter discusses the joint optimization of the trainable parameter vectors $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_R$ to meet application-specific performance requirements. Two training algorithms are proposed to train the end-to-end radar system. The first algorithm alternates between training of the receiver and of the transmitter. This algorithm is referred to as *alternating training*, and is inspired by the approach used in [35] to train encoder and decoder of a digital communication system. In contrast, the second algorithm trains the receiver and transmitter simultaneously. This approach is referred to as *simultaneous training*. Note that the proposed two training algorithms are applicable to other differentiable parametric functions implementing the transmitter $f_{\boldsymbol{\theta}_T}(\cdot)$ and the receiver $f_{\boldsymbol{\theta}_R}(\cdot)$, such as recurrent neural network or its variants [43]. In the following, we first discuss alternating training and then we detail simultaneous training.

3.1 Alternating Training: Receiver Design

Alternating training consists of iterations encompassing separate receiver and transmitter updates. In this section, we focus on the receiver updates. A receiver training update optimizes the receiver parameter vector $\boldsymbol{\theta}_R$ for a fixed transmitter waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$. Receiver design is supervised in the sense that we assume the target state indicator i to be available to the receiver during training. Supervised training of the receiver for a fixed transmitter's parameter vector $\boldsymbol{\theta}_T$ is illustrated in Figure 3.1.

The standard cross-entropy loss [42] is adopted as the loss function for the receiver. For a given transmitted waveform $\mathbf{y}_{\boldsymbol{\theta}_T} = f_{\boldsymbol{\theta}_T}(\mathbf{s})$, the receiver average loss

function is accordingly given by

$$\mathcal{L}_R(\boldsymbol{\theta}_R) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\big\}, \tag{3.1}$$

where $P(\mathcal{H}_i)$ is the prior probability of the target state indicator i, and $\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)$ is the instantaneous cross-entropy loss for a pair $\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)$, namely,

$$\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big) = -i\ln f_{\boldsymbol{\theta}_R}(\mathbf{z}) - (1-i)\ln\big[1-f_{\boldsymbol{\theta}_R}(\mathbf{z})\big]. \tag{3.2}$$

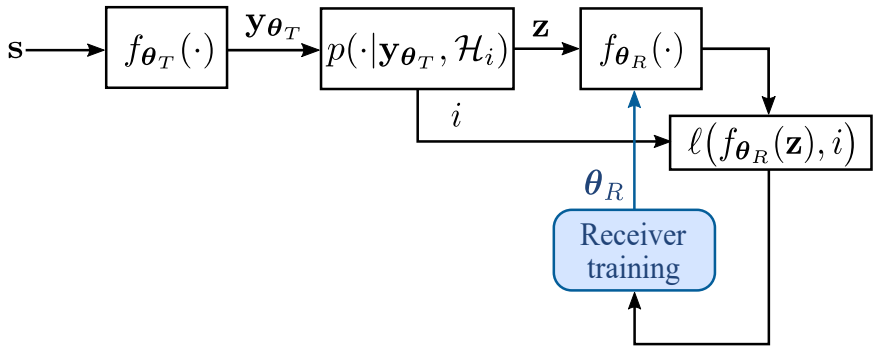

Figure 3.1 Supervised training of the receiver for a fixed transmitted waveform.

For a fixed transmitted waveform, the receiver parameter vector $\boldsymbol{\theta}_R$ should be ideally optimized by minimizing average loss (3.1), e.g., via gradient descent or one of its variants [44]. The gradient of average loss (3.1) with respect to the receiver parameter vector $\boldsymbol{\theta}_R$ is

$$\nabla_{\boldsymbol{\theta}_R}\mathcal{L}_R(\boldsymbol{\theta}_R) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)}\big\{\nabla_{\boldsymbol{\theta}_R}\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\big\}. \tag{3.3}$$

This being a data-driven approach, rather than assuming known prior probability of the target state indicator $P(\mathcal{H}_i)$ and likelihood $p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$, the receiver is assumed to have access to Q_R independent and identically distributed (i.i.d.) samples $\mathcal{D}_R = \left\{\mathbf{z}^{(q)} \sim p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_{i^{(q)}}), i^{(q)} \in \{0,1\}\right\}_{q=1}^{Q_R}$.

Given the output of the receiver function $f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)})$ for a received sample vector $\mathbf{z}^{(q)}$ and the indicator $i^{(q)} \in \{0,1\}$, the instantaneous cross-entropy loss is computed from Equation (3.2), and the estimated receiver gradient is given by

$$\nabla_{\boldsymbol{\theta}_R} \widehat{\mathcal{L}}_R(\boldsymbol{\theta}_R) = \frac{1}{Q_R} \sum_{q=1}^{Q_R} \nabla_{\boldsymbol{\theta}_R} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)}), i^{(q)}\big). \tag{3.4}$$

Using (3.4), the receiver parameter vector $\boldsymbol{\theta}_R$ is adjusted according to stochastic gradient descent updates

$$\boldsymbol{\theta}_R^{(n+1)} = \boldsymbol{\theta}_R^{(n)} - \epsilon \nabla_{\boldsymbol{\theta}_R} \widehat{\mathcal{L}}_R(\boldsymbol{\theta}_R^{(n)}) \tag{3.5}$$

across iterations $n = 1, 2, \ldots$, where $\epsilon > 0$ is the learning rate.

3.2 Alternating Training: Transmitter Design

In the transmitter training phase of alternating training, the receiver parameter vector $\boldsymbol{\theta}_R$ is held constant, and the function $f_{\boldsymbol{\theta}_T}(\cdot)$ implementing the transmitter is optimized. The goal of transmitter training is to find an optimized parameter vector $\boldsymbol{\theta}_T$ that minimizes the cross-entropy loss function (3.1) seen as a function of $\boldsymbol{\theta}_T$.

As illustrated in Figure 3.2, a stochastic transmitter outputs a waveform $\mathbf{a}$ drawn from a distribution $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ conditioned on $\mathbf{y}_{\boldsymbol{\theta}_T} = f_{\boldsymbol{\theta}_T}(\mathbf{s})$. The introduction of the randomization $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ of the designed waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$ is useful to enable

exploration of the design space in a manner akin to standard RL policies. To train the transmitter, we aim to minimize the average cross-entropy loss

$$\mathcal{L}_T^{\pi}(\boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\big\}. \tag{3.6}$$

Note that this is consistent with Equation (3.1), with the caveat that an expectation is taken over policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$. This is indicated by the superscript "π".

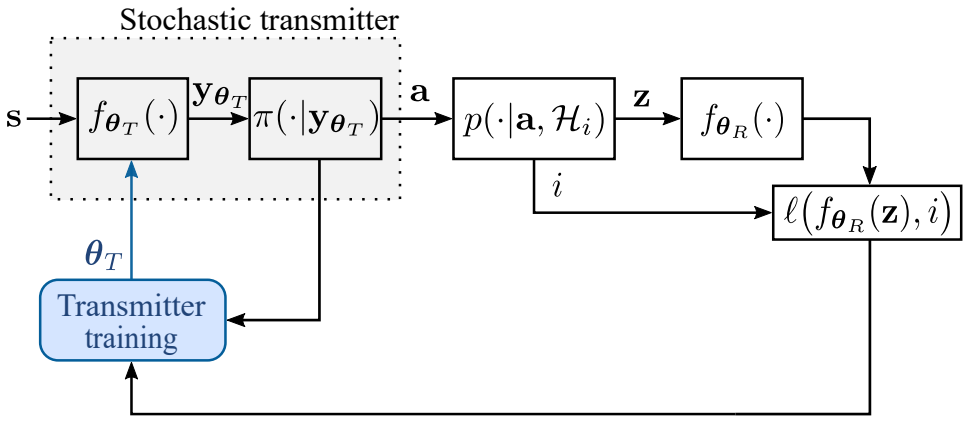

Figure 3.2 RL-based transmitter training for a fixed receiver design.

Assume that the policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ is differentiable with respect to the transmitter parameter vector $\boldsymbol{\theta}_T$, i.e., that the gradient $\nabla_{\boldsymbol{\theta}_T}\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ exists. The policy gradient theorem [45] states that the gradient of the average loss (3.6) can be written as

$$\nabla_{\boldsymbol{\theta}_T}\mathcal{L}_T^{\pi}(\boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\nabla_{\boldsymbol{\theta}_T}\ln\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\big\}. \tag{3.7}$$

The gradient (3.7) has the important advantage that it may be estimated via Q_T i.i.d. samples $\mathcal{D}_T = \big\{\mathbf{a}^{(q)}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}), \mathbf{z}^{(q)}\sim p(\mathbf{z}|\mathbf{a}^{(q)},\mathcal{H}_{i^{(q)}}), i^{(q)}\in\{0,1\}\big\}_{q=1}^{Q_T}$, yielding the estimate

$$\nabla_{\boldsymbol{\theta}_T} \widehat{\mathcal{L}}_T^{\pi}(\boldsymbol{\theta}_T) = \frac{1}{Q_T} \sum_{q=1}^{Q_T} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)}), i^{(q)}\big) \nabla_{\boldsymbol{\theta}_T} \ln \pi(\mathbf{a}^{(q)} | \mathbf{y}_{\boldsymbol{\theta}_T}). \tag{3.8}$$

With estimate (3.8), in a manner similar to Equation (3.5), the transmitter parameter vector $\boldsymbol{\theta}_T$ may be optimized iteratively according to the stochastic gradient descent update rule

$$\boldsymbol{\theta}_T^{(n+1)} = \boldsymbol{\theta}_T^{(n)} - \epsilon \nabla_{\boldsymbol{\theta}_T} \widehat{\mathcal{L}}_T^{\pi}(\boldsymbol{\theta}_T^{(n)}) \tag{3.9}$$

over iterations $n = 1, 2, \ldots$. The alternating training algorithm is summarized in Algorithm 1. The training process is carried out until a stopping criterion is satisfied. For example, a prescribed number of iterations may have been reached, or a number of iterations may have elapsed during which the training loss (3.6), estimated using samples $\mathcal{D}_T$, may have not decreased by more than a given amount.

Algorithm 1: Alternating Training

Input: initialization waveform $\mathbf{s}$; stochastic policy $\pi_{\boldsymbol{\theta}_T}(\cdot|\mathbf{y})$; learning rate ϵ

Output: learned parameter vectors $\boldsymbol{\theta}_R$ and $\boldsymbol{\theta}_T$

initialize $\boldsymbol{\theta}_R^{(0)}$ and $\boldsymbol{\theta}_T^{(0)}$, and set $n = 0$

while *stopping criterion not satisfied* **do**

`/* receiver training phase */`

evaluate the receiver loss gradient $\nabla_{\boldsymbol{\theta}_R}\widehat{\mathcal{L}}_R(\boldsymbol{\theta}_R^{(n)})$ from (3.4) with $\boldsymbol{\theta}_T = \boldsymbol{\theta}_T^{(n)}$

update receiver parameter vector $\boldsymbol{\theta}_R$ via

$$\boldsymbol{\theta}_R^{(n+1)} = \boldsymbol{\theta}_R^{(n)} - \epsilon\nabla_{\boldsymbol{\theta}_R}\widehat{\mathcal{L}}_R(\boldsymbol{\theta}_R^{(n)})$$

and stochastic transmitter policy turned off

`/* transmitter training phase */`

evaluate the transmitter loss gradient $\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_T^{\pi}(\boldsymbol{\theta}_T^{(n)})$ from (3.8) with $\boldsymbol{\theta}_R = \boldsymbol{\theta}_R^{(n+1)}$

update transmitter parameter vector $\boldsymbol{\theta}_T$ via

$$\boldsymbol{\theta}_T^{(n+1)} = \boldsymbol{\theta}_T^{(n)} - \epsilon\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_T^{\pi}(\boldsymbol{\theta}_T^{(n)})$$

$n \leftarrow n + 1$

end

3.3 Simultaneous Training

This section discusses simultaneous training, in which the receiver and transmitter are updated simultaneously as illustrated in Figure 3.3. To this end, the objective function is the average loss

$$\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) \\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}} \big\{ \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) \big\}. \tag{3.10}$$

This function is minimized over both parameters $\boldsymbol{\theta}_R$ and $\boldsymbol{\theta}_T$ via stochastic gradient descent.

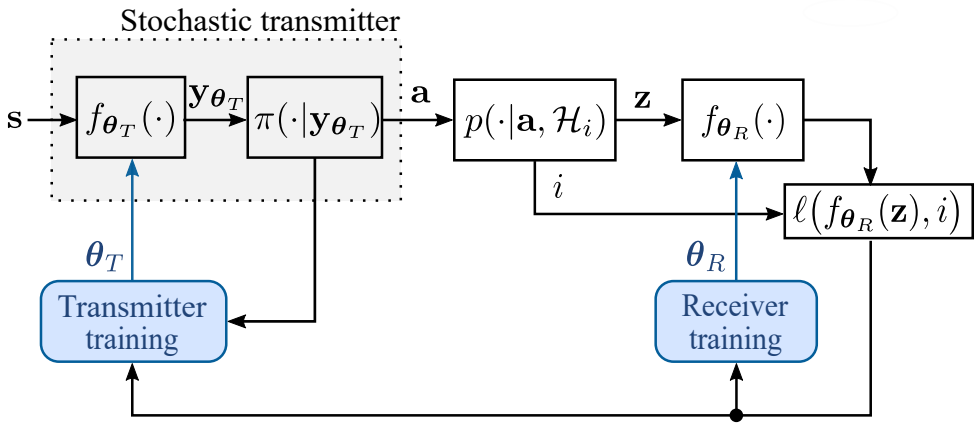

Figure 3.3 Simultaneous training of the end-to-end radar system. The receiver is trained by supervised learning, while the transmitter is trained by RL.

