1

Information-Theoretic Waveform Design for MIMO Radar Target Detection

Bo Tang[1], Jun Tang[2], and Petre Stoica[3]

[1] *College of Electronic Engineering, National University of Defense Technology, Hefei, China*
[2] *Department of Electronic Engineering, Tsinghua University, Beijing, China*
[3] *Department of Information Technology, Uppsala University, Uppsala, Sweden*

1.1 Introduction

Information theory is a fundamental tool to study how to communicate reliably in noisy channels [1]. The analysis of the ultimate data compression rate and the highest data transmission rate in communications is a main focus of information theory [2]. After nearly one century of research, many elegant results about the data compression and transmission rate have been developed in information theory. These studies have achieved tremendous success in many areas, including but not limited to communication system design, computer science, statistics, and stock market investment.

An early attempt at using information theory as a guideline for designing radar systems dates back to Woodward's work in the 1950s [3–5]. These cited references pointed out that a well-established and easy-to-understand concept in the radar community, viz., the signal-to-interference-plus-noise ratio (SINR), did not measure information except for the Gaussian case and certain special elliptical contoured distributions [6]. Thus, an "information function" was defined, which was related to the posterior distributions of the observations in the radar receiver. In addition, a posterior receive filter was proposed based on the information function. Besides the receiver design based on information-theoretic approaches, there have been even more attempts to design radar transmit waveforms based on information theory. In [7], the author considered the design of radar waveforms to maximize the mutual information between the target response and the received signals. By connecting the mutual information to the rate-distortion function, the study showed that the mutual information maximization resulted in the

Information-Theoretic Radar Signal Processing, First Edition. Edited by Yujie Gu and Yimin D. Zhang.

reduction of the parameter estimation uncertainty. Therefore, it is meaningful to maximize the mutual information through waveform design. The results in [7] showed that the optimal waveforms maximizing the mutual information had a "water-filling" interpretation. In [8, 9], the authors investigated the design of clutter-resistant radar waveforms. The results therein also showed that the optimal waveforms admitted a "water-filling"-like solution. Note that only the spectral distribution of the radar waveform was derived in [7, 9]. In real-world radar systems, we need to synthesize radar waveforms under practical constraints. In [10, 11], several iterative algorithms, which leveraged fast Fourier transform, were proposed to synthesize waveforms with arbitrary spectral shape under some practical constraints, including the constant-modulus constraint, the low peak-to-average-power ratio (PAPR) constraint, and the similarity constraint. Different from the studies in [7, 9], which focused on the design of continuous-time waveforms, the authors of [12] considered the design of discrete-time waveforms to maximize the mutual information. Moreover, they showed the connection between the waveform maximizing mutual information and that minimizing minimum mean square error (MMSE).

Intuitively, the mutual information between the target response and the received signals stands for the amount of useful information that can be extracted from the received signals. Therefore, the waveforms designed based on mutual information maximization are beneficial for target parameter estimation as well as classification. However, the waveforms maximizing the mutual information might not have a satisfactory detection performance. To improve the target detection performance via designing waveforms, information-theoretic approaches have also been investigated [13–15]. Indeed, Stein's lemma [2] states that, for a fixed probability of false alarm, the probability of miss detection asymptotically decays with an exponential rate equal to the relative entropy (aka Kullback–Leibler divergence) between the distributions of the observations under the two hypotheses (viz., the target is present/absent). Therefore, maximizing the relative entropy results in minimizing the probability of miss detection. In [13], the author studied the problem of detecting point-like targets in colored noise. He proved that the performance of Neyman–Pearson detector is determined by the relative entropy. Moreover, the optimal waveform based on relative entropy maximization was derived. Interestingly, the optimal waveforms also had a "water-filling" interpretation. In [15], the authors considered the detection of weak extended targets based on an information-theoretic approach. Different from the results in [13], the authors showed that the optimal waveforms achieving the maximum relative entropy place all the transmit energy into the frequency bin with the highest signal-to-noise-ratio (SNR) (a similar observation is also presented in [16]).

Note that the above works focus on information-theoretic methods for designing single-input single-output (SISO) radar systems. Unlike SISO radar systems, multiple-input multiple-output (MIMO) radar employs a multi-antenna array for transmitting diverse waveforms and receiving returned signals, showing significantly improved target detection and parameter estimation performance [17–19]. This improvement is partly attributable to the waveform diversity provided by MIMO radar. Therefore, waveform design for MIMO radar is an important topic, and there are many studies devoted to designing MIMO radar waveforms based on information theory [20–24]. In [20], the authors considered waveform design based on maximizing the mutual information for MIMO radar in the white noise case. It is shown that the optimal waveforms that maximize the mutual information also achieve the lowest MMSE. In [21], the authors used an assumption different from that in [20]. By assuming that the target response is uncorrelated (i.e., the target covariance matrix is a scaled identity matrix) and the receiver noise is colored, it is proved that the optimal waveforms maximizing the mutual information are identical to those achieving the Chernoff bound.

In this chapter, we consider information-theoretic approaches for MIMO radar waveform design. Our purpose is to boost the detection performance of MIMO radar in the presence of external disturbance. The external disturbance can be due to hostile jamming (which is independent of the transmit waveforms) or clutter (which is dependent on the waveforms). Considering that the detection probability associated with the detector has a complicated expression, we resort to the relative entropy between the distributions of the observations under the two hypotheses as the waveform design metric. We present an algebraic solution to the waveform design problem for a distributed MIMO radar (namely, the radar antennas are widely separated) in the absence of clutter, as well as a computational approach to design MIMO radar waveforms for more general cases. The performance of the proposed solutions is demonstrated through numerical simulations.

The remaining parts of this chapter are structured as follows. Section 1.2 introduces the signal model and formulates the waveform design problem. Section 1.3 investigates the design of waveforms for a distributed MIMO radar in the absence of clutter. Section 1.4 presents an iterative algorithm to design waveforms for MIMO radar (which can be either distributed or colocated) in the range-spread clutter case. Section 1.5 provides numerical examples to assess the performance of the synthesized MIMO radar waveforms. Finally, Section 1.6 concludes this chapter.

Notations: See Table 1.1.

Table 1.1 List of notations.