The gradient of (3.10) with respect to $\boldsymbol{\theta}_R$ is

$$\nabla_{\boldsymbol{\theta}_R}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) \\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}} \big\{ \nabla_{\boldsymbol{\theta}_R}\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) \big\}, \tag{3.11}$$

and the gradient of Equation (3.10) with respect to $\boldsymbol{\theta}_T$ is

$$\begin{aligned}\nabla_{\boldsymbol{\theta}_T}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T) &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\nabla_{\boldsymbol{\theta}_T}\mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\big\} \\ &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\nabla_{\boldsymbol{\theta}_T}\ln\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})\big\}.\end{aligned} \tag{3.12}$$

To estimate gradients (3.11) and (3.12), we assume access to Q i.i.d. samples $\mathcal{D} = \big\{\mathbf{a}^{(q)} \sim \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}), \mathbf{z}^{(q)} \sim p(\mathbf{z}|\mathbf{a}^{(q)},\mathcal{H}_{i^{(q)}}), i^{(q)} \in \{0,1\}\big\}_{q=1}^{Q}$. From (3.11), the estimated receiver gradient is

$$\nabla_{\boldsymbol{\theta}_R}\widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T) = \frac{1}{Q}\sum_{q=1}^{Q}\nabla_{\boldsymbol{\theta}_R}\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)}),i^{(q)}\big). \tag{3.13}$$

Note that, in (3.13), the received vector $\mathbf{z}^{(q)}$ is obtained based on a given waveform $\mathbf{a}^{(q)}$ sampled from policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$. Thus, the estimated receiver gradient (3.13) is averaged over the stochastic waveforms $\mathbf{a}$. This is in contrast to alternating training, in which the receiver gradient depends directly on the transmitted waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$.

From (3.12), the estimated transmitter gradient is given by

$$\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T) = \frac{1}{Q}\sum_{q=1}^{Q}\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)}),i^{(q)}\big)\nabla_{\boldsymbol{\theta}_T}\ln\pi(\mathbf{a}^{(q)}|\mathbf{y}_{\boldsymbol{\theta}_T}). \tag{3.14}$$

Finally, denote the parameter set $\boldsymbol{\theta} = \{\boldsymbol{\theta}_R,\boldsymbol{\theta}_T\}$, from (3.13) and (3.14), the trainable parameter set $\boldsymbol{\theta}$ is updated according to the stochastic gradient descent rule

$$\boldsymbol{\theta}^{(n+1)} = \boldsymbol{\theta}^{(n)} - \epsilon\nabla_{\boldsymbol{\theta}}\widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R^{(n)},\boldsymbol{\theta}_T^{(n)}) \tag{3.15}$$

across iterations $n = 1, 2, \ldots$. The simultaneous training algorithm is summarized in Algorithm 2.

Algorithm 2: Simultaneous Training

Input: initialization waveform $\mathbf{s}$; stochastic policy $\pi(\cdot|\mathbf{y}_{\boldsymbol{\theta}_T})$; learning rate ϵ

Output: learned parameter vectors $\boldsymbol{\theta}_R$ and $\boldsymbol{\theta}_T$

initialize $\boldsymbol{\theta}_R^{(0)}$ and $\boldsymbol{\theta}_T^{(0)}$, and set $n = 0$

while *stopping criterion not satisfied* **do**

evaluate the receiver gradient $\nabla_{\boldsymbol{\theta}_R} \widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R^{(n)}, \boldsymbol{\theta}_T^{(n)})$ and the transmitter gradient $\nabla_{\boldsymbol{\theta}_T} \widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R^{(n)}, \boldsymbol{\theta}_T^{(n)})$ from (3.13) and (3.14), respectively

update receiver parameter vector $\boldsymbol{\theta}_R$ and transmitter parameter vector $\boldsymbol{\theta}_T$ simultaneously via

$$\boldsymbol{\theta}_R^{(n+1)} = \boldsymbol{\theta}_R^{(n)} - \epsilon \nabla_{\boldsymbol{\theta}_R} \widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R^{(n)}, \boldsymbol{\theta}_T^{(n)})$$

and

$$\boldsymbol{\theta}_T^{(n+1)} = \boldsymbol{\theta}_T^{(n)} - \epsilon \nabla_{\boldsymbol{\theta}_T} \widehat{\mathcal{L}}^{\pi}(\boldsymbol{\theta}_R^{(n)}, \boldsymbol{\theta}_T^{(n)})$$

$n \leftarrow n + 1$

end

CHAPTER 4

TRANSMITTER DESIGN WITH CONSTRAINTS

We extend the transmitter training discussed in the previous chapter to incorporate waveform constraints on PAR and spectral compatibility. To this end, we introduce penalty functions that are used to modify the training criterion (3.6) to meet these constraints.

4.1 PAR Constraint

Low PAR waveforms are preferred in radar systems due to hardware limitations related to waveform generation. A lower PAR entails a lower dynamic range of the power amplifier, which in turn allows an increase in average transmitted power. The PAR of a radar waveform $\mathbf{y}_{\boldsymbol{\theta}_T} = f_{\boldsymbol{\theta}_T}(\mathbf{s})$ may be expressed

$$J_{\mathrm{PAR}}(\boldsymbol{\theta}_T) = \frac{\max\limits_{k=1,\cdots,K} |y_{\boldsymbol{\theta}_T,k}|^2}{||\mathbf{y}_{\boldsymbol{\theta}_T}||^2/K}, \tag{4.1}$$

which is bounded according to $1 \leq J_{\mathrm{PAR}}(\boldsymbol{\theta}_T) \leq K$.

4.2 Spectral Compatibility Constraint

A spectral constraint is imposed when a radar system is required to operate over a spectrum partially shared with other systems such as wireless communication networks. Suppose there are D frequency bands $\{\Gamma_d\}_{d=1}^{D}$ shared by the radar and by the coexisting systems, where $\Gamma_d = [f_{d,l}, f_{d,u}]$, with $f_{d,l}$ and $f_{d,u}$ denoting the lower and upper normalized frequencies of the dth band, respectively. The amount

of interfering energy generated by the radar waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$ in the dth shared band is

$$\int_{f_{d,l}}^{f_{d,u}} \left| \sum_{k=0}^{K-1} y_{\boldsymbol{\theta}_T,k} e^{-j2\pi fk} \right|^2 df = \mathbf{y}_{\boldsymbol{\theta}_T}^H \boldsymbol{\Omega}_d \mathbf{y}_{\boldsymbol{\theta}_T}, \tag{4.2}$$

where

$$\left[\boldsymbol{\Omega}_d\right]_{v,h} = \begin{cases} f_{d,u} - f_{d,l} & \text{if } v = h \\ \dfrac{e^{j2\pi f_{d,u}(v-h)} - e^{j2\pi f_{d,l}(v-h)}}{j2\pi(v-h)} & \text{if } v \neq h \end{cases} \tag{4.3}$$

for $(v, h) \in \{1, \ldots, K\}^2$. Let $\boldsymbol{\Omega} = \sum_{d=1}^{D} \omega_d \boldsymbol{\Omega}_d$ be a weighted interference covariance matrix, where the weights $\{\omega_d\}_{d=1}^{D}$ are assigned based on practical considerations regarding the impact of interference in the D bands. These include distance between the radar transmitter and interferenced systems, and tactical importance of the coexisting systems [46]. Given a radar waveform $\mathbf{y}_{\boldsymbol{\theta}_T} = f_{\boldsymbol{\theta}_T}(\mathbf{s})$, we define the spectral compatibility penalty function as

$$J_{\text{spectrum}}(\boldsymbol{\theta}_T) = \mathbf{y}_{\boldsymbol{\theta}_T}^H \boldsymbol{\Omega} \mathbf{y}_{\boldsymbol{\theta}_T}, \tag{4.4}$$

which is the total interfering energy from the radar waveform produced on the shared frequency bands.

4.3 Constrained Transmitter Design

In alternating training, the average loss (3.6) is modified by introducing a penalty function $J \in \{J_{\text{PAR}}, J_{\text{spectrum}}\}$ for a fixed receiver parameter vector $\boldsymbol{\theta}_R$. Accordingly,

we formulate the transmitter loss function, encompassing (3.6), (4.1) and (4.4), as

$$\begin{aligned}\mathcal{L}_{T,c}^{\pi}(\boldsymbol{\theta}_T) &= \mathcal{L}_{T}^{\pi}(\boldsymbol{\theta}_T) + \lambda J(\boldsymbol{\theta}_T) \\ &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\substack{\mathbf{a}\sim\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) \\ \mathbf{z}\sim p(\mathbf{z}|\mathbf{a},\mathcal{H}_i)}}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\big\} + \lambda J(\boldsymbol{\theta}_T).\end{aligned} \tag{4.5}$$

where λ controls the weight of the penalty $J(\boldsymbol{\theta}_T)$, and is referred to as the *penalty parameter*. When the penalty parameter λ is small, the transmitter is trained to improve its ability to adapt to the environment, while placing less emphasis on reducing the PAR level or interference energy from the radar waveform; and vice versa for large values of λ. Note that the waveform penalty function $J(\boldsymbol{\theta}_T)$ depends only on the transmitter trainable parameters $\boldsymbol{\theta}_T$. Thus, imposing the waveform constraint does not affect the receiver training.

It is straightforward to write the estimated version of the gradient (4.5) with respect to $\boldsymbol{\theta}_T$ by introducing the penalty as

$$\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_{T,c}^{\pi}(\boldsymbol{\theta}_T) = \nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_{T}^{\pi}(\boldsymbol{\theta}_T) + \lambda\nabla_{\boldsymbol{\theta}_T}J(\boldsymbol{\theta}_T), \tag{4.6}$$

where the gradient of the penalty function $\nabla_{\boldsymbol{\theta}_T}J(\boldsymbol{\theta}_T)$ is provided in Appendix A.

Substituting (3.8) into (4.6), we finally have the estimated gradient

$$\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_{T,c}^{\pi}(\boldsymbol{\theta}_T) = \frac{1}{Q_T}\sum_{q=1}^{Q_T}\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}^{(q)}),i^{(q)}\big)\nabla_{\boldsymbol{\theta}_T}\ln\pi(\mathbf{a}^{(q)}|\mathbf{y}_{\boldsymbol{\theta}_T}) + \lambda\nabla_{\boldsymbol{\theta}_T}J(\boldsymbol{\theta}_T), \tag{4.7}$$

which is used in the stochastic gradient update rule

$$\boldsymbol{\theta}_T^{(n+1)} = \boldsymbol{\theta}_T^{(n)} - \epsilon\nabla_{\boldsymbol{\theta}_T}\widehat{\mathcal{L}}_{T,c}^{\pi}(\boldsymbol{\theta}_T^{(n)}) \quad \text{for } n = 1,2,\ldots. \tag{4.8}$$

Like alternating training, simultaneous training can be directly extended to incorporate prescribed waveform constraints by adding the penalty term $\lambda J(\boldsymbol{\theta}_T)$ to the average loss (3.10).

CHAPTER 5

THEORETICAL PROPERTIES OF THE GRADIENTS

In this chapter, we discuss two useful theoretical properties of the gradients used for learning receiver and transmitter.

5.1 Receiver Gradient

As discussed previously, end-to-end learning of transmitted waveform and detector may be accomplished either by alternating or simultaneous training. The main difference between alternating and simultaneous training concerns the update of the receiver trainable parameter vector $\boldsymbol{\theta}_R$. Alternating training of $\boldsymbol{\theta}_R$ relies on a fixed waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$ (see Figure 3.1), while simultaneous training relies on random waveforms $\mathbf{a}$ generated in accordance with a preset policy, i.e., $\mathbf{a} \sim \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$, as shown in Figure 3.3. The relation between the gradient applied by alternating training, $\nabla_{\boldsymbol{\theta}_R}\mathcal{L}_R(\boldsymbol{\theta}_R)$, and the gradient of simultaneous training, $\nabla_{\boldsymbol{\theta}_R}L^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T)$, with respect to $\boldsymbol{\theta}_R$ is stated by the following proposition.

Proposition 1. *For the loss function (3.1) computed based on a waveform* $\mathbf{y}_{\boldsymbol{\theta}_T}$ *and loss function (3.10) computed based on a stochastic policy* $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ *continuous in* $\mathbf{a}$*, the following equality holds:*

$$\nabla_{\boldsymbol{\theta}_R}\mathcal{L}_R(\boldsymbol{\theta}_R) = \nabla_{\boldsymbol{\theta}_R}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T). \tag{5.1}$$

Proof. See Appendix B. ■

Proposition 1 states that the gradient of simultaneous training, $\nabla_{\boldsymbol{\theta}_R}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T)$, equals the gradient of alternating training, $\nabla_{\boldsymbol{\theta}_R}\mathcal{L}_R(\boldsymbol{\theta}_R)$, even though simultaneous

training applies a random waveform $\mathbf{a} \sim \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ to train the receiver. Note that this result applies only to ensemble means according to Equation (3.3) and Equation (3.11), and not to the empirical estimates used by Algorithms 1 and 2. Nevertheless, Proposition 1 suggests that training updates of the receiver are unaffected by the choice of alternating or simultaneous training. That said, given the distinct updates of the transmitter's parameter, the overall trajectory of the parameters $(\boldsymbol{\theta}_R,\ \boldsymbol{\theta}_T)$ during training may differ according to the two algorithms.

5.2 Transmitter Gradient

As shown in Chapter 3, the gradients used for learning receiver parameters $\boldsymbol{\theta}_R$ by alternating training (3.4) or simultaneous training (3.13) may be directly estimated from the channel output samples $\mathbf{z}^{(q)}$. In contrast, the gradient used for learning transmitter parameters $\boldsymbol{\theta}_T$ according to Equation (3.1) cannot be directly estimated from the channel output samples. To obviate this problem, in Algorithms 1 and 2, the transmitter is trained by exploring the space of transmitted waveforms according to a policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$. We refer to the transmitter loss gradient obtained via policy gradient (3.12) as the *RL transmitter gradient*. The benefit of RL-based transmitter training is that it renders unnecessary access to the likelihood function $p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$ to evaluate the RL transmitter gradient, rather the gradient is estimated via samples. We now formalize the relation of the RL transmitter gradient (3.12) and the transmitter gradient for a known likelihood obtained according to Equation (3.1).

As mentioned, if the likelihood $p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$ were known, and if it were differentiable with respect to the transmitter parameter vector $\boldsymbol{\theta}_T$, the transmitter parameter vector $\boldsymbol{\theta}_T$ may be learned by minimizing the average loss (3.1), which we rewrite as a function of both $\boldsymbol{\theta}_R$ and $\boldsymbol{\theta}_T$ as

$$\mathcal{L}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i)\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big)\big\}. \tag{5.2}$$

The gradient of Equation (5.2) with respect to $\boldsymbol{\theta}_T$ is expressed as

$$\nabla_{\boldsymbol{\theta}_T}\mathcal{L}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)=\sum_{i\in\{0,1\}}P(\mathcal{H}_i)\mathbb{E}_{\mathbf{z}\sim p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)}\big\{\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}),i\big)\nabla_{\boldsymbol{\theta}_T}\ln p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)\big\}, \tag{5.3}$$

where the equality leverages the following relation

$$\nabla_{\boldsymbol{\theta}_T}p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)=p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)\nabla_{\boldsymbol{\theta}_T}\ln p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i). \tag{5.4}$$

The relation between the RL transmitter gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ in Equation (3.12) and the transmitter gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ in (5.3) is elucidated by the following proposition.

Proposition 2. *If likelihood function $p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T},\mathcal{H}_i)$ is differentiable with respect to the transmitter parameter vector $\boldsymbol{\theta}_T$ for $i\in\{0,1\}$, the following equality holds*

$$\nabla_{\boldsymbol{\theta}_T}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)=\nabla_{\boldsymbol{\theta}_T}\mathcal{L}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T). \tag{5.5}$$

Proof. See Appendix C. ■

Proposition 2 establishes that the RL transmitter gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ equals the transmitter gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ for any given receiver parameters $\boldsymbol{\theta}_R$. Proposition 2 hence provides a theoretical justification for replacing the gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ with the RL gradient $\nabla_{\boldsymbol{\theta}_T}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R,\boldsymbol{\theta}_T)$ to perform transmitter training as done in Algorithms 1 and 2.

CHAPTER 6

NUMERICAL RESULTS

This chapter first introduces the simulation setup, and then it presents numerical examples of waveform design and detection performance that compare the proposed data-driven methodology with existing model-based approaches. While simulation results presented in this chapter rely on various models of target, clutter and interference, this work expressly distinguishes data-driven learning from model-based design. Learning schemes rely solely on data and not on model information. In contrast, model-based design implies a system structure that is based on a specific and known model. Furthermore, learning may rely on synthetic data containing diverse data that is generated according to a variety of models. On the contrary, model-based design typically relies on a single model. For example, as we will see, a synthetic dataset for learning may contain multiple clutter sample sets, each generated according to a different clutter model. Conversely, a single clutter model is typically assumed for model-based design.

6.1 Models, Policy, and Parameters

6.1.1 Models of target, clutter, and noise

The target is stationary, and has a Rayleigh envelope, i.e., $\alpha \sim \mathcal{CN}(0, \sigma_\alpha^2)$. where σ_α^2 is the target power. The noise has a zero-mean Gaussian distribution with correlation matrix $\mathbf{\Omega}_n = \sigma_n^2 \mathbf{I}_K + \mathbf{\Omega}_I$, where σ_n^2 is the thermal noise power level and $\mathbf{\Omega}_I$ is the correlation matrix of signal-independent interference. Elements of $\mathbf{\Omega}_I$ are $[\mathbf{\Omega}_I]_{v,h} = \rho^{|v-h|}$ for $(v, h) \in \{1, \ldots, K\}^2$, where ρ is the one-lag correlation coefficient. The clutter vector in Equation (2.4) is the superposition of returns from $2K - 1$ consecutive range cells, reflecting all clutter illuminated by the K-length signal as it

sweeps in range across the target. Accordingly, the clutter vector may be expressed as [3]

$$\mathbf{c} = \sum_{g=-K+1}^{K-1} \gamma_g \mathbf{J}_g \mathbf{y}, \tag{6.1}$$

where $\mathbf{J}_g$ represents the shifting matrix at the gth range cell with elements

$$\left[\mathbf{J}_g\right]_{v,h} = \begin{cases} 1 & \text{if} \quad v-h=g \\ 0 & \text{if} \quad v-h \neq g \end{cases} \quad (v,h) \in \{1, \cdots, K\}^2 \, . \tag{6.2}$$

The magnitude $|\gamma_g|$ of the gth clutter scattering coefficient is generated according to a Weibull distribution [5]

$$p(|\gamma_g|) = \frac{\beta}{\nu^\beta} |\gamma_g|^{\beta-1} \exp\left(- \frac{|\gamma_g|^\beta}{\nu^\beta} \right), \tag{6.3}$$

where β is the shape parameter and ν is the scale parameter of the distribution. Let $\sigma^2_{\gamma_g}$ represent the power of the clutter scattering coefficient γ_g. The relation between $\sigma^2_{\gamma_g}$ and the Weibull distribution parameters $\{\beta, \nu\}$ is [47]

$$\sigma^2_{\gamma_g} = \mathrm{E}\{|\gamma_g|^2\} = \frac{2\nu^2}{\beta} \Gamma\left(\frac{2}{\beta}\right), \tag{6.4}$$

where $\Gamma(\cdot)$ is the Gamma function. The nominal range of the shape parameter is $0.25 \leq \beta \leq 2$ [48]. In the simulation, the complex-valued clutter scattering coefficient γ_g is obtained by multiplying a real-valued Weibull random variable $|\gamma_g|$ with the factor $\exp(j\psi_g)$, where ψ_g is the phase of γ_g distributed uniformly in the interval

$(0, 2\pi)$. When the shape parameter $\beta = 2$, the clutter scattering coefficient γ_g follows the Gaussian distribution $\gamma_g \sim \mathcal{CN}(0, \sigma^2_{\gamma_g})$. A homogeneous clutter environment is considered. Based on the assumed mathematical models of the target, clutter and noise, the optimal detector in the NP sense is available (see details in Appendix D), and the adaptive waveform for target detection can be obtained by maximizing the signal-to-clutter-plus-noise ratio at the receiver output at the time of target detection (see details in Appendix E).

6.1.2 Transmitter and receiver models

Waveform generation and detection is implemented using feedforward neural networks as explained in Section 2.2. The transmitter $\tilde{f}_{\boldsymbol{\theta}_T}(\cdot)$ is a feedforward neural network with four layers, i.e., an input layer with $2K$ neurons, two hidden layers with $M = 48$ neurons, and an output layer with $2K$ neurons. The activation function is exponential linear unit (ELU) [49]. The receiver $\tilde{f}_{\boldsymbol{\theta}_R}(\cdot)$ is implemented as a feedforward neural network with four layers, i.e., an input layer with $2K$ neurons, two hidden layers with M neurons, and an output layer with one neuron. The sigmoid function is chosen as the activation function. The layout of transmitter and receiver networks is summarized in Table 6.1.

Table 6.1 Layout of the Transmitter and Receiver Networks

	Transmitter $\tilde{f}_{\boldsymbol{\theta}_T}(\cdot)$			Receiver $\tilde{f}_{\boldsymbol{\theta}_R}(\cdot)$		
Layer	1	2-3	4	1	2-3	4
Dimension	$2K$	M	$2K$	$2K$	M	1
Activation	-	ELU	Linear	-	Sigmoid	Sigmoid

6.1.3 Gaussian policy

A Gaussian policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ is adopted for RL-based transmitter training. Accordingly, the output of the stochastic transmitter follows a complex Gaussian distribution $\mathbf{a} \sim \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) = \mathcal{CN}\big(\sqrt{1-\sigma_p^2}\mathbf{y}_{\boldsymbol{\theta}_T}, \frac{\sigma_p^2}{K}\mathbf{I}_K\big)$, where the per-chip variance σ_p^2 is referred to as

the *policy hyperparameter*. When $\sigma_p^2 = 0$, the stochastic policy becomes deterministic [50], i.e., the policy is governed by a Dirac function at $\mathbf{y}_{\boldsymbol{\theta}_T}$. In this case, the policy does not explore the space of transmitted waveforms, but it "exploits" the current waveform. At the opposite end, when $\sigma_p^2 = 1$, the output of the stochastic transmitter is independent of $\mathbf{y}_{\boldsymbol{\theta}_T}$, and the policy becomes zero-mean complex-Gaussian noise with covariance matrix $\mathbf{I}_K/K$. Thus, the policy hyperparameter σ_p^2 is selected in the range $(0, 1)$, and its value sets a trade-off between exploration of new waveforms versus exploitation of current waveform.

6.1.4 Training parameters

The initialization waveform $\mathbf{s}$ is a linear frequency modulated (LFM) pulse with $K = 16$ complex-valued chips and chirp rate $R = (100 \times 10^3)/(40 \times 10^{-6})$ Hz/s. Specifically, the kth chip of $\mathbf{s}$ is given by

$$\mathbf{s}(k) = \frac{1}{\sqrt{K}} \exp\left\{j\pi R\big(k/f_s\big)^2\right\} \tag{6.5}$$

$\forall k \in \{0, \ldots, K-1\}$, where $f_s = 200$ kHz. The signal-to-clutter ratio is defined as

$$\text{SCR} \triangleq 10 \log_{10}\left\{\sigma_\alpha^2/\sigma_\gamma^2\right\}, \tag{6.6}$$

where σ_γ^2 is the aggregated clutter power of the range cells covered by the radar waveform. The clutter-to-noise ratio is defined as

$$\text{CNR} \triangleq 10 \log_{10}\left\{\sigma_\gamma^2/\sigma_n^2\right\}. \tag{6.7}$$

The training is performed at SCR = 10 dB and CNR = 20 dB. The one-lag correlation coefficient of the signal-independent interference is $\rho = 0.7$. Letting the subscripts "train" and "test" represent quantities associated with the training stage and testing stage, respectively. Unless stated otherwise, we set $\text{SCR}_{\text{train}} = \text{SCR}_{\text{test}}$, and the shape parameters of the clutter distribution (6.3) $\beta_{\text{train}} = \beta_{\text{test}} = 2$.

To obtain a balanced classification dataset, the training set is populated by samples belonging to either hypothesis with equal prior probability, i.e., $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 0.5$. The number of training samples is set as $Q_R = Q_T = Q = 2^{12}$ in the estimated gradients (3.4), (3.8), (3.13), and (3.14). Unless stated otherwise, the policy parameter is set to $\sigma_p^2 = 10^{-2}$, and the penalty parameter is $\lambda = 0$, i.e., there are no waveform constraints. The Adam optimizer [51] is adopted to train the system over a number of iterations chosen by trial and error. The learning rate is $\epsilon = 0.005$. In the testing phase, 2×10^5 samples are used to estimate the probability of false alarm (P_{fa}) under hypothesis $\mathcal{H}_0$, while 5×10^4 samples are used to estimate the probability of detection (P_d) under hypothesis $\mathcal{H}_1$. Receiver operating characteristic (ROC) curves are obtained via Monte Carlo simulations by varying the threshold applied at the output of the receiver. Results are obtained by averaging over fifty trials. Numerical results presented in this section assume simultaneous training, unless stated otherwise.

6.2 Results and Discussion

This section shows results and discussion relating to the proposed approaches.

6.2.1 Simultaneous training versus training with known likelihood

We first analyze the impact of the choice of the policy hyperparameter σ_p^2 on the performance on the training set. Figure 6.1 shows the empirical cross-entropy loss of simultaneous training versus the policy hyperparameter σ_p^2 upon the completion of the training process. The empirical loss of the system training with a known channel

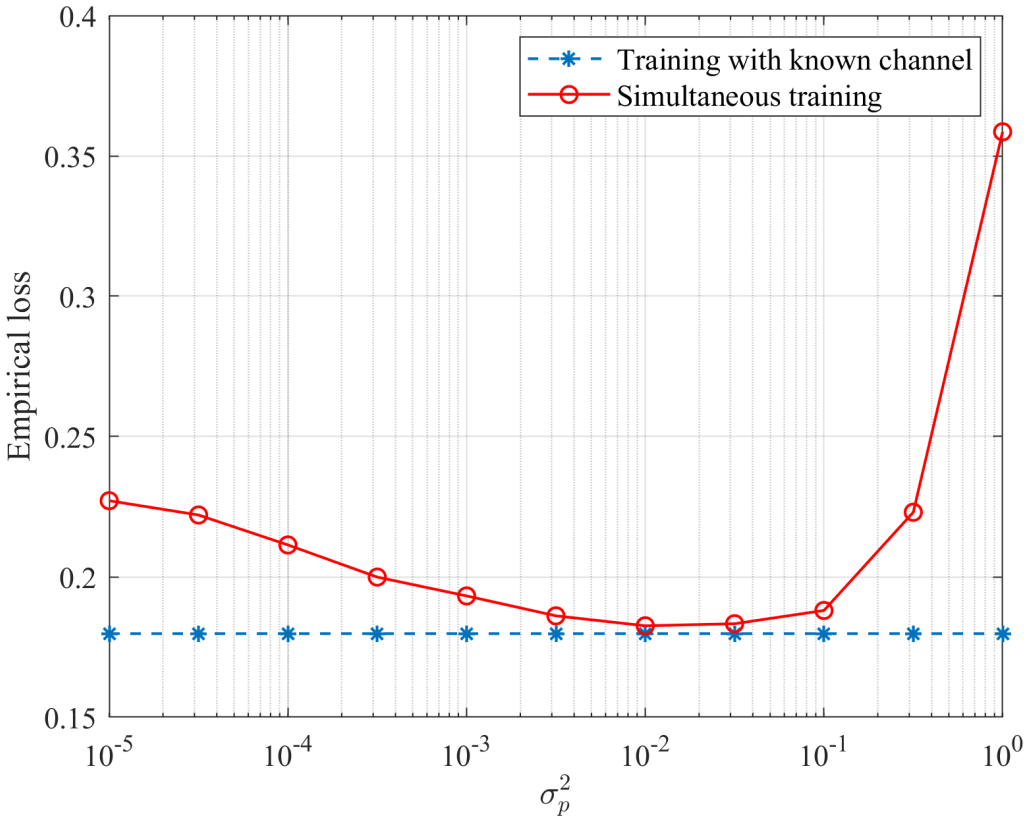

Figure 6.1 Empirical training loss versus policy hyperparameter σ_p^2 for simultaneous training algorithm and training with known channel, respectively.