Symbol	Meaning
$\mathbb{R}$, $\mathbb{C}$	Domain of the real and complex numbers
$\mathbb{R}^{M\times N}(\mathbb{C}^{M\times N})$	$M \times N$ real (complex)-valued matrix
$\boldsymbol{X}$, $\boldsymbol{x}$, and x	Matrix, vector, and scalar
$\boldsymbol{X}(:, m)$	The mth column of $\boldsymbol{X}$
$\boldsymbol{I}$	Identity matrix
$\mathbf{0}$	Matrix of zeros
$\mathrm{Diag}(\boldsymbol{x})$	The diagonal matrix formed by $\boldsymbol{x}$
$\mathrm{BlkDiag}(\boldsymbol{A}; \boldsymbol{B})$	The block diagonal matrix formed by $\boldsymbol{A}$ and $\boldsymbol{B}$
$(\cdot)^*$, $(\cdot)^\top$, $(\cdot)^\dagger$	Conjugate, transpose, and conjugate transpose
$(\cdot)^{1/2}$	Square root of a positive semi-definite matrix
$\det(\cdot)$	Determinant of a square matrix
$\mathrm{tr}(\cdot)$	Trace of a square matrix
$\lvert\cdot\rvert$	Magnitude
$\lVert\cdot\rVert_2$	Euclidean norm
$\lVert\cdot\rVert_\infty$	Infinity norm
$\lVert\cdot\rVert_F$	Frobenius norm
$\mathbb{E}\{x\}$	Expectation of the random variable x
$\mathrm{vec}(\cdot)$	Vectorization
$\otimes$	Kronecker product
$\arg(\cdot)$, $\mathrm{Re}(\cdot)$, $\mathrm{Im}(\cdot)$	The argument, real and imaginary part of a complex number
$\mathcal{CN}(\boldsymbol{m}, \boldsymbol{R})$	Complex Gaussian distribution with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{R}$
$\boldsymbol{A} \succ \mathbf{0}$ ($\boldsymbol{A} \succeq \mathbf{0}$)	$\boldsymbol{A}$ is positive definite (semi-definite)
$\lambda_{\min}(\boldsymbol{X})$	The smallest eigenvalue of $\boldsymbol{X}$
$\lambda_{\max}(\boldsymbol{X})$	The largest eigenvalue of $\boldsymbol{X}$
$(x)^+$	$\max(x, 0)$

1.2 Signal Model and Problem Formulation

1.2.1 Signal Model

Consider a MIMO radar with N_T transmit antennas and N_R receive antennas. Let $\boldsymbol{S} = [\boldsymbol{s}_1, \dots, \boldsymbol{s}_{N_T}] \in \mathbb{C}^{L\times N_T}$ denote the transmit waveform matrix, where $\boldsymbol{s}_m = [s_m(1), s_m(2), \dots, s_m(L)]^\top \in \mathbb{C}^L$ is the (discrete-time) transmit waveform of the mth transmitter, $m = 1, \dots, N_T$, and L is the code length. By assuming that the Doppler shift is negligible, the complex baseband signal of the target return received by the radar system can be written as [25–28]:

$$\boldsymbol{S}\boldsymbol{H}, \tag{1.1}$$

where $\boldsymbol{H} \in \mathbb{C}^{N_T \times N_R}$ denotes the target response matrix, whose (m, n)th element, denoted as h_{mn}, represents the response from the mth transmit antenna to the nth receive antenna. The expression and the statistical properties of $\boldsymbol{H}$ are determined by the array configuration. For a colocated MIMO radar in the presence of D targets, the target response matrix can be described as $\boldsymbol{H} = \sum_{d=1}^{D} a_{t,d}\boldsymbol{a}_T(\theta_{t,d})\boldsymbol{a}_R^\top(\theta_{t,d})$, where $a_{t,d} \in \mathbb{C}$ is the amplitude of the dth target signal, $\theta_{t,d}$ denotes the associated direction-of-arrival (DOA), and $\boldsymbol{a}_T(\theta) \in \mathbb{C}^{N_T}$ and $\boldsymbol{a}_R(\theta) \in \mathbb{C}^{N_R}$ are the transmit and the receive array steering vectors at θ. For a distributed MIMO radar, it is usually assumed (see, e.g., [23–25, 29]) that the columns of $\boldsymbol{H}$ are independent and identically distributed, such that the covariance matrix of $\boldsymbol{h} = \mathrm{vec}(\boldsymbol{H})$ is given by $\boldsymbol{I} \otimes \bar{\boldsymbol{R}}_h$, where $\bar{\boldsymbol{R}}_h$ is the covariance matrix of each column of $\boldsymbol{H}$.

In radar systems, the target return is usually contaminated by clutter as well as disturbances from inside and outside the receiver. The clutter refers to unwanted returns, e.g., from ground, sea, and clouds, and it might be in the same range cell as the target or occupy the neighboring cells, as illustrated in Figure 1.1. By taking the clutter and disturbance into consideration, the received signals of a MIMO radar can be modeled by:

$$\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{H} + \sum_{p=-P}^{P} \boldsymbol{J}_p\boldsymbol{S}\boldsymbol{C}_p + \boldsymbol{N}, \tag{1.2}$$

where $\boldsymbol{J}_p \in \mathbb{R}^{L \times L}$ is the shift matrix, whose (i, j)th element equals 1 if $i - j + p = 0$, and equals 0 otherwise. $\boldsymbol{C}_p \in \mathbb{C}^{N_T \times N_R}$ denotes the response matrix of the clutter

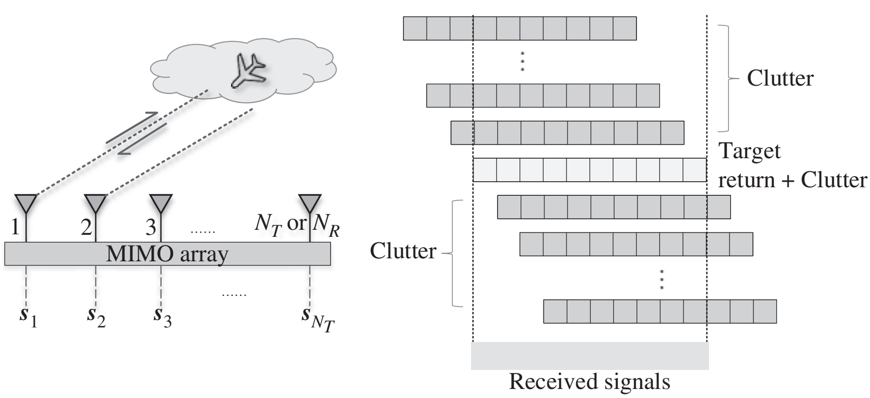

Figure 1.1 MIMO radar target detection in clutter.

in the pth range cell (note that the number of clutter range cells adjacent to the target range cell is $N_c = 2P + 1$), and $\boldsymbol{N} \in \mathbb{C}^{L\times N_R}$ is the (signal-independent) interference (including thermal noise inside the receiver and disturbances outside the receiver).

To facilitate the algebraic derivations in the sequel, we define $\boldsymbol{y} = \mathrm{vec}(\boldsymbol{Y}) \in \mathbb{C}^{LN_R}$, which can be written as

$$\boldsymbol{y} = \widetilde{\boldsymbol{S}}\boldsymbol{h} + \sum_{p=-P}^{P} \widetilde{\boldsymbol{S}}_p\boldsymbol{c}_p + \boldsymbol{n}, \tag{1.3}$$

where $\widetilde{\boldsymbol{S}} = \boldsymbol{I}_{N_R} \otimes \boldsymbol{S}$, $\widetilde{\boldsymbol{S}}_p = \boldsymbol{I}_{N_R} \otimes \boldsymbol{S}_p$, $\boldsymbol{S}_p = \boldsymbol{J}_p\boldsymbol{S}$, $\boldsymbol{c}_p = \mathrm{vec}(\boldsymbol{C}_p)$, and $\boldsymbol{n} = \mathrm{vec}(\boldsymbol{N})$.

Remark 1.1 If the Doppler shift of the signals is taken into consideration, the signal model will be more complicated. Take a colocated MIMO radar as a case study. Denote the Doppler frequency of the dth target by $f_{t,d}$. Then the target return is given by

$$\boldsymbol{Y}_t = \sum_{d=1}^{D} a_{t,d}\, \mathrm{Diag}(\boldsymbol{f}_{t,d})\boldsymbol{S}\boldsymbol{a}_T(\theta_{t,d})\boldsymbol{a}_R^\top(\theta_{t,d}), \tag{1.4}$$

where $\boldsymbol{f}_{t,d} = [1, \exp(\jmath 2\pi f_{t,d}T_s), \dots, \exp(\jmath 2\pi(L-1)f_{t,d}T_s)]^\top \in \mathbb{C}^L$, and T_s is the sampling interval. Similarly, the clutter can be modeled by

$$\boldsymbol{Y}_c = \sum_{p=-P}^{P}\sum_{m_p=1}^{M_p} a_{c,p,m_p}\boldsymbol{J}_p\, \mathrm{Diag}(\boldsymbol{f}_{c,p,m_p})\boldsymbol{S}\boldsymbol{a}_T(\theta_{c,p,m_p})\boldsymbol{a}_R^\top(\theta_{c,p,m_p}), \tag{1.5}$$

where $\boldsymbol{f}_{c,p,m_p} = [1, \exp(\jmath 2\pi f_{c,p,m_p}T_s), \dots, \exp(\jmath 2\pi(L-1)f_{c,p,m_p}T_s)]^\top \in \mathbb{C}^L$, a_{c,p,m_p}, f_{c,p,m_p}, and θ_{c,p,m_p} are, respectively, the amplitude, the Doppler frequency, and the DOA associated with the m_pth clutter patch in the pth range cell, and M_p is the number of clutter patches in the pth range cell. Note that if $f_{t,d}T_s \approx 0$, and $f_{c,p,m_p}T_s \approx 0$, $\mathrm{Diag}(\boldsymbol{f}_{t,d}) \approx \mathrm{Diag}(\boldsymbol{f}_{c,p,m_p}) \approx \boldsymbol{I}_L$, the target and the clutter models degrade to those that are valid in the absence of Doppler shift (see, e.g., (1.2)).

1.2.2 Problem Formulation

To determine whether a target is present in the cell under test (CUT), we consider the following binary hypothesis testing problem:

$$\begin{cases} \mathcal{H}_0\colon\ \boldsymbol{y} = \sum\limits_{p=-P}^{P} \widetilde{\boldsymbol{S}}_p\boldsymbol{c}_p + \boldsymbol{n}, \\ \mathcal{H}_1\colon\ \boldsymbol{y} = \widetilde{\boldsymbol{S}}\boldsymbol{h} + \sum\limits_{p=-P}^{P} \widetilde{\boldsymbol{S}}_p\boldsymbol{c}_p + \boldsymbol{n}. \end{cases} \tag{1.6}$$

Assume that $\boldsymbol{h} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_h)$, $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_n)$, and $\{\boldsymbol{c}_p \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_{c,p})\}_{p=-P}^{P}$. Moreover, the target response vector (i.e., $\boldsymbol{h}$) and the clutter response vectors (i.e., $\boldsymbol{c}_p, p = -P, \dots, P$) are independently distributed. Given these assumptions, the probability density functions (PDFs) of $\boldsymbol{y}$ under the two hypotheses are, respectively, given by

$$P_0(\boldsymbol{y}) = \frac{1}{\pi^{LN_R} \det(\boldsymbol{R}_0)} \exp(-\boldsymbol{y}^\dagger \boldsymbol{R}_0^{-1} \boldsymbol{y})$$

and

$$P_1(\boldsymbol{y}) = \frac{1}{\pi^{LN_R} \det(\boldsymbol{R}_1)} \exp(-\boldsymbol{y}^\dagger \boldsymbol{R}_1^{-1} \boldsymbol{y}),$$

where $\boldsymbol{R}_0 = \sum_{p=-P}^{P} \widetilde{\boldsymbol{S}}_p \boldsymbol{R}_{c,p} \widetilde{\boldsymbol{S}}_p^\dagger + \boldsymbol{R}_n$ and $\boldsymbol{R}_1 = \boldsymbol{R}_0 + \widetilde{\boldsymbol{S}} \boldsymbol{R}_h \widetilde{\boldsymbol{S}}^\dagger$.