(5.2) is plotted as a comparison. It is seen that there is an optimal policy parameter σ_p^2 for which the empirical loss of simultaneous training approaches the loss with known channel. As the policy hyperparameter σ_p^2 tends to 0, the output of the stochastic transmitter $\mathbf{a}$ is close to the waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$, which leads to no exploration of the space of transmitted waveforms. In contrast, when the policy parameter σ_p^2 tends to 1, the output of the stochastic transmitter becomes a complex Gaussian noise with zero mean and covariance matrix $\mathbf{I}_K/K$. In both cases, the RL transmitter gradient is difficult to estimate accurately.

While Figure 6.1 evaluates the performance on the training set in terms of empirical cross-entropy loss, the choice of the policy hyperparameter σ_p^2 should be based on validation data and in terms of the testing criterion that is ultimately of

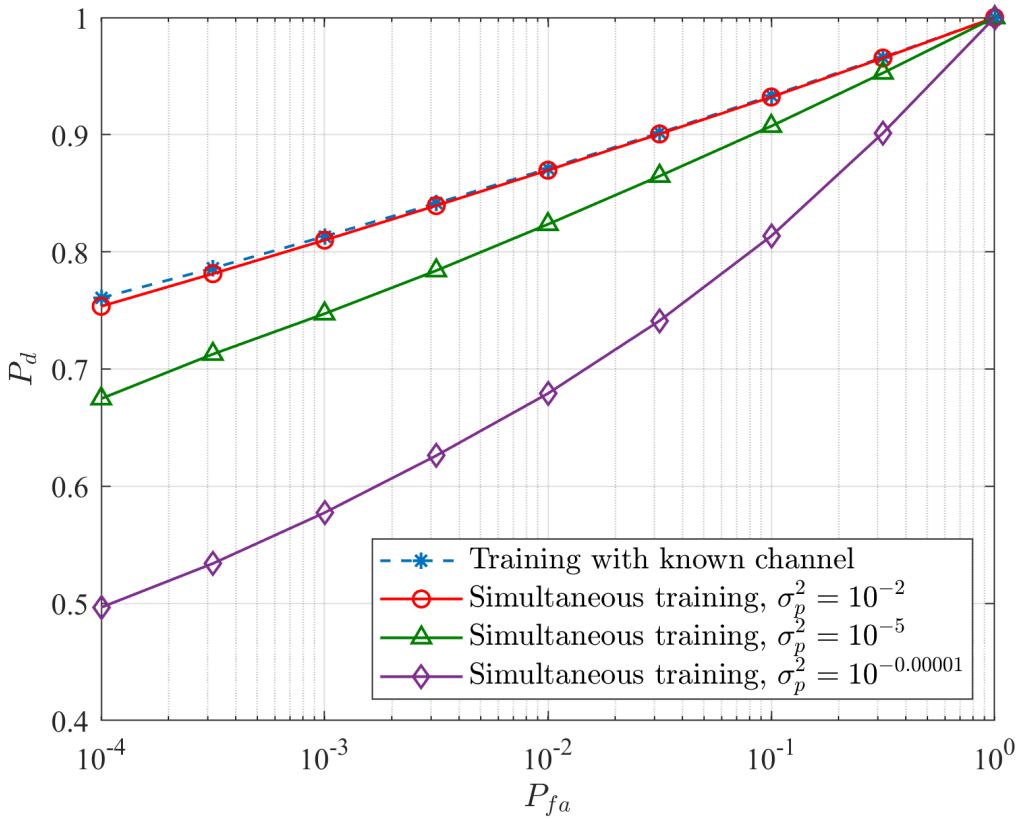

Figure 6.2 ROC curves for training with known channel and simultaneous training with different values of policy parameter σ_p^2.

interest. To elaborate on this point, ROC curves obtained by simultaneous training with different values of the policy hyperparameter σ_p^2 and training with known channel are shown in Figure 6.2. As shown in the figure, simultaneous training with $\sigma_p^2 = 10^{-2}$ achieves a similar ROC as training with known channel. The choice $\sigma_p^2 = 10^{-2}$, also has the lowest empirical training loss in Figure 6.1. These results suggest that training is not subject to overfitting [52].

6.2.2 Simultaneous training versus alternating training

We now compare simultaneous and alternating training in terms of ROC curves in Figure 6.3. ROC curves based on the optimal detector in the NP sense, namely, the clairvoyant detector (D.5) and the adaptive/LFM waveform are plotted as

benchmarks. Moreover, the ROC curve based on the matched filter to the LFM waveform (6.5) is illustrated as well. As shown in the figure, simultaneous training provides a similar detection performance as alternating training. Furthermore, both simultaneous training and alternating training are seen to result in significant improvements as compared to training of only the receiver, and provide detection performance comparable to adaptive waveform and clairvoyant detector.

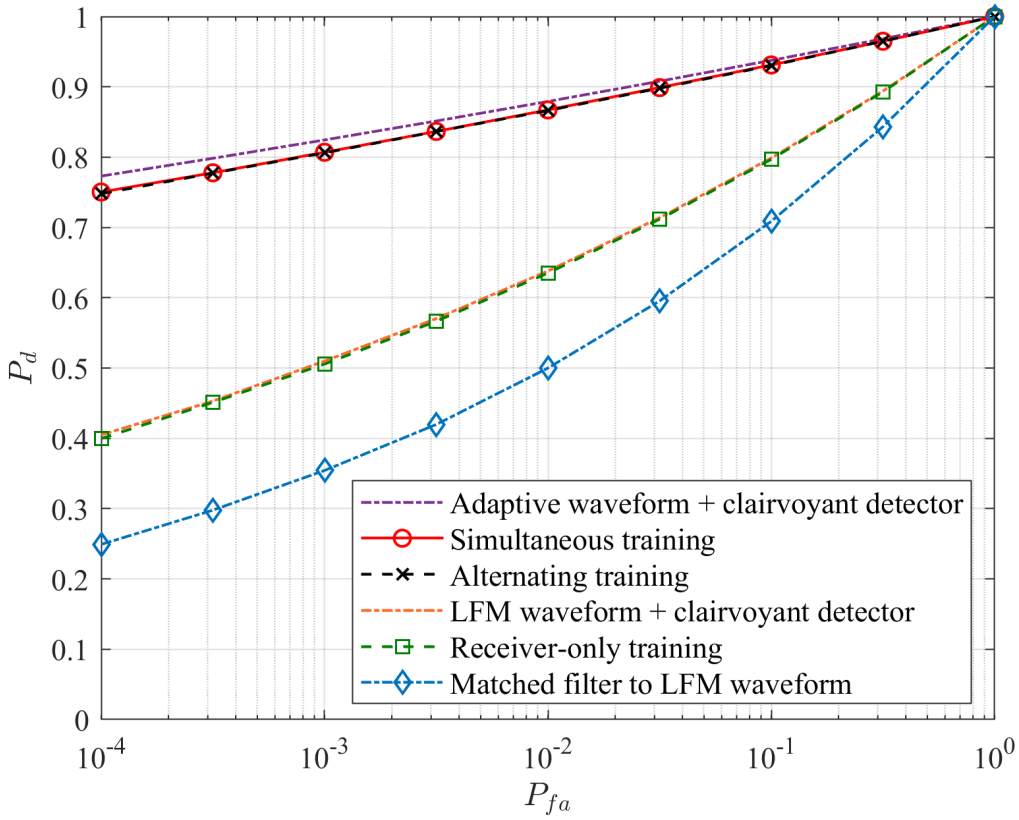

Figure 6.3 ROC curves with and without transmitter training.

6.2.3 Simultaneous training with different $\mathrm{SCR}_{\mathrm{test}}$ levels

Figure 6.4 shows how the end-to-end system trained by simultaneous training works at different $\mathrm{SCR}_{\mathrm{test}}$ levels, given $P_{fa} = 10^{-4}$ and $\mathrm{SCR}_{\mathrm{train}} = 10$ dB. As shown in the figure, even though there is a mismatch in terms of SCR level utilized during the training and testing stages, simultaneous training still provides a comparable

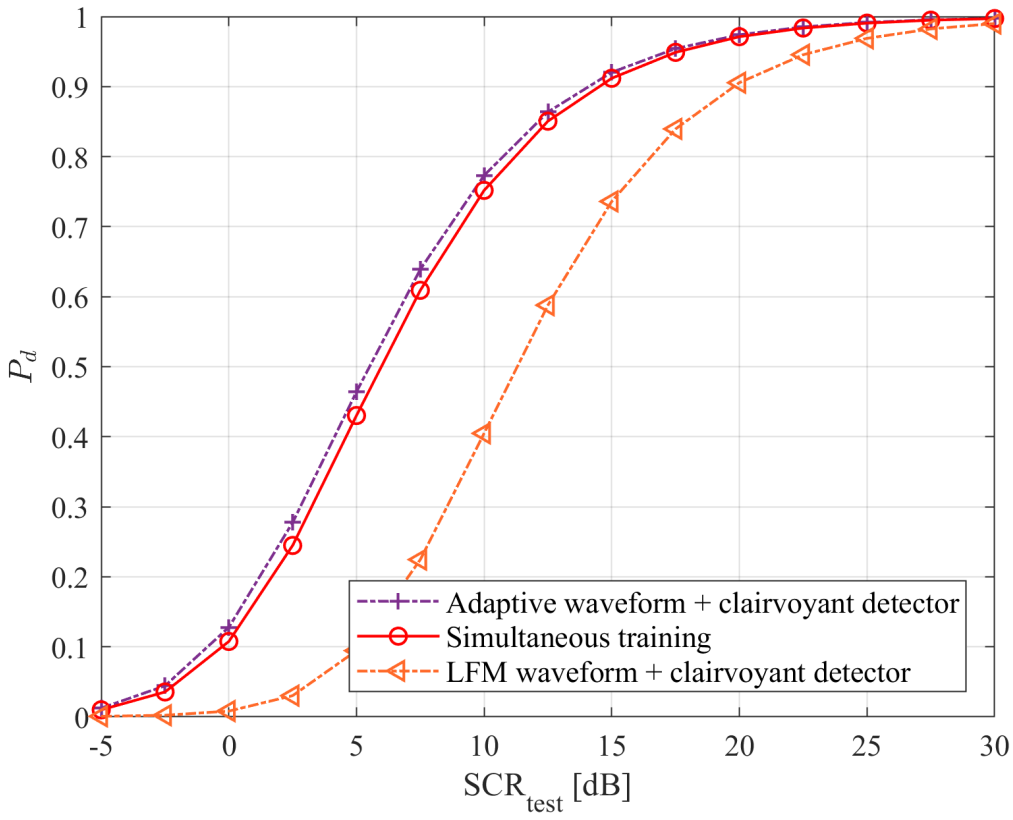

Figure 6.4 Probability of detection P_d versus SCR_{test} level for simultaneous training with probability of false alarm $P_{fa} = 10^{-4}$ and $\text{SCR}_{\text{train}} = 10$ dB.

detection performance with the adaptive waveform and the clairvoyant detector (D.5). For instance, for $\text{SCR}_{\text{test}} = 15$ dB, both simultaneous training and detection via adaptive waveform and clairvoyant detector yield $P_d = 0.91$, while detection via LFM waveform (6.5) and clairvoyant detector yields $P_d = 0.74$.

6.2.4 Learning Gaussian and non-Gaussian clutter

Two sets of ROC curves under different clutter statistics are illustrated in Figure 6.5. Each set contains two ROC curves with the same clutter statistics: one curve is obtained based on simultaneous training, and the other one is based on model-based design. For simultaneous training, the shape parameter of the clutter distribution (6.3) in the training stage is the same as that in the test stage, i.e., $\beta_{\text{train}} = \beta_{\text{test}}$.

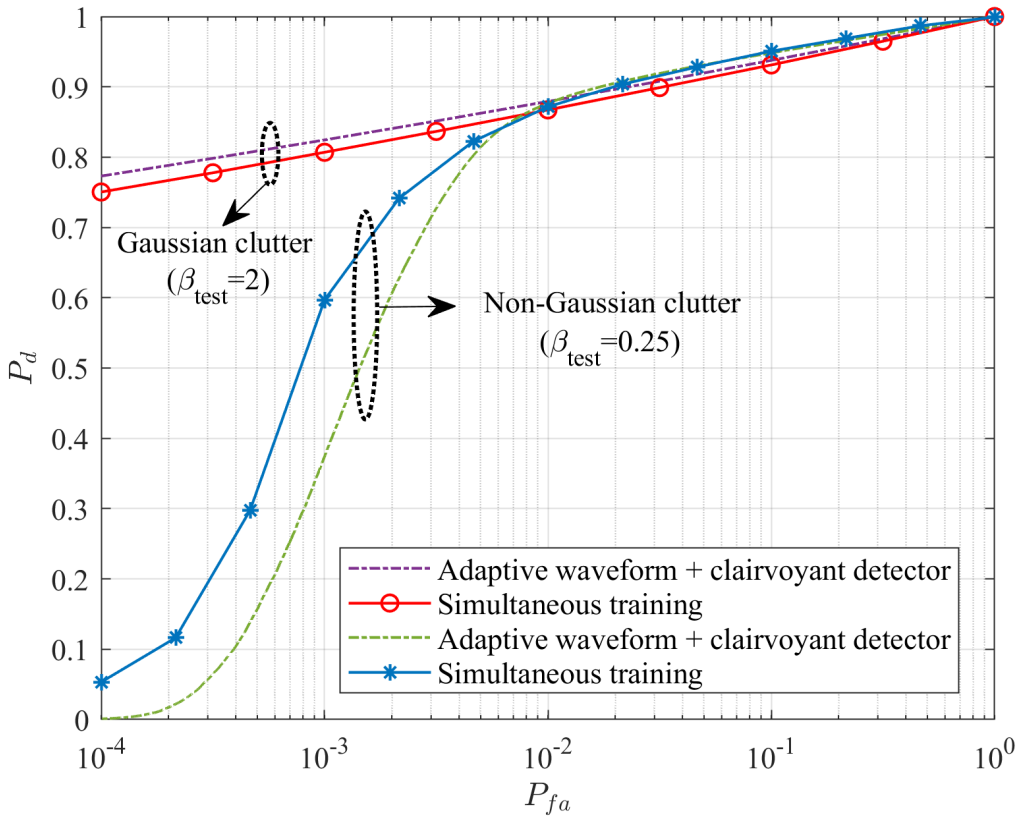

Figure 6.5 ROC curves for Gaussian/non-Gaussian clutter. The end-to-end radar system is trained and tested by the same clutter statistics, i.e., $\beta_{\text{train}} = \beta_{\text{test}}$.