Using the Neyman–Pearson criterion [30], one can verify that the optimal detector for (1.6) is given by

$$\boldsymbol{y}^\dagger (\boldsymbol{R}_0^{-1} - \boldsymbol{R}_1^{-1}) \boldsymbol{y} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T_h, \tag{1.7}$$

where T_h is the detection threshold. Moreover, the test statistic under either hypothesis is a weighted sum of chi-squared random variables (which follows from Lemma 8 in [25]). However, the probability of detection has a very complicated expression. Consequently, an information-theoretic approach is employed herein to design the waveforms, namely, the waveforms are designed to maximize the relative entropy between $P_0(\boldsymbol{y})$ and $P_1(\boldsymbol{y})$.

The relative entropy between $P_0(\boldsymbol{y})$ and $P_1(\boldsymbol{y})$ is given by

$$\begin{aligned} D(P_0 \| P_1) &= \int P_0(\boldsymbol{y}) \log \left(\frac{P_0(\boldsymbol{y})}{P_1(\boldsymbol{y})} \right) d\boldsymbol{y} \\ &= \log\det(\boldsymbol{R}_0^{-1} \boldsymbol{R}_1) + \operatorname{tr}(\boldsymbol{R}_1^{-1} \boldsymbol{R}_0) - LN_R. \end{aligned} \tag{1.8}$$

Thus, the waveform design problem based on relative entropy maximization can be stated as:

$$\begin{aligned} &\max_{\boldsymbol{S}} f(\boldsymbol{S}) = \log\det(\boldsymbol{R}_0^{-1} \boldsymbol{R}_1) + \operatorname{tr}(\boldsymbol{R}_1^{-1} \boldsymbol{R}_0) \\ &\quad \text{s.t.}\, \boldsymbol{S} \in \mathcal{S}, \end{aligned} \tag{1.9}$$

where $\mathcal{S}$ represents the constraint set of the transmit waveforms. The following constraints are considered in this chapter.

1. *Energy constraint*: Considering that the energy of the transmit waveforms in practical radar systems is limited, we impose the following constraint:

$$\operatorname{tr}(\boldsymbol{S}\boldsymbol{S}^\dagger) = e_t, \tag{1.10}$$

where e_t is the transmit energy.

2. *Low PAPR constraint*: To facilitate the radio frequency amplifier to operate in the saturated mode and reduce unnecessary distortions, low PAPR waveforms are preferable. To make the transmit energy across the transmit antennas uniformly distributed and control the waveform PAPR, we enforce the constraint below

$$\boldsymbol{s}_m^\dagger \boldsymbol{s}_m = e_t/N_T, \mathrm{PAPR}(\boldsymbol{s}_m) \leq \rho, m = 1, \dots, N_T, \tag{1.11}$$

where $1 \leq \rho \leq L$, and $\mathrm{PAPR}(\boldsymbol{s}_m) = \frac{\max_l |s_m(l)|^2}{\frac{1}{L}\sum_{l=1}^{L} |s_m(l)|^2}, m = 1, \dots, N_T$. Particularly, if $\rho = 1$, the PAPR constraint becomes the constant-modulus constraint:

$$|s_m(l)| = a_s, m = 1, \dots, N_T, l = 1, \dots, L, \tag{1.12}$$

where $a_s = \sqrt{e_t/(LN_T)}$.

3. *Similarity constraint*: To achieve desired properties like some existing waveforms, a similarity constraint can be enforced on the waveforms:

$$\mathrm{tr}(\boldsymbol{S}\boldsymbol{S}^\dagger) = e_t, \|\boldsymbol{S} - \boldsymbol{S}_0\|_F^2 \leq \epsilon, \tag{1.13}$$

where $\boldsymbol{S}_0 \in \mathbb{C}^{L \times N_T}$ stands for the matrix of reference waveforms with certain favorable properties, $\mathrm{tr}(\boldsymbol{S}_0\boldsymbol{S}_0^\dagger) = e_t$, $\epsilon \in [0, 2e_t]$ is the user-specified parameter determining the similarity between the designed waveforms and the reference waveforms.

4. *Constant-modulus and similarity (CMS) constraints*: To ensure that the designed waveforms have a constant modulus and exhibit certain desirable properties, we consider the following constraint:

$$|s(l)| = a_s, \|\boldsymbol{s} - \boldsymbol{s}_0\|_\infty \leq \epsilon_\infty, l = 1, \dots, LN_T, \tag{1.14}$$

where $s(l)$ represents the lth element of $\boldsymbol{s} = \mathrm{vec}(\boldsymbol{S})$, $\boldsymbol{s}_0 = \mathrm{vec}(\boldsymbol{S}_0)$, $\boldsymbol{S}_0$ denotes a reference waveform matrix (the modulus of each element of $\boldsymbol{S}_0$ is equal to a_s), and $\epsilon_\infty \in [0, 2a_s]$ controls the similarity between $\boldsymbol{s}$ and $\boldsymbol{s}_0$.

In the formulation of the waveform design problem (1.9), we have assumed that the prior knowledge on $\boldsymbol{R}_h$, $\boldsymbol{R}_n$, and $\{\boldsymbol{R}_{c,p}\}_{p=-P}^{P}$ are available. To justify the assumption, we highlight that $\boldsymbol{R}_h$ and $\{\boldsymbol{R}_{c,p}\}_{p=-P}^{P}$ can be estimated from the received signals in previous scans of the MIMO radar (see, e.g., [31–33]). Moreover, the knowledge on the statistical property of the external disturbance can be gained by configuring the radar in a passive mode.

1.3 Optimal Waveforms for Distributed MIMO Radar in the Absence of Clutter

In this section, we consider a simple case, in which the receive antennas of the MIMO radar are widely separated (i.e., a distributed MIMO radar), and the clutter in the received signals is ignored. Due to the widely separated receive antennas, the spatial correlation at the receiver end is negligible. Moreover, similar to [23, 25, 29], we assume that $\boldsymbol{R}_h = \boldsymbol{I} \otimes \bar{\boldsymbol{R}}_h$ and $\boldsymbol{R}_n = \boldsymbol{I} \otimes \bar{\boldsymbol{R}}_n$, i.e., the columns of $\boldsymbol{H}$ and $\boldsymbol{N}$ are independent and identically distributed, where $\bar{\boldsymbol{R}}_n$ stands for the temporal correlation matrix of the interference. Based on this assumption, we have that $\boldsymbol{R}_0 = \boldsymbol{I} \otimes \bar{\boldsymbol{R}}_n$ and $\boldsymbol{R}_1 = \boldsymbol{I} \otimes (\bar{\boldsymbol{R}}_n + \boldsymbol{S}\bar{\boldsymbol{R}}_h\boldsymbol{S}^\dagger)$. Thus, $f(\boldsymbol{S})$ is given by

$$f(\boldsymbol{S}) = N_R \left(\log\det \left(\boldsymbol{I} + \bar{\boldsymbol{R}}_n^{-1}\boldsymbol{S}\bar{\boldsymbol{R}}_h\boldsymbol{S}^\dagger \right) + \mathrm{tr}\left[(\boldsymbol{I} + \bar{\boldsymbol{R}}_n^{-1}\boldsymbol{S}\bar{\boldsymbol{R}}_h\boldsymbol{S}^\dagger)^{-1} \right] \right). \tag{1.15}$$

Therefore, the energy constrained waveform design problem based on relative entropy maximization can be formulated as

$$\begin{aligned} &\max_{\boldsymbol{S}}\ \log\det \left(\boldsymbol{I} + \bar{\boldsymbol{R}}_n^{-1}\boldsymbol{S}\bar{\boldsymbol{R}}_h\boldsymbol{S}^\dagger \right) + \mathrm{tr}\left[\left(\boldsymbol{I} + \bar{\boldsymbol{R}}_n^{-1}\boldsymbol{S}\bar{\boldsymbol{R}}_h\boldsymbol{S}^\dagger \right)^{-1} \right] \\ &\ \text{s.t. } \mathrm{tr}(\boldsymbol{S}\boldsymbol{S}^\dagger) = e_t, \end{aligned} \tag{1.16}$$

where we have ignored the constant terms.

In [25], it is proved that the optimal solution to (1.16) is given by

$$\boldsymbol{S}_{\text{opt}} = \boldsymbol{V}_n \left[\boldsymbol{0}_{N_T \times (L-N_T)}\ \boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_h^{-1/2} \right]^\top \boldsymbol{V}_h^\dagger, \tag{1.17}$$

where $\boldsymbol{V}_n$ is a unitary matrix whose columns are the eigenvectors corresponding to the eigenvalues of $\bar{\boldsymbol{R}}_n$ in decreasing order, $\boldsymbol{V}_h$ is a unitary matrix whose columns are the eigenvectors corresponding to the eigenvalues of $\bar{\boldsymbol{R}}_h$ in increasing order, $\boldsymbol{\Sigma}_h = \mathrm{Diag}([\sigma_{h,1}, \dots, \sigma_{h,N_T}])$, $\sigma_{h,1} \le \cdots \le \sigma_{h,N_T}$ are the eigenvalues of $\bar{\boldsymbol{R}}_h$, $\boldsymbol{\Sigma}_1 = \mathrm{Diag}([\sigma_{1,N_T}, \dots, \sigma_{1,1}])$, $\sigma_{1,m}$ $(m = 1, \dots, N_T)$ is given by

$$\sigma_{1,m} = \frac{-p_m - \sqrt{p_m^2 - 4\varpi^2\sigma_{n,L+1-m}^2}}{2\varpi},$$

$p_m = 2\varpi\sigma_{n,L+1-m} + \sigma_{h,N_T+1-m}$, ϖ is the solution to the equation $\sum_{m=1}^{N_T} (\sigma_{1,m})^+ \sigma_{h,N_T+1-m}^{-1} = e_t$ (which can be obtained, for instance, using Newton's method), and $\sigma_{n,1} \ge \cdots \ge \sigma_{n,L}$ are the eigenvalues of $\bar{\boldsymbol{R}}_n$. According to (1.17), the maximum relative entropy that can be achieved by an MIMO radar is given by

$$D_{\text{opt}} = \sum_{m=1}^{N_T} N_R \left(\log\left(1 + \frac{\sigma_{1,m}}{\sigma_{n,L+1-m}} \right) + \left(1 + \frac{\sigma_{1,m}}{\sigma_{n,L+1-m}} \right)^{-1} - N_T \right). \tag{1.18}$$

We can see from (1.17) that the left and right singular vectors of the optimal waveforms should be aligned with the eigenvectors of $\bar{\boldsymbol{R}}_n$ and $\bar{\boldsymbol{R}}_h$, respectively. Moreover, the eigenvector corresponding to the mth smallest eigenvalue of $\bar{\boldsymbol{R}}_n$ should be paired with the eigenvector corresponding to the mth largest eigenvalue of $\bar{\boldsymbol{R}}_h$. This pairing is referred to as "mode alignment" in the area of compressive sensing [34]. We highlight that the relevance of "mode alignment" is also revealed in waveform design for MIMO radar in spectrally crowded environments [28, 35].

In [25], it is also proved that the maximum relative entropy that can be achieved by a MIMO radar transmitting orthogonal waveforms (i.e., the waveform matrix $\boldsymbol{S}$ is a semi-unitary matrix that satisfies $\boldsymbol{S}^\dagger \boldsymbol{S} = e_t/N_T \cdot \boldsymbol{I}$) is given by

$$D_{\text{orth}} = \sum_{m=1}^{N_T} N_R \left(\log\left(1 + \frac{e_t}{N_T} \frac{\sigma_{h,m}}{\sigma_{n,L+1-m}}\right) + \left(1 + \frac{e_t}{N_T} \frac{\sigma_{h,m}}{\sigma_{n,L+1-m}}\right)^{-1} - N_T \right), \tag{1.19}$$

and the orthogonal waveform matrix achieving the maximum relative entropy is given by

$$\boldsymbol{S}_{\text{orth}} = \sqrt{\frac{e_t}{N_T}} \boldsymbol{V}_n \left[\boldsymbol{0}_{N_T \times (L-N_T)}\ \boldsymbol{I}_{N_T}\right]^\top \boldsymbol{V}_h^\dagger. \tag{1.20}$$

We proceed to give a simple case to showcase the performance of the optimal waveforms maximizing the relative entropy. In this example, the distributed MIMO radar has $N_T = 4$ transmit antennas and $N_R = 4$ receive antennas. The code length is $L = 5$. The eigenvalues of $\bar{\boldsymbol{R}}_h$ and $\bar{\boldsymbol{R}}_n$ are $\{0.5, 1, 2, 5\}$ and $\{10, 8, 4, 3, 2\}$, respectively. The eigenvectors of $\bar{\boldsymbol{R}}_h$ and $\bar{\boldsymbol{R}}_n$ are randomly generated. First, we comment on the optimal power allocation of the transmit waveforms. Figure 1.2 shows the allocated power on each mode of the optimal waveform matrix for $e_t = 10$ dB and $e_t = 25$ dB, respectively. We can observe that for the mode that corresponds to a larger $\sigma_{h,m}/\sigma_{n,L+1-m}$ (i.e., the target response is stronger and the interference is weaker), more power is allocated. Moreover, in the case of $e_t = 10$ dB, only two modes are active, implying a higher coherency between the transmit waveforms. If we increase the transmit energy to $e_t = 25$ dB, all the modes are activated, meaning a higher waveform diversity.