In the test stage, for Gaussian clutter ($\beta_{\text{test}} = 2$), the model-based ROC curve is obtained by the adaptive waveform and the optimal detector in the NP sense. As expected, simultaneous training provides a comparable detection performance with the adaptive waveform and clairvoyant detector (also shown in Figure 6.3). In contrast, when the clutter is non-Gaussian ($\beta_{\text{test}} = 0.25$), the optimal detector in the NP sense is mathematically intractable. Under this scenario, the data-driven approach is beneficial since it relies on data rather than a model. As observed in the figure, for non-Gaussian clutter with a shape parameter $\beta_{\text{test}} = 0.25$, simultaneous training outperforms the adaptive waveform and clairvoyant detector.

6.2.5 Simultaneous training with mixed clutter statistics

The robustness of the trained radar system to the clutter statistics is investigated next. As discussed previously, model-based design relies on a single clutter model, whereas data-driven learning depends on a training dataset. The dataset may contain samples from multiple clutter models. Thus, the system based on data-driven learning may be robustified by drawing samples from a mixture of clutter models. In the test stage, the clutter model may not be the same as any of the clutter models used in the training stage. As shown in Figure 6.6, for simultaneous training, the training dataset contains clutter samples generated from Equation (6.3) with four different values of shape parameter $\beta_{\text{train}} \in \{0.25, 0.5, 0.75, 1\}$. The test data is generated with a clutter shape parameter $\beta_{\text{test}} = 0.3$ not included in the training dataset. The end-to-end leaning radar system trained by mixing clutter samples provides performance gains compared to a model-based system using an adaptive waveform and clairvoyant detector.

6.2.6 Simultaneous training under PAR constraint

Detection performance with waveforms learned subject to a PAR constraint is shown in Figure 6.7. The end-to-end system trained with no PAR constraint, i.e., $\lambda = 0$, serves as the reference. It is observed the detection performance degrades as the value of the penalty parameter λ increases. Moreover, PAR values of waveforms with different λ are shown in Table 6.2. As shown in Figure 6.7 and Table 6.2, there is a trade-off between detection performance and PAR level. For instance, given $P_{fa} = 5 \times 10^{-4}$, training the transmitter with the largest penalty parameter $\lambda = 0.1$ yields the lowest $P_d = 0.786$ with the lowest PAR value 0.47 dB. In contrast, training the transmitter with no PAR constraint, i.e., $\lambda = 0$, yields the best detection with the largest PAR value 6.02 dB. Figure 6.8 compares the normalized modulus of waveforms with different values of the penalty parameter λ. As shown in Figure 6.8 and Table

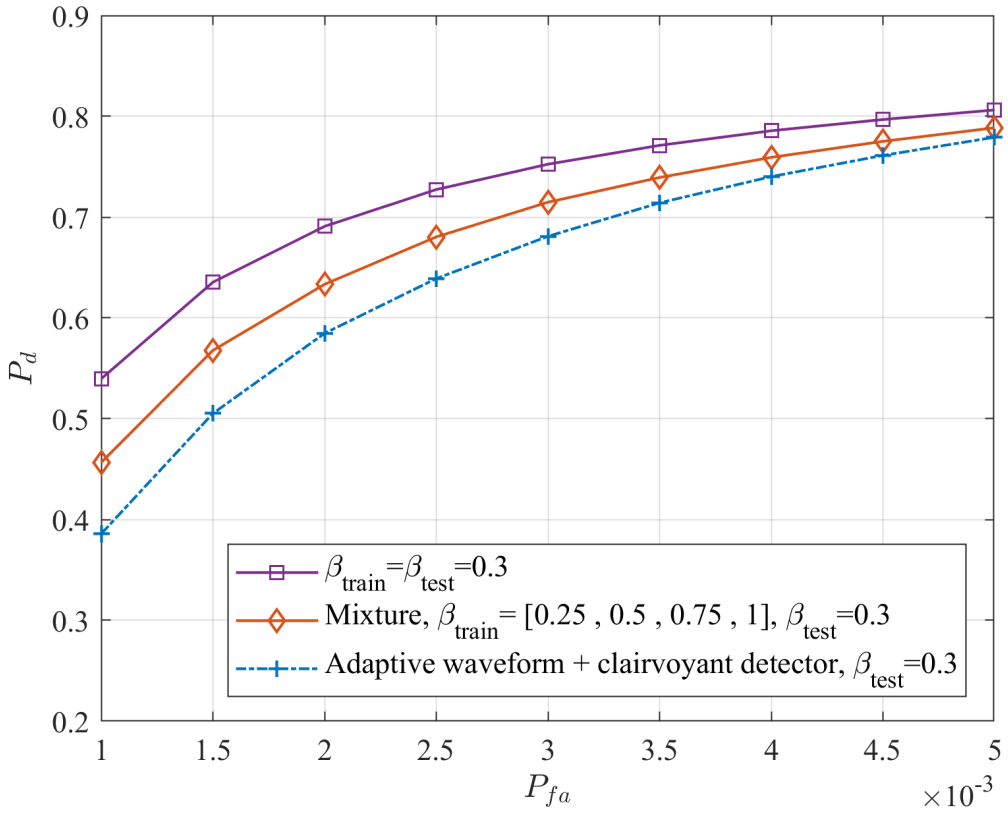

Figure 6.6 ROC curves for non-Gaussian clutter. To robustify detection performance, the end-to-end leaning radar system is trained with mixed clutter statistics, while testing for a clutter model different than used for training.

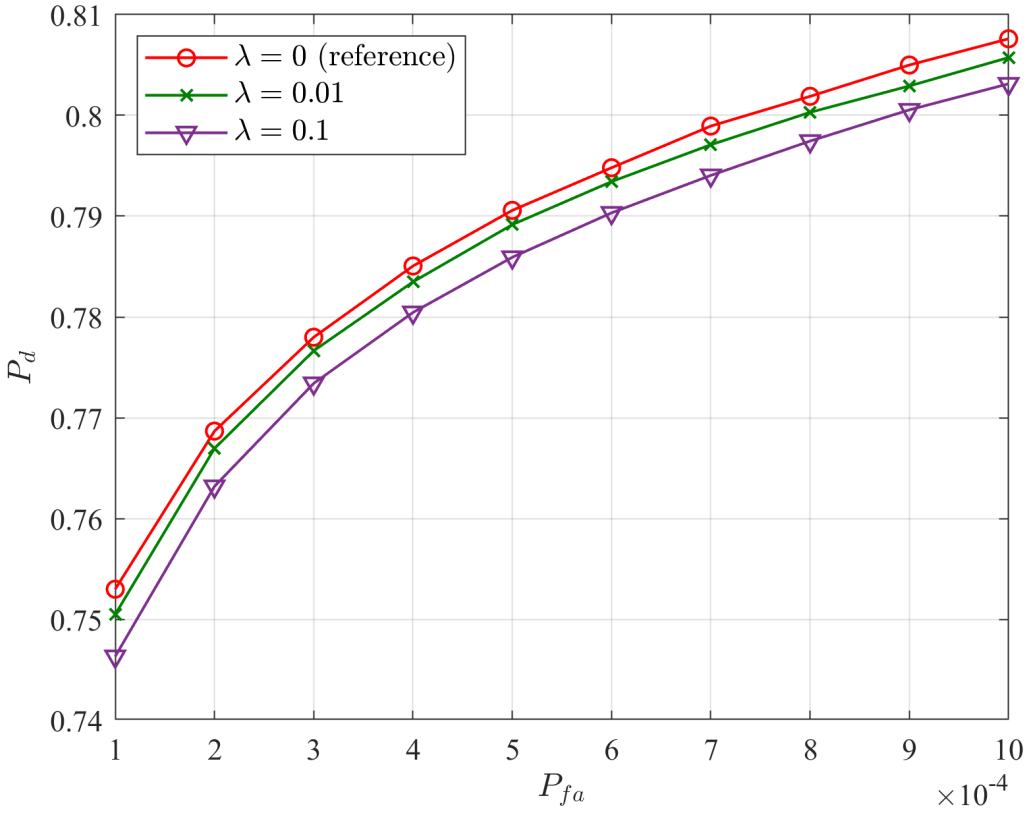

Figure 6.7 ROC curves for PAR constraint with different values of the penalty parameter λ.

6.2, the larger the penalty parameter λ adopted in the simultaneous training, the smaller the PAR value of the waveform.

Table 6.2 PAR Values of Waveforms with Different Values of Penalty Parameter λ

	$\lambda = 0$ (reference)	$\lambda = 0.01$	$\lambda = 0.1$
PAR [dB] (4.1)	6.02	2.70	0.47

6.2.7 Simultaneous training under spectral compatibility constraint

ROC curves for spectral compatibility constraint with different values of the penalty parameter λ are illustrated in Figure 6.9. The shared frequency bands are $\Gamma_1 = [0.3, 0.35]$ and $\Gamma_2 = [0.5, 0.6]$. The end-to-end system trained with no spectral

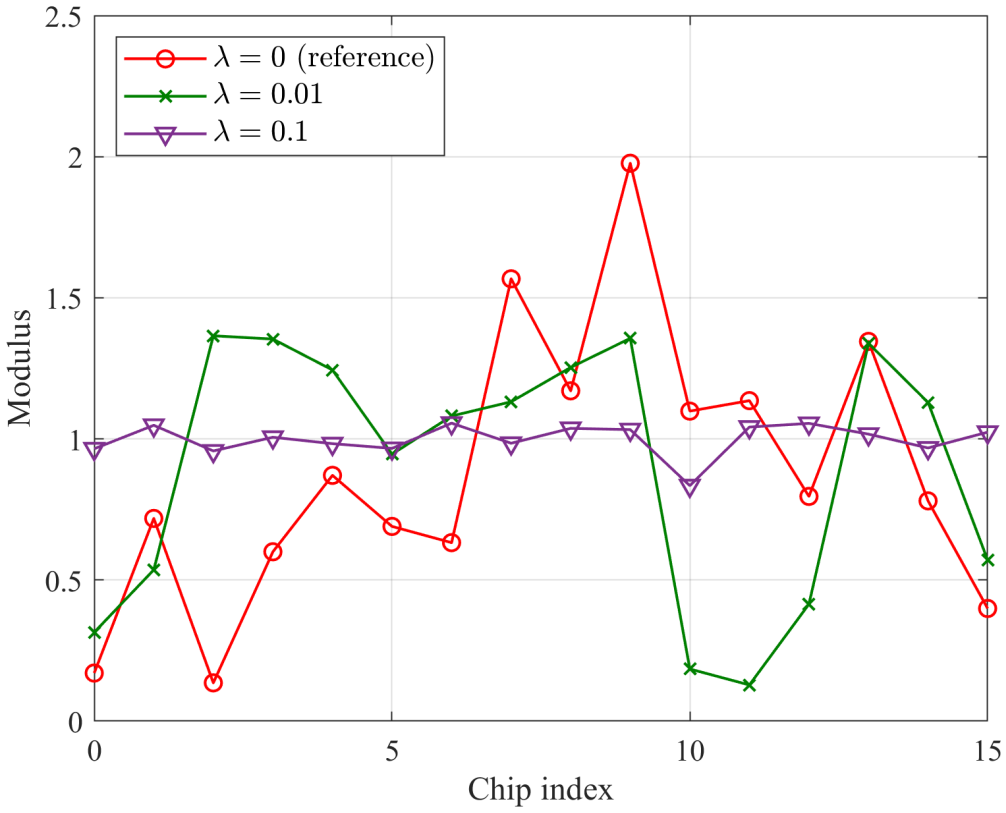

Figure 6.8 Normalized modulus of transmitted waveforms with different values of penalty parameter λ.

Table 6.3 Interfering Energy from Radar Waveforms with Different Values of Penalty Parameter λ

	$\lambda = 0$ (reference)	$\lambda = 0.5$	$\lambda = 5$
Interfering energy [dB] (4.4)	-5.95	-14.47	-26.40

compatibility constraint, i.e., $\lambda = 0$, serves as the reference. Training the transmitter with a large value of the penalty parameter λ is seen to result in performance degradation. Interfering energy from radar waveforms trained with different values of λ are shown in Table 6.3. It is observed that λ plays an important role in controlling the tradeoff between detection performance and spectral compatibility of the waveform. For instance, for a fixed $P_{fa} = 5 \times 10^{-4}$, training the transmitter with $\lambda = 0$ yields $P_d = 0.79$ with an amount of interfering energy -5.95 dB on the shared frequency bands, while training the transmitter with $\lambda = 5$ creates notches in the spectrum of the transmitted waveform at the shared frequency bands. Energy spectral densities of transmitted waveforms with different values of λ are illustrated in Figure 6.10. A larger the penalty parameter λ results in a lower amount of interfering energy in the prescribed frequency shared regions. Note, for instance, that the nulls of the energy spectrum density of the waveform for $\lambda = 5$ are much deeper than their counterparts for $\lambda = 0.5$.

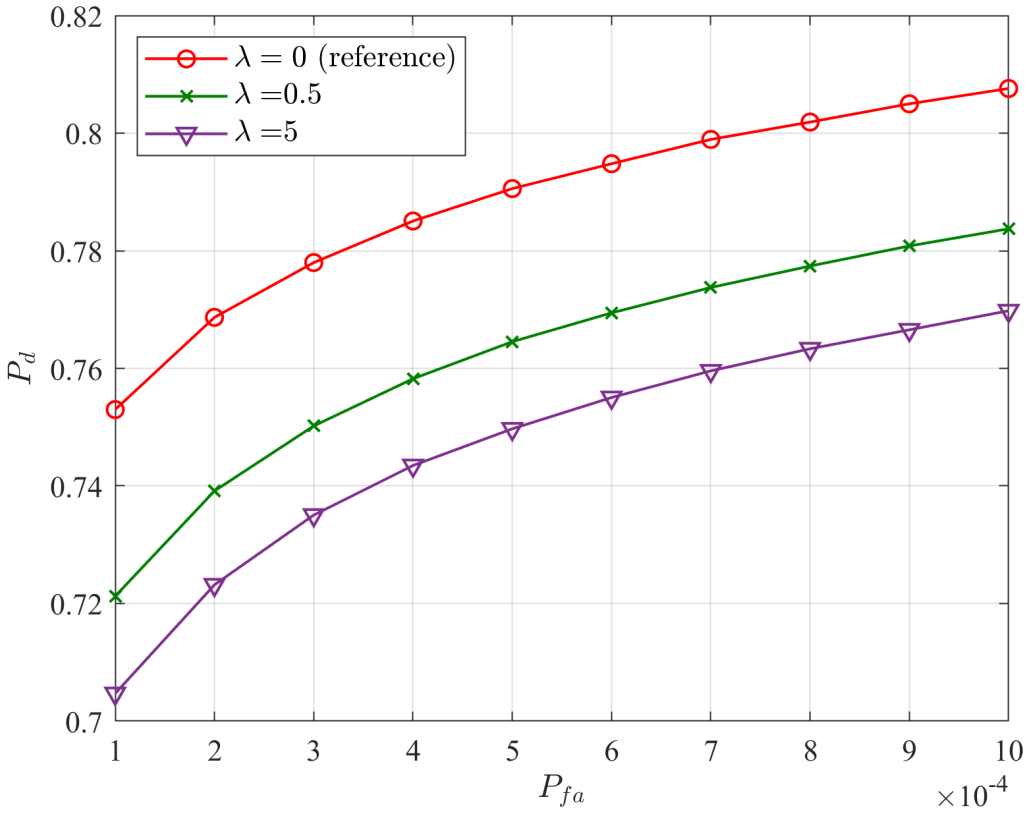

Figure 6.9 ROC curves for spectral compatibility constraint with different values of penalty parameter λ.

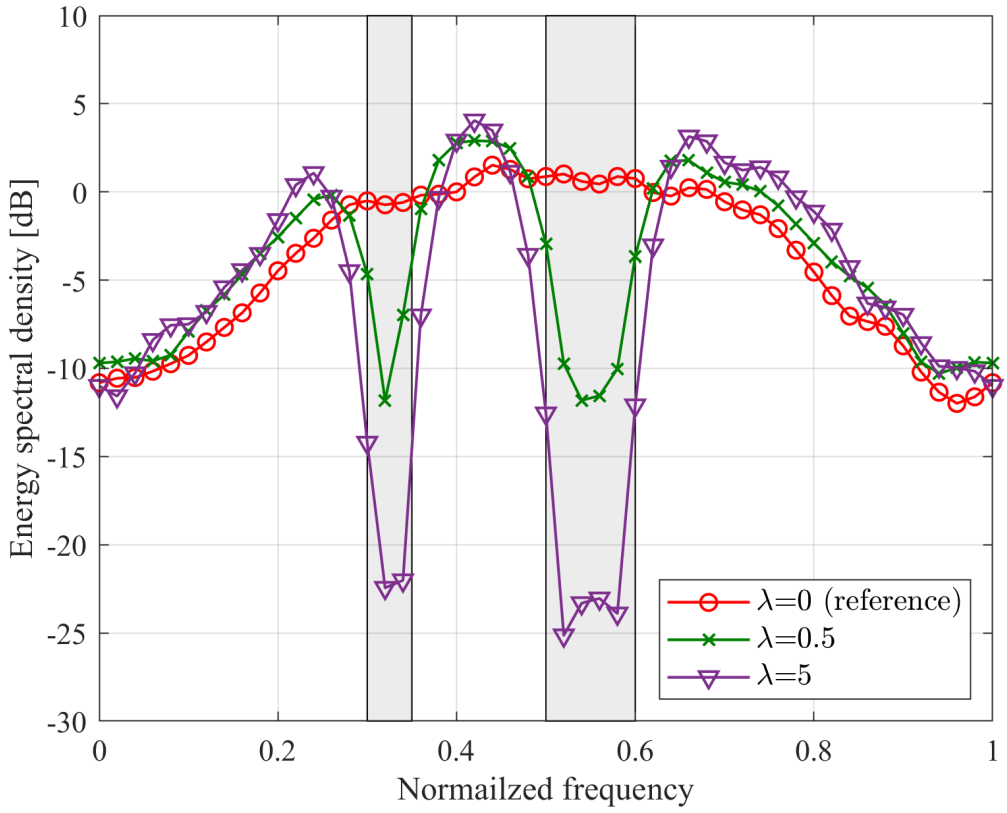

Figure 6.10 Energy spectral density of waveforms with different values of penalty parameter λ.

CHAPTER 7

END NOTES

In this dissertation, we have formulated the radar design problem as end-to-end learning of waveform generation and detection. We have developed two training algorithms, both of which are able to incorporate various waveform constraints into the system design. Training may be implemented either as simultaneous supervised training of the receiver and RL-based training of the transmitter, or as alternating between training of the receiver and of the transmitter. Both training algorithms have similar performance. We have also robustified the detection performance by training the system with mixed clutter statistics. Numerical results have shown that the proposed end-to-end learning approaches are beneficial under non-Gaussian clutter, and successfully learn the transmitted waveform to adapt to actual statistics of environmental conditions, while satisfying operational constraints.

APPENDIX A

GRADIENT OF PENALTY FUNCTIONS

In this appendix, the respective gradients of the penalty functions (4.1) and (4.4) are derived with respect to the transmitter parameter vector $\boldsymbol{\theta}_T$. To facilitate the presentation, let $\overline{\mathbf{y}}_{\boldsymbol{\theta}_T}$ represent a $2K \times 1$ real vector comprising the real and imaginary parts of the waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$, i.e., $\overline{\mathbf{y}}_{\boldsymbol{\theta}_T} = \left[\Re(\mathbf{y}_{\boldsymbol{\theta}_T}), \Im(\mathbf{y}_{\boldsymbol{\theta}_T})\right]^T$.

Gradient of PAR Penalty Function As discussed in Section 4.1, the transmitted power is normalized such that $||\mathbf{y}_{\boldsymbol{\theta}_T}||^2 = ||\overline{\mathbf{y}}_{\boldsymbol{\theta}_T}||^2 = 1$. Let subscript "max" represent the chip index associated with the PAR value (4.1). By leveraging the chain rule, the gradient of Equation (4.1) with respect to $\boldsymbol{\theta}_T$ is written

$$\nabla_{\boldsymbol{\theta}_T} J_{\text{PAR}}(\boldsymbol{\theta}_T) = \nabla_{\boldsymbol{\theta}_T} \overline{\mathbf{y}}_{\boldsymbol{\theta}_T} \cdot \mathbf{g}_{\text{PAR}}, \tag{A.1}$$

where $\mathbf{g}_{\text{PAR}}$ represents the gradient of the PAR penalty function $J_{\text{PAR}}(\boldsymbol{\theta}_T)$ with respect to $\overline{\mathbf{y}}_{\boldsymbol{\theta}_T}$, and is given by

$$\mathbf{g}_{\text{PAR}} = \begin{bmatrix} [0, \ldots, 0, 2K\Re(y_{\boldsymbol{\theta}_T,\max}), 0, \ldots, 0]^T \\ \hdashline [0, \ldots, 0, 2K\Im(y_{\boldsymbol{\theta}_T,\max}), 0, \ldots, 0]^T \end{bmatrix}. \tag{A.2}$$

Gradient of Spectral Compatibility Penalty Function According to the chain rule, the gradient of Equation (4.4) with respect to $\boldsymbol{\theta}_T$ is expressed

$$\nabla_{\boldsymbol{\theta}_T} J_{\text{spectrum}}(\boldsymbol{\theta}_T) = \nabla_{\boldsymbol{\theta}_T} \overline{\mathbf{y}}_{\boldsymbol{\theta}_T} \cdot \mathbf{g}_{\text{spectrum}}, \tag{A.3}$$

where $\mathbf{g}_{\text{spectrum}}$ denotes the gradient of the spectral compatibility penalty function $J_{\text{spectrum}}(\boldsymbol{\theta}_T)$ with respect to $\overline{\mathbf{y}}_{\boldsymbol{\theta}_T}$, and is given by

$$\mathbf{g}_{\text{spectrum}} = \left[\begin{array}{c} 2\Re\big[(\boldsymbol{\Omega}\mathbf{y}_{\boldsymbol{\theta}_T})^*\big] \\ \hdashline -2\Im\big[(\boldsymbol{\Omega}\mathbf{y}_{\boldsymbol{\theta}_T})^*\big] \end{array}\right]. \tag{A.4}$$

APPENDIX B

PROOF OF PROPOSITION 1

This appendix provides the proof of Proposition 1.

Proof. The average loss function of simultaneous training $\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T)$ (3.10) could be expressed

$$\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{A}} \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) \int_{\mathcal{Z}} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) p(\mathbf{z}|\mathbf{a}, \mathcal{H}_i) d\mathbf{z} d\mathbf{a}. \tag{B.1}$$

As discussed in Section 2.2, the last layer of the receiver implementation consists of a sigmoid activation function, which leads to the output of the receiver $f_{\boldsymbol{\theta}_R}(\mathbf{z}) \in (0, 1)$. Thus there exists a constant b such that $\sup_{\mathbf{z},i} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) < b < \infty$. Furthermore, for $i \in \{0, 1\}$, the instantaneous values of the cross-entropy loss $\ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big)$, the policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$, and the likelihood $p(\mathbf{z}|\mathbf{a}, \mathcal{H}_i)$ are continuous in variables $\mathbf{a}$ and $\mathbf{z}$. By leveraging Fubini's theorem [53] to exchange the order of integration in (B.1), we have

$$\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) = \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{Z}} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) \int_{\mathcal{A}} p(\mathbf{z}|\mathbf{a}, \mathcal{H}_i) \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) d\mathbf{a} d\mathbf{z}. \tag{B.2}$$

Note that for a waveform $\mathbf{y}_{\boldsymbol{\theta}_T}$ and a target state indicator i, the product between the likelihood $p(\mathbf{z}|\mathbf{a}, \mathcal{H}_i)$ and the policy $\pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T})$ becomes a joint PDF of two random variables $\mathbf{a}$ and $\mathbf{z}$, namely,

$$p(\mathbf{z}|\mathbf{a}, \mathcal{H}_i) \pi(\mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}) = p(\mathbf{a}, \mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i). \tag{B.3}$$

Substituting (B.3) into (B.2), we obtain

$$\begin{aligned}\mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{Z}} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) \int_{\mathcal{A}} p(\mathbf{a}, \mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i) d\mathbf{a} d\mathbf{z} \\ &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{Z}} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i) d\mathbf{z},\end{aligned} \tag{B.4}$$

where the second equality holds by integrating the joint PDF $p(\mathbf{z}, \mathbf{a}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$ over the random variable $\mathbf{a}$, i.e., $\int_{\mathcal{A}} p(\mathbf{a}, \mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i) d\mathbf{a} = p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i)$.

Taking the gradient of (B.4) with respect to $\boldsymbol{\theta}_R$, we have

$$\begin{aligned}\nabla_{\boldsymbol{\theta}_R} \mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) &= \sum_{i\in\{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{Z}} p(\mathbf{z}|\mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i) \nabla_{\boldsymbol{\theta}_R} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) d\mathbf{z} \\ &= \nabla_{\boldsymbol{\theta}_R} \mathcal{L}_R(\boldsymbol{\theta}_R),\end{aligned} \tag{B.5}$$

where the second equality holds via Equation (3.3). Thus, the proof of Proposition 1 is completed. ■

APPENDIX C

PROOF OF PROPOSITION 2

In this appendix, the proof of Proposition 2 is provided.