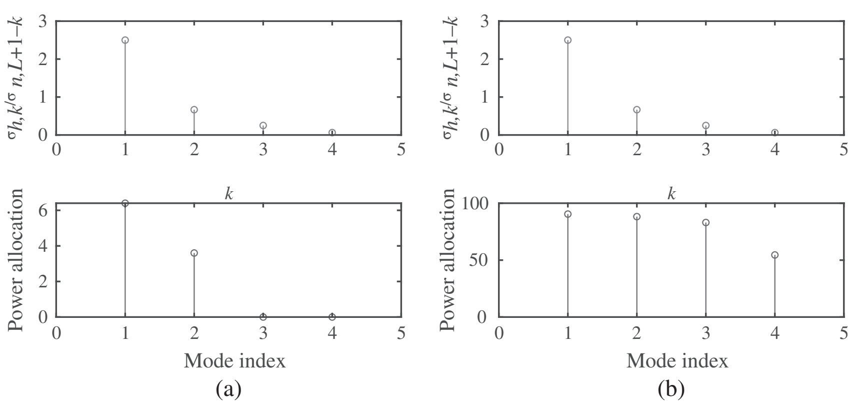

Figure 1.2 Optimal power allocation of the transmit waveforms. (a) $e_t = 10$ dB. (b) $e_t = 25$ dB.

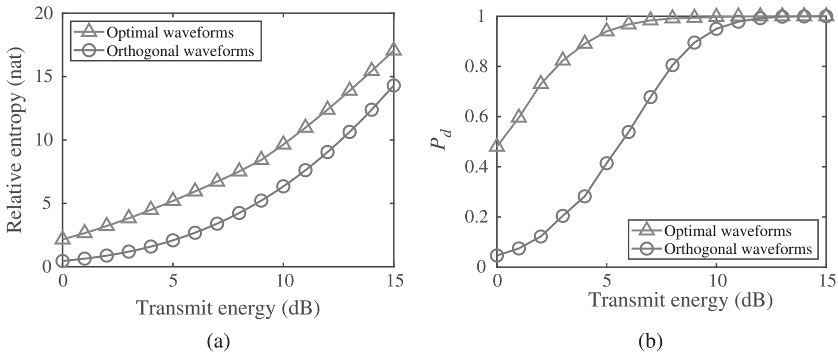

Figure 1.3 Performance comparison between the optimal waveforms and the optimal orthogonal waveforms. (a) Relative entropy versus transmit energy. (b) P_d versus transmit energy.

Now, we compare the performance of the optimal waveforms in (1.17) with that of the optimal orthogonal waveforms in (1.20). Figure 1.3(a) shows the relative entropy of the optimal waveforms and the orthogonal waveforms versus the transmit energy. We can see that, for a given transmit energy, the optimal waveforms achieve a larger relative entropy than the orthogonal waveforms. Figure 1.3(b) compares the detection probability (P_d) of the waveforms corresponding to Figure 1.3(a). In the analysis, the optimal Neyman–Pearson detector (see (1.7)) is used. In addition, the probability of false alarm is $P_{fa} = 10^{-3}$, and 10^5 independent Monte Carlo trials are conducted to determine the detection threshold and the detection probability, respectively. The results in Figure 1.3(b) are consistent with those in Figure 1.3(a), demonstrating that the performance of the optimal waveforms is visibly superior to that of the optimal orthogonal waveforms. This is attributed to the fact that the optimal waveforms have a larger number of degrees of freedom and thus achieve additional waveform diversity.

1.4 MM-Based Waveform Design in the Presence of Range-Spread Clutter

In this section, we design waveforms to boost the detection performance of MIMO radar in the presence of range-spread clutter. To this end, we consider the optimization problem (1.9). In general, this problem is nonconvex. In the following, we derive a minorization–maximization (MM) algorithm [36] to tackle this nonconvex problem. To this end, we split the objective function in (1.9) into three parts:

$$f(\boldsymbol{S}) = \log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_1) - \log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_0) + \mathrm{tr}(\boldsymbol{R}_1^{-1}\boldsymbol{R}_0)$$
$$= f_1(\boldsymbol{S}) + f_2(\boldsymbol{S}) + f_3(\boldsymbol{S}), \tag{1.21}$$

where $f_1(\boldsymbol{S}) = \log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_1)$, $f_2(\boldsymbol{S}) = -\log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_0)$, and $f_3(\boldsymbol{S}) = \mathrm{tr}(\boldsymbol{R}_1^{-1}\boldsymbol{R}_0)$. Next, we derive a quadratic surrogate function (i.e., a minorizer) for each $f_r(\boldsymbol{S})$, $r = 1, 2, 3$.

1.4.1 Surrogate Function Construction

Define $\widetilde{\boldsymbol{X}} = [\widetilde{\boldsymbol{S}}_{-P}, \dots, \widetilde{\boldsymbol{S}}_P] \in \mathbb{C}^{LN_R \times N_T N_R N_c}$, $\boldsymbol{R}_C = \mathrm{BlkDiag}(\boldsymbol{R}_{c,-P}; \dots; \boldsymbol{R}_{c,P}) \in \mathbb{C}^{N_T N_R N_c \times N_T N_R N_c}$, and $\boldsymbol{R}_{TC} = \mathrm{BlkDiag}(\boldsymbol{R}_{hc,-P}; \dots; \boldsymbol{R}_{hc,P}) \in \mathbb{C}^{N_T N_R N_c \times N_T N_R N_c}$, where $\boldsymbol{R}_{hc,p} = \boldsymbol{R}_{c,p}$ if $p \neq 0$, and $\boldsymbol{R}_{hc,0} = \boldsymbol{R}_{c,0} + \boldsymbol{R}_h$. Then $\boldsymbol{R}_0 = \widetilde{\boldsymbol{X}}\boldsymbol{R}_C\widetilde{\boldsymbol{X}}^\dagger + \boldsymbol{R}_n$ and $\boldsymbol{R}_1 = \widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}\widetilde{\boldsymbol{X}}^\dagger + \boldsymbol{R}_n$. Following the derivations in [27, 37], we can show that

$$f_1(\boldsymbol{S}) = \log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_1) = \log\det(\boldsymbol{K}_1\boldsymbol{T}^{-1}\boldsymbol{K}_1^\top), \tag{1.22}$$

where $\boldsymbol{K}_1 = [\boldsymbol{I}_{N_T N_R N_c}, \boldsymbol{0}_{N_T N_R N_c \times LN_R}]$, and