Proof. According to Equation (B.4), the gradient of the average loss function of simultaneous training with respect to $\boldsymbol{\theta}_T$ is given by

$$\begin{aligned} \nabla_{\boldsymbol{\theta}_T} \mathcal{L}^{\pi}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T) &= \sum_{i \in \{0,1\}} P(\mathcal{H}_i) \int_{\mathcal{Z}} \ell\big(f_{\boldsymbol{\theta}_R}(\mathbf{z}), i\big) \nabla_{\boldsymbol{\theta}_T} p(\mathbf{z} | \mathbf{y}_{\boldsymbol{\theta}_T}, \mathcal{H}_i) d\mathbf{z} \\ &= \nabla_{\boldsymbol{\theta}_T} \mathcal{L}(\boldsymbol{\theta}_R, \boldsymbol{\theta}_T), \end{aligned} \tag{C.1}$$

where the last equality holds by Equation (5.3). The proof of Proposition 2 is completed. ■

APPENDIX D

OPTIMAL NP DETECTOR IN GAUSSIAN CLUTTER

In this appendix, the optimal detector in the NP sense is derived for Gaussian clutter, i.e., when $\beta = 2$. Denote the covariance matrices of the target response and the clutter as $\mathbf{\Omega}_s$ and $\mathbf{\Omega}_c$, respectively. With the assumptions in Section 6.1, we have covariance matrix $\mathbf{\Omega}_s$ given by

$$\begin{aligned} \mathbf{\Omega}_s &= \mathbb{E}\big[(\alpha\mathbf{y})(\alpha\mathbf{y})^H\big] \\ &= \sigma_\alpha^2 \mathbf{y}\mathbf{y}^H, \end{aligned} \tag{D.1}$$

and covariance matrix $\mathbf{\Omega}_c$ given by

$$\begin{aligned} \mathbf{\Omega}_c &= \mathbb{E}\big[\mathbf{c}\mathbf{c}^H\big] \\ &= \sum_{g=-K+1}^{K-1} \sigma_{\gamma_g}^2 \mathbf{J}_g \mathbf{y}\mathbf{y}^H \mathbf{J}_g^H. \end{aligned} \tag{D.2}$$

After applying the logarithm to Equation (2.6) and neglecting the terms that are independent of the received signal $\mathbf{z}$, the NP decision rule can be expressed as

$$\mathbf{z}^H \big[(\mathbf{\Omega}_c + \mathbf{\Omega}_n)^{-1} - (\mathbf{\Omega}_s + \mathbf{\Omega}_c + \mathbf{\Omega}_n)^{-1}\big]\mathbf{z} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T'_\Lambda, \tag{D.3}$$

where $T_\Lambda^{'}$ is a modified threshold. By leveraging the Woodbury matrix inversion lemma, we have

$$\begin{aligned}&(\mathbf{\Omega}_s+\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1}\\&=(\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1}-\kappa(\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1}\mathbf{y}\mathbf{y}^H(\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1},\end{aligned}\tag{D.4}$$

where $\kappa=\left[1/\sigma_\alpha^2+\mathbf{y}^H(\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1}\mathbf{y}\right]^{-1}$ is a positive scalar, which is independent of the received signal $\mathbf{z}$. Substituting (D.4) into (D.3), we obtain the optimal detector

$$\left|\mathbf{z}^H(\mathbf{\Omega}_c+\mathbf{\Omega}_n)^{-1}\mathbf{y}\right|^2\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\kappa^{-1}T_\Lambda^{'}.\tag{D.5}$$

Note that analytical expressions of the probability of false alarm and probability of detection are available when the clutter is Gaussian. Let $\mathbf{z}_w=\mathbf{D}\mathbf{z}$ with $\mathbf{D}=\left(\mathbf{\Omega}_c+\mathbf{\Omega}_n\right)^{-1/2}$. The binary hypothesis testing problem (2.5) could be equivalently expressed as

$$\begin{cases}\mathcal{H}_0:\mathbf{z}_w\sim\mathcal{CN}\left(\mathbf{0},\mathbf{I}_K\right)\\\mathcal{H}_1:\mathbf{z}_w\sim\mathcal{CN}\left(\mathbf{0},\mathbf{I}_K+\mathbf{D}\mathbf{\Omega}_s\mathbf{D}\right).\end{cases}\tag{D.6}$$

According to the estimator-correlator theorem [1], the test statistic for (D.6) is given by

$$T(\mathbf{z}_w)=\mathbf{z}_w^H\mathbf{D}\mathbf{\Omega}_s\mathbf{D}(\mathbf{I}_K+\mathbf{D}\mathbf{\Omega}_s\mathbf{D})^{-1}\mathbf{z}_w.\tag{D.7}$$

Let $\mathbf{V}\boldsymbol{\xi}\mathbf{V}^H$ represent the eigenvalue decomposition of $\mathbf{D}\boldsymbol{\Omega}_s\mathbf{D}$ in Equation (D.7), where $\mathbf{V}$ is a unitary matrix including column eigenvectors, and $\boldsymbol{\xi}$ is a diagonal matrix containing eigenvalues of $\mathbf{D}\boldsymbol{\Omega}_s\mathbf{D}$. Then, the test statistic (D.7) can be written as

$$T(\mathbf{z}_w) = \mathbf{z}_w^H \mathbf{V}\boldsymbol{\xi}(\mathbf{I}_K + \boldsymbol{\xi})^{-1}\mathbf{V}^H\mathbf{z}_w. \tag{D.8}$$

By defining $\mathbf{r} = \mathbf{V}^H\mathbf{z}_w$, the covariance matrix of the random vector $\mathbf{r}$ under $\mathcal{H}_0$ hypothesis is given by

$$\begin{aligned}\mathbb{E}\{\mathbf{r}\mathbf{r}^H|\mathcal{H}_0\} &= \mathbb{E}\{\mathbf{V}^H\mathbf{z}_w\mathbf{z}_w^H\mathbf{V}|\mathcal{H}_0\} \\ &= \mathbf{V}^H\mathbf{I}_K\mathbf{V} \\ &= \mathbf{I}_K,\end{aligned} \tag{D.9}$$

and the covariance matrix of the random vector $\mathbf{r}$ under $\mathcal{H}_1$ hypothesis is given by

$$\begin{aligned}\mathbb{E}\{\mathbf{r}\mathbf{r}^H|\mathcal{H}_1\} &= \mathbb{E}\{\mathbf{V}^H\mathbf{z}_w\mathbf{z}_w^H\mathbf{V}|\mathcal{H}_1\} \\ &= \mathbf{V}^H(\mathbf{I}_K + \mathbf{D}\boldsymbol{\Omega}_s\mathbf{D})\mathbf{V} \\ &= \mathbf{I}_K + \boldsymbol{\xi}.\end{aligned} \tag{D.10}$$

As indicated in Equation (D.1), the rank of the target covariance matrix $\boldsymbol{\Omega}_s$ equals one. Accordingly, we rewrite the test statistic (D.8) as

$$T(\mathbf{z}_w) = \frac{\xi}{\xi + 1}|r|^2, \tag{D.11}$$

where

$$\xi = \sigma_\alpha^2 \mathbf{y}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n)^{-1} \mathbf{y} \tag{D.12}$$

is the positive eigenvalue. From Equations (D.9) and (D.10), the binary hypothesis testing problem (D.6) can be further written as

$$\begin{cases} \mathcal{H}_0 : r \sim \mathcal{CN}\big(0, 1\big) \\ \mathcal{H}_1 : r \sim \mathcal{CN}\big(0, 1+\xi\big). \end{cases} \tag{D.13}$$

In order to characterize the statistics of $|r|^2$ in Equation (D.11), we express $|r|^2$ in terms of the real and imaginary parts of r as follows

$$\begin{cases} \mathcal{H}_0 : |r|^2 = \dfrac{1}{2}\left\{ \left[\dfrac{\Re(r)}{\sqrt{1/2}}\right]^2 + \left[\dfrac{\Im(r)}{\sqrt{1/2}}\right]^2 \right\} \\ \mathcal{H}_1 : |r|^2 = \dfrac{1+\xi}{2}\left\{ \left[\dfrac{\Re(r)}{\sqrt{(1+\xi)/2}}\right]^2 + \left[\dfrac{\Im(r)}{\sqrt{(1+\xi)/2}}\right]^2 \right\}. \end{cases} \tag{D.14}$$

Thus, we have

$$\begin{cases} \mathcal{H}_0 : 2|r|^2 \sim \mathcal{X}_2^2 \\ \mathcal{H}_1 : \dfrac{2}{1+\xi}|r|^2 \sim \mathcal{X}_2^2, \end{cases} \tag{D.15}$$

where $\mathcal{X}_2^2$ represents the chi-square distribution with 2 degrees of freedom. From Equations (D.11) and (D.15), the detection performance can by found by noting that

$$\begin{cases} \mathcal{H}_0 : \dfrac{2(1+\xi)}{\xi} T(\mathbf{z}_w) \sim \mathcal{X}_2^2 \\ \mathcal{H}_1 : \dfrac{2}{\xi} T(\mathbf{z}_w) \sim \mathcal{X}_2^2. \end{cases} \tag{D.16}$$

Given a threshold $T_\Lambda^{''}$, the probability of false alarm is therefore

$$\begin{aligned} P_{fa} &= Pr\{T(\mathbf{z}_w) > T_\Lambda^{''} | \mathcal{H}_0\} \\ &= Pr\left\{ \frac{2(1+\xi)}{\xi} T(\mathbf{z}_w) > \frac{2(1+\xi)}{\xi} T_\Lambda^{''} \middle| \mathcal{H}_0 \right\} \\ &= Q_{\mathcal{X}_2^2}\left(\frac{2(1+\xi)}{\xi} T_\Lambda^{''} \right), \end{aligned} \tag{D.17}$$

and the probability of detection is given by

$$\begin{aligned} P_d &= Pr\{T(\mathbf{z}_w) > T_\Lambda^{''} | \mathcal{H}_1\} \\ &= Pr\left\{ \frac{2}{\xi} T(\mathbf{z}_w) > \frac{2}{\xi} T_\Lambda^{''} \middle| \mathcal{H}_1 \right\} \\ &= Q_{\mathcal{X}_2^2}\left(\frac{2}{\xi} T_\Lambda^{''} \right), \end{aligned} \tag{D.18}$$

where $Q_{\mathcal{X}_2^2}(\cdot)$ denotes the right-tail probability for a chi-square random variable with 2 degrees of freedom.

From (D.17), the probability of false alarm is

$$\begin{aligned} P_{fa} &= \int_{\frac{2(1+\xi)}{\xi} T_\Lambda^{''}}^{+\infty} \frac{1}{2} \exp^{-x/2} \, dx \\ &= \exp\left\{ -\frac{1+\xi}{\xi} T_\Lambda^{''} \right\}. \end{aligned} \tag{D.19}$$

The threshold setting in terms of P_{fa} is

$$T_{\Lambda}^{''} = -\frac{\xi}{1+\xi} \ln P_{fa}. \tag{D.20}$$

From Equation (D.18), the probability of detection is

$$\begin{aligned} P_d &= \int_{\frac{2}{\xi}T_{\Lambda}^{''}}^{+\infty} \frac{1}{2} \exp^{-x/2} \, dx \\ &= \exp\left\{ -\frac{1}{\xi} T_{\Lambda}^{''} \right\}. \end{aligned} \tag{D.21}$$

Substituting (D.20) into (D.21), the analytical expression of the detection probability P_d as a function of the false alarm probability P_{fa} is given by

$$P_d = P_{fa}^{\frac{1}{1+\xi}}. \tag{D.22}$$

APPENDIX E

RADAR WAVEFORM DESIGN FOR TARGET DETECTION

This appendix reviews the radar waveform design for target detection via the sequential optimization algorithm.

Assume that the received signal $\mathbf{z}$ is filtered through $\mathbf{h}$, such that the signal-to-clutter-plus-noise ratio at the output of the filter is given by

$$\text{SCNR} = \frac{\sigma_\alpha^2 |\mathbf{h}^H \mathbf{y}|^2}{\mathbf{h}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n) \mathbf{h}}. \tag{E.1}$$

Thus, the radar waveform $\mathbf{y}$ and the receiver filter $\mathbf{h}$ can be obtained jointly by solving the following constrained optimization problem

$$\begin{aligned} \max_{\mathbf{y},\mathbf{h}} \quad & \frac{|\mathbf{h}^H \mathbf{y}|^2}{\mathbf{h}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n) \mathbf{h}} \\ \text{s.t.} \quad & \mathbf{y}^H \mathbf{y} = 1. \end{aligned} \tag{E.2}$$

Since the optimization problem (E.2) is not convex with a non-convex objective function and a constant waveform power constraint, it is difficult to obtain a globally optimal solution. We aim to find a locally optimal solution of the waveform $\mathbf{y}$ and the receiver filter $\mathbf{h}$ iteratively. Specifically, at the nth iteration, we optimize the receiver filter $\mathbf{h}^{(n)}$ given the waveform $\mathbf{y} = \mathbf{y}^{(n-1)}$, attained at the previous iteration; subsequently, the transmitted waveform is optimized with the receiver filter $\mathbf{h}^{(n)}$ obtained at the previous step. The details of the iterative algorithm are presented next.

Given the waveform $\mathbf{y} = \mathbf{y}^{(n-1)}$, we aim at finding the optimal value of the receiver filter $\mathbf{h}^{(n)}$ for the optimization problem (E.2). At the nth iteration, we evaluate the receiver filter $\mathbf{h}^{(n)}$ by solving the following unconstrained optimization problem

$$\max_{\mathbf{h}} \quad \frac{|\mathbf{h}^H \mathbf{y}|^2}{\mathbf{h}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n)\mathbf{h}}. \tag{E.3}$$

As discussed in [54], the problem (E.3) may be recast as

$$\begin{aligned} \min_{\mathbf{h}} \quad & \mathbf{h}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n)\mathbf{h} \\ \text{s.t.} \quad & \mathbf{h}^H \mathbf{y} = 1. \end{aligned} \tag{E.4}$$

Note that the solution to (E.4) is well-known [55], and is given by

$$\mathbf{h}_* = \mu_{\mathbf{y}} (\mathbf{\Omega}_c + \mathbf{\Omega}_n)^{-1} \mathbf{y}, \tag{E.5}$$

where $\mu_{\mathbf{y}} = [\mathbf{y}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n)^{-1} \mathbf{y}]^{-1}$ is a constant, which is used to ensure the equality constraint in (E.4). Thus, the optimal value of the receiver filter at the nth iteration is $\mathbf{h}^{(n)} = \mathbf{h}_*$.

Next we aim to optimize the waveform $\mathbf{y}^{(n)}$ given the receiver filter $\mathbf{h} = \mathbf{h}^{(n)}$. From Equation (E.2), the waveform $\mathbf{y}^{(n)}$ is evaluated by solving the following problem

$$\begin{aligned} \max_{\mathbf{y}} \quad & \frac{|\mathbf{h}^H \mathbf{y}|^2}{\mathbf{h}^H (\mathbf{\Omega}_c + \mathbf{\Omega}_n)\mathbf{h}} \\ \text{s.t.} \quad & \mathbf{y}^H \mathbf{y} = 1. \end{aligned} \tag{E.6}$$

Plugging Equation (D.2) into the objective function of (E.6), we have

$$\begin{aligned}\frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{h}^H(\mathbf{\Omega}_c+\mathbf{\Omega}_n)\mathbf{h}} &= \frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{h}^H\big(\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{J}_g\mathbf{y}\mathbf{y}^H\mathbf{J}_g^H+\mathbf{\Omega}_n\big)\mathbf{h}} \\ &= \frac{|\mathbf{h}^H\mathbf{y}|^2}{\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{h}^H\mathbf{J}_g\mathbf{y}\mathbf{y}^H\mathbf{J}_g^H\mathbf{h}+\mathbf{h}^H\mathbf{\Omega}_n\mathbf{h}\cdot 1}.\end{aligned} \tag{E.7}$$

Substituting the waveform power constraint into (E.7), the objective function of (E.6) could be further rewritten

$$\begin{aligned}\frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{h}^H(\mathbf{\Omega}_c+\mathbf{\Omega}_n)\mathbf{h}} &= \frac{|\mathbf{h}^H\mathbf{y}|^2}{\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{h}^H\mathbf{J}_g\mathbf{y}\mathbf{y}^H\mathbf{J}_g^H\mathbf{h}+\mathbf{h}^H\mathbf{\Omega}_n\mathbf{h}\cdot\mathbf{y}^H\mathbf{y}} \\ &= \frac{|\mathbf{h}^H\mathbf{y}|^2}{\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{y}^H\mathbf{J}_g^H\mathbf{h}\mathbf{h}^H\mathbf{J}_g\mathbf{y}+\mathbf{y}^H(\mathbf{h}^H\mathbf{\Omega}_n\mathbf{h}\cdot\mathbf{I}_K)\mathbf{y}} \\ &= \frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{y}^H\bigg(\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{J}_g^H\mathbf{h}\mathbf{h}^H\mathbf{J}_g+\mathbf{h}^H\mathbf{\Omega}_n\mathbf{h}\cdot\mathbf{I}_K\bigg)\mathbf{y}}\end{aligned} \tag{E.8}$$

For notational simplicity, let

$$\mathbf{\Omega}_{c,h}=\sum_{g=-K+1}^{K-1}\sigma_{\gamma_g}^2\mathbf{J}_g^H\mathbf{h}\mathbf{h}^H\mathbf{J}_g, \tag{E.9}$$

and

$$\mathbf{\Omega}_{n,h}=\mathbf{h}^H\mathbf{\Omega}_n\mathbf{h}\cdot\mathbf{I}_K. \tag{E.10}$$

Based on (E.8)-(E.10), the optimization problem (E.6) becomes

$$\begin{aligned} \max_{\mathbf{y}} \quad & \frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{y}^H(\mathbf{\Omega}_{c,h}+\mathbf{\Omega}_{n,h})\mathbf{y}} \\ \text{s.t.} \quad & \mathbf{y}^H\mathbf{y}=1. \end{aligned} \tag{E.11}$$

According to the Proposition 1 in [54], the above optimization problem could be relaxed to the following unconstrained problem

$$\max_{\mathbf{y}} \quad \frac{|\mathbf{h}^H\mathbf{y}|^2}{\mathbf{y}^H(\mathbf{\Omega}_{c,h}+\mathbf{\Omega}_{n,h})\mathbf{y}}. \tag{E.12}$$

Similar to the problem (E.3), the problem (E.12) could be recast as

$$\begin{aligned} \min_{\mathbf{y}} \quad & \mathbf{y}^H(\mathbf{\Omega}_{c,h}+\mathbf{\Omega}_{n,h})\mathbf{y} \\ \text{s.t.} \quad & \mathbf{h}^H\mathbf{y}=1. \end{aligned} \tag{E.13}$$

The solution to (E.13) is well-known [55], and is given by

$$\mathbf{y}_* = \mu_{\mathbf{h}}(\mathbf{\Omega}_{c,h}+\mathbf{\Omega}_{n,h})^{-1}\mathbf{h}, \tag{E.14}$$

where $\mu_{\mathbf{h}} = [\mathbf{h}^H(\mathbf{\Omega}_{c,h}+\mathbf{\Omega}_{n,h})^{-1}\mathbf{h}]^{-1}$ is a constant, which is used to ensure the constraint in (E.13). Then, the optimal value of the waveform at the nth iteration is obtained via $\mathbf{y}^{(n)} = \mathbf{y}_*/||\mathbf{y}_*||$.

The optimization process is carried out until the improvement of the signal-to-clutter-plus-noise ratio level (E.1) becomes insignificant. The design of radar waveform for target detection is summarized in Algorithm 3.

Algorithm 3: Waveform Optimization for Target Detection

Input: noise covariance matrix $\mathbf{\Omega}_n$; clutter power $\{\sigma^2_{\gamma_g}\}_{g=-K+1}^{K-1}$

Output: optimal waveform $\mathbf{y}$

initialize waveform $\mathbf{y}^{(0)}$ and receiver filter $\mathbf{h}^{(0)}$, and set $n = 0$

while *stopping criterion not satisfied* **do**

  calculate $\mathbf{\Omega}_c$ (D.2) with $\mathbf{y} = \mathbf{y}^{(n-1)}$

  find the receiver filter $\mathbf{h}^{(n)} = \mathbf{h}_*$ via (E.5)

  calculate $\mathbf{\Omega}_{c,h}$ (E.9) and $\mathbf{\Omega}_{n,h}$ (E.10) with $\mathbf{h} = \mathbf{h}^{(n)}$

  find $\mathbf{y}_*$ via (E.14)

  update the waveform $\mathbf{y}^{(n)} = \mathbf{y}_*/||\mathbf{y}_*||$

  $n \leftarrow n + 1$

end

REFERENCES

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory.* New York, NY, USA: Pearson, 1998.

[2] D. F. Delong and E. M. Hofstetter, "On the design of optimum radar waveforms for clutter rejection," *IEEE Transactions on Information Theory*, vol. 13, no. 3, pp. 454–463, 1967.

[3] P. Stoica, H. He, and J. Li, "Optimization of the receive filter and transmit sequence for active sensing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1730–1740, 2012.

[4] S. M. Kay, "Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 31–41, 2007.

[5] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar.* Edison, NJ, USA: Scitech Publishing, 2010.

[6] F. Gini, A. De Maio, and L. Patton, *Waveform Design and Diversity for Advanced Radar Systems.* London, UK: Institution of Engineering and Technology, 2012.

[7] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation and Linear Modulation Theory.* New York, NY, USA: John Wiley & Sons, 2004.

[8] D. P. Meyer and H. A. Mayer, *Radar Target Detection.* New York, NY, USA: Academic Press, 1973.

[9] M. A. Richards, *Fundamentals of Radar Signal processing.* New York, NY, USA: McGraw-Hill, 2005.

[10] K. J. Sangston and K. R. Gerlach, "Coherent detection of radar targets in a non-Gaussian background," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 2, pp. 330–340, 1994.

[11] F. Gini, "Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter," *IEE Proceedings Radar, Sonar and Navigation*, vol. 144, no. 1, pp. 39–48, 1997.

[12] K. J. Sangston, F. Gini, M. V. Greco, and A. Farina, "Structures for radar detection in compound Gaussian clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 445–458, 1999.

[13] A. Farina, A. Russo, and F. Studer, "Coherent radar detection in log-normal clutter," *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 133, no. 1, pp. 39–53, 1986.

[14] F. A. Pentini, A. Farina, and F. Zirilli, "Radar detection of targets located in a coherent K distributed clutter background," *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, no. 3, pp. 239–245, 1992.

[15] S. Kay, "Waveform design for multistatic radar detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1153–1166, 2009.

[16] M. M. Naghsh, M. Modarres-Hashemi, S. ShahbazPanahi, M. Soltanalian, and P. Stoica, "Unified optimization framework for multi-static radar code design using information-theoretic criteria," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5401–5416, 2013.

[17] S. Khalili, O. Simeone, and A. M. Haimovich, "Cloud radio-multistatic radar: joint optimization of code vector and backhaul quantization," *IEEE Signal Processing Letters*, vol. 22, no. 4, pp. 494–498, 2014.

[18] S. Jeong, S. Khalili, O. Simeone, A. Haimovich, and J. Kang, "Multistatic cloud radar systems: joint sensing and communication design," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 5, pp. 716–730, 2016.

[19] A. De Maio, Y. Huang, M. Piezzo, S. Zhang, and A. Farina, "Design of optimized radar codes with a peak to average power ratio constraint," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2683–2697, 2011.

[20] B. Tang, M. M. Naghsh, and J. Tang, "Relative entropy-based waveform design for MIMO radar detection in the presence of clutter and interference," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3783–3796, 2015.

[21] M. M. Naghsh, M. Modarres-Hashemi, M. A. Kerahroodi, and E. H. M. Alian, "An information theoretic approach to robust constrained code design for MIMO radars," *IEEE Transactions on Signal Processing*, vol. 65, no. 14, pp. 3647–3661, 2017.

[22] L. Wu, P. Babu, and D. P. Palomar, "Transmit waveform/receive filter design for MIMO radar with multiple waveform constraints," *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1526–1540, 2017.

[23] G. Locke, L. E. Strickling, and A. Secretary, "An assessment of the near-term viability of accommodating wireless broadband systems in the 1675-1710 MHz, 1755-1780 MHz, 3500-3650 MHz, and 4200-4220 MHz, 4380-4400 MHz bands (fast track evaluation report)," pp. 3500–3650, 2010.

[24] A. Aubry, V. Carotenuto, A. De Maio, A. Farina, and L. Pallotta, "Optimization theory-based radar waveform design for spectrally dense environments," *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 12, pp. 14–25, 2016.

[25] W. Jiang and A. M. Haimovich, "Joint optimization of waveform and quantization in spectral congestion conditions," in *IEEE 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 1894–1898, 2018.

[26] L. Zheng, M. Lops, X. Wang, and E. Grossi, "Joint design of overlaid communication systems and pulsed radars," *IEEE Transactions on Signal Processing*, vol. 66, no. 1, pp. 139–154, 2017.

[27] W. Jiang and A. M. Haimovich, "Waveform optimization in cloud radar with spectral congestion constraints," in *IEEE Radar Conference*, pp. 1–6, 2019.

[28] B. Tang and J. Li, "Spectrally constrained MIMO radar waveform design based on mutual information," *IEEE Transactions on Signal Processing*, vol. 67, no. 3, pp. 821–834, 2018.

[29] N. Sebe, I. Cohen, A. Garg, and T. S. Huang, *Machine Learning in Computer Vision*, vol. 29. Dordrecht, Netherlands: Springer, 2005.

[30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: a large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.

[31] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[32] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[33] T. O'shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[34] M. Kim, W. Lee, and D.-H. Cho, "A novel PAPR reduction scheme for OFDM system based on deep learning," *IEEE Communications Letters*, vol. 22, no. 3, pp. 510–513, 2017.

[35] F. A. Aoudia and J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2503–2516, 2019.

[36] V. Raj and S. Kalyani, "Backpropagating through the air: deep learning at physical layer without channel models," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2278–2281, 2018.

[37] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, 2018.

[38] O. Simeone, S. Park, and J. Kang, "From learning to meta-learning: reduced training overhead and complexity for communication systems," in *IEEE 2nd 6G Wireless Summit*, pp. 1–5, 2020.

[39] S. Park, O. Simeone, and J. Kang, "Meta-learning to communicate: fast end-to-end training for fading channels," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5075–5079, 2020.

[40] S. Park, O. Simeone, and J. Kang, "End-to-end fast training of communication links without a channel model via online meta-learning," in *IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications*, pp. 1–5, 2020.

[41] M.-P. Jarabo-Amores, M. Rosa-Zurera, R. Gil-Pita, and F. Lopez-Ferreras, "Study of two error functions to approximate the Neyman–Pearson detector using supervised learning machines," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4175–4181, 2009.

[42] M.-P. Jarabo-Amores, D. De la Mata-Moya, R. Gil-Pita, and M. Rosa-Zurera, "Radar detection with the Neyman–Pearson criterion using supervised-learning-machines trained with the cross-entropy error," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 1–10, 2013.

[43] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*. Cambridge, Massachusetts, USA: MIT press, 2016.

[44] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[45] R. S. Sutton, D. A. McAllester, S. P. Singh, Y. Mansour, *et al.*, "Policy gradient methods for reinforcement learning with function approximation.," in *Conference on Neural Information Processing Systems*, pp. 1057–1063, 1999.

[46] A. Aubry, A. De Maio, Y. Huang, M. Piezzo, and A. Farina, "A new radar waveform design algorithm with improved feasibility for spectral coexistence," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 1029–1038, 2015.

[47] A. Farina, A. Russo, F. Scannapieco, and S. Barbarossa, "Theory of radar detection in coherent Weibull clutter," *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 134, no. 2, pp. 174–190, 1987.

[48] D. A. Shnidman, "Generalized radar clutter model," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 857–865, 1999.

[49] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," *arXiv preprint arXiv:1511.07289*, 2015.

[50] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International Conference on Machine Learning*, pp. 387–395, 2014.

[51] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[52] O. Simeone, "A brief introduction to machine learning for engineers," *Foundations Trends® in Signal Processing*, vol. 12, no. 3-4, pp. 200–431, 2017.

[53] W. Rudin, *Real and Complex Analysis*. New York, NY, USA: McGraw-Hill, 1987.

[54] C. Y. Chen and P. Vaidyanathan, "MIMO radar waveform optimization with prior information of the extended target and clutter," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3533–3544, 2009.

[55] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.