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{I}_{N_T N_R N_c} & \boldsymbol{R}_{TC}^{\frac{1}{2}}\widetilde{\boldsymbol{X}}^\dagger \\ \widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}^{\frac{1}{2}} & \boldsymbol{R}_n + \widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}\widetilde{\boldsymbol{X}}^\dagger \end{bmatrix}.$$

It follows from Lemma 1 in [27] that $\log\det(\boldsymbol{K}_1\boldsymbol{T}^{-1}\boldsymbol{K}_1^\top)$ is convex with respect to $\boldsymbol{T}$. By using a property of convex functions [38], we have

$$\log\det(\boldsymbol{K}_1\boldsymbol{T}^{-1}\boldsymbol{K}_1^\top) \geq \log\det(\boldsymbol{K}_1\boldsymbol{T}_k^{-1}\boldsymbol{K}_1^\top) + \mathrm{tr}(\boldsymbol{G}_{T_k}(\boldsymbol{T} - \boldsymbol{T}_k)),$$

where $\boldsymbol{G}_{T_k} = -\boldsymbol{T}_k^{-1}\boldsymbol{K}_1^\top(\boldsymbol{K}_1\boldsymbol{T}_k^{-1}\boldsymbol{K}_1^\top)^{-1}\boldsymbol{K}_1\boldsymbol{T}_k^{-1}$ is the gradient of $\log\det(\boldsymbol{K}_1\boldsymbol{T}^{-1}\boldsymbol{K}_1^\top)$ at $\boldsymbol{T}_k$ (which can be shown using results from [39]). Let us partition $\boldsymbol{G}_{T_k}$ as:

$$\boldsymbol{G}_{T_k} = \begin{bmatrix} \boldsymbol{G}_{T_k}^{11} & \boldsymbol{G}_{T_k}^{12} \\ \boldsymbol{G}_{T_k}^{21} & \boldsymbol{G}_{T_k}^{22} \end{bmatrix},$$

where $\boldsymbol{G}_{T_k}^{11} \in \mathbb{C}^{N_T N_R N_c \times N_T N_R N_c}$, $\boldsymbol{G}_{T_k}^{12} = (\boldsymbol{G}_{T_k}^{21})^\dagger \in \mathbb{C}^{N_T N_R N_c \times LN_R}$, and $\boldsymbol{G}_{T_k}^{22} \in \mathbb{C}^{LN_R \times LN_R}$. One can verify that $\boldsymbol{G}_{T_k}^{21} = \boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{X}}_k\boldsymbol{R}_{TC}^{\frac{1}{2}}$, $\boldsymbol{G}_{T_k}^{22} = -\boldsymbol{G}_{T_k}^{21}(\boldsymbol{K}_1\boldsymbol{T}_k^{-1}\boldsymbol{K}_1^\top)^{-1}\boldsymbol{G}_{T_k}^{12}$, and $\boldsymbol{G}_{T_k}^{22} = \boldsymbol{R}_{1,k}^{-1} - \boldsymbol{R}_n^{-1}$ [37]. As a result,

$$\mathrm{tr}(\boldsymbol{G}_{T_k}\boldsymbol{T}) = \mathrm{tr}(\boldsymbol{G}_{T_k}^{11}) + \mathrm{tr}(\boldsymbol{G}_{T_k}^{22}\boldsymbol{R}_n) + 2\mathrm{Re}(\mathrm{tr}(\boldsymbol{G}_{T_k}^{12}\widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}^{\frac{1}{2}})) + \mathrm{tr}(\boldsymbol{G}_{T_k}^{22}\widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}\widetilde{\boldsymbol{X}}^\dagger).$$

Define $c_1 = \log\det(\boldsymbol{K}_1\boldsymbol{T}_k^{-1}\boldsymbol{K}_1^\top) + \mathrm{tr}(\boldsymbol{G}_{T_k}^{11}) + \mathrm{tr}(\boldsymbol{G}_{T_k}^{22}\boldsymbol{R}_n) - \mathrm{tr}(\boldsymbol{G}_{T_k}\boldsymbol{T}_k)$. Then a surrogate function of $f_1(\boldsymbol{S})$ is given by

$$g_1(\boldsymbol{S}; \boldsymbol{S}_k) = 2\mathrm{Re}\left(\mathrm{tr}(\widetilde{\boldsymbol{X}}\boldsymbol{Z}_k^\dagger)\right) + \mathrm{tr}\left(\boldsymbol{G}_{T_k}^{22}\widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}\widetilde{\boldsymbol{X}}^\dagger\right) + c_1. \tag{1.23}$$

Next, we rewrite $f_2(\boldsymbol{S})$ as

$$f_2(\boldsymbol{S}) = -\log\det(\boldsymbol{R}_n^{-1}\boldsymbol{R}_0) = -\log\det(\boldsymbol{R}_0) + \log\det(\boldsymbol{R}_n). \tag{1.24}$$

Using the fact that $-\log\det(\boldsymbol{A})$ is convex with respect to $\boldsymbol{A} \succeq \boldsymbol{0}$ [38], we can obtain

$$-\log\det(\boldsymbol{R}_0) \geq -\log\det(\boldsymbol{R}_{0,k}) - \mathrm{tr}\left(\boldsymbol{R}_{0,k}^{-1}(\boldsymbol{R}_0 - \boldsymbol{R}_{0,k})\right).$$

Define $c_2 = \log\det(\boldsymbol{R}_n) + LN_R - \log\det(\boldsymbol{R}_{0,k}) - \mathrm{tr}(\boldsymbol{R}_{0,k}^{-1}\boldsymbol{R}_n)$. Then one can verify that a surrogate function of $f_2(\boldsymbol{S})$ can be written as

$$g_2(\boldsymbol{S};\boldsymbol{S}_k) = -\mathrm{tr}\left(\boldsymbol{R}_{0,k}^{-1}\widetilde{\boldsymbol{X}}\boldsymbol{R}_C\widetilde{\boldsymbol{X}}^\dagger\right) + c_2. \tag{1.25}$$

Finally, we note that $f_3(\boldsymbol{S}) = \mathrm{tr}(\boldsymbol{R}_1^{-1}\boldsymbol{R}_0) = -\mathrm{tr}(\widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_1^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{R}_h) + LN_R$. Moreover, by using the matrix inversion lemma, we have $\widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_1^{-1}\widetilde{\boldsymbol{S}} = (\boldsymbol{K}_3\boldsymbol{M}^{-1}\boldsymbol{K}_3^\top)^{-1}$, where $\boldsymbol{K}_3 = [\boldsymbol{I}_{N_TN_R}, \boldsymbol{0}_{N_TN_R\times N_TN_RN_c}]$,

$$\boldsymbol{M} = \begin{bmatrix}\boldsymbol{M}^{11} & \boldsymbol{M}^{12}\\ \boldsymbol{M}^{21} & \boldsymbol{M}^{22}\end{bmatrix},$$

$\boldsymbol{M}^{11} = \widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{S}}$, $\boldsymbol{M}^{12} = (\boldsymbol{M}^{21})^\dagger = \widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}^{\frac{1}{2}}$, and $\boldsymbol{M}^{22} = \boldsymbol{I} + \boldsymbol{R}_{TC}^{\frac{1}{2}}\widetilde{\boldsymbol{X}}^\dagger\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{X}}\boldsymbol{R}_{TC}^{\frac{1}{2}}$. Thus,

$$f_3(\boldsymbol{S}) = -\mathrm{tr}\left(\boldsymbol{R}_h\left(\boldsymbol{K}_3\boldsymbol{M}^{-1}\boldsymbol{K}_3^\top\right)^{-1}\right) + LN_R. \tag{1.26}$$

Note that $-\mathrm{tr}(\boldsymbol{R}_h(\boldsymbol{K}_3\boldsymbol{M}^{-1}\boldsymbol{K}_3^\top)^{-1})$ is convex with respect to $\boldsymbol{M}$ [27]. Hence, we have

$$-\mathrm{tr}(\boldsymbol{R}_h(\boldsymbol{K}_3\boldsymbol{M}^{-1}\boldsymbol{K}_3^\top)^{-1}) \geq -\mathrm{tr}(\boldsymbol{R}_h(\boldsymbol{K}_3\boldsymbol{M}_k^{-1}\boldsymbol{K}_3^\top)^{-1}) + \mathrm{tr}(\boldsymbol{G}_{M_k}(\boldsymbol{M} - \boldsymbol{M}_k)),$$

where $\boldsymbol{G}_{M_k} = -\boldsymbol{M}_k^{-1}\boldsymbol{K}_3\boldsymbol{A}(\boldsymbol{S}_k)\boldsymbol{R}_h\boldsymbol{A}(\boldsymbol{S}_k)\boldsymbol{K}_3^\top\boldsymbol{M}_k^{-1}$ is the gradient of $-\mathrm{tr}(\boldsymbol{R}_h(\boldsymbol{K}_3\boldsymbol{M}^{-1}\boldsymbol{K}_3^\top)^{-1})$ at $\boldsymbol{M}_k$, and $\boldsymbol{A}(\boldsymbol{S}) = \widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_1^{-1}\widetilde{\boldsymbol{S}}$. Let us partition $\boldsymbol{G}_{M_k}$ as

$$\boldsymbol{G}_{M_k} = \begin{bmatrix}\boldsymbol{G}_{M_k}^{11} & \boldsymbol{G}_{M_k}^{12}\\ \boldsymbol{G}_{M_k}^{21} & \boldsymbol{G}_{M_k}^{22}\end{bmatrix},$$

where $\boldsymbol{G}_{M_k}^{11} \in \mathbb{C}^{N_TN_R\times N_TN_R}$, $\boldsymbol{G}_{M_k}^{12} = (\boldsymbol{G}_{M_k}^{21})^\dagger \in \mathbb{C}^{N_TN_R\times N_TN_RN_c}$, and $\boldsymbol{G}_{M_k}^{22} \in \mathbb{C}^{N_TN_RN_c\times N_TN_RN_c}$. It can be checked that $\boldsymbol{G}_{M_k}^{11} = -\boldsymbol{R}_h$, $\boldsymbol{G}_{M_k}^{12} = \boldsymbol{R}_h\boldsymbol{M}_k^{12}(\boldsymbol{M}_k^{22})^{-1}$, and $\boldsymbol{G}_{M_k}^{22} = -(\boldsymbol{M}_k^{22})^{-1}\boldsymbol{M}_k^{21}\boldsymbol{R}_h\boldsymbol{M}_k^{12}(\boldsymbol{M}_k^{22})^{-1}$ [37]. Therefore, we have

$$\mathrm{tr}(\boldsymbol{G}_{M_k}\boldsymbol{M}) = \mathrm{tr}(\boldsymbol{G}_{M_k}^{11}\boldsymbol{M}^{11}) + \mathrm{tr}(\boldsymbol{G}_{M_k}^{22}\boldsymbol{M}^{22}) + 2\mathrm{Re}(\mathrm{tr}(\boldsymbol{G}_{M_k}^{12}\boldsymbol{M}^{21})).$$

Define $\hat{c}_3 = LN_R - \mathrm{tr}(\boldsymbol{R}_h(\boldsymbol{K}_3\boldsymbol{M}_k^{-1}\boldsymbol{K}_3^\top)^{-1}) - \mathrm{tr}(\boldsymbol{G}_{M_k}\boldsymbol{M}_k)$. Then $f_3(\boldsymbol{S})$ is minorized by

$$g_3(\boldsymbol{S};\boldsymbol{S}_k) = \mathrm{tr}(\boldsymbol{G}_{M_k}^{11}\boldsymbol{M}^{11}) + \mathrm{tr}(\boldsymbol{G}_{M_k}^{22}\boldsymbol{M}^{22}) + 2\mathrm{Re}(\mathrm{tr}(\boldsymbol{G}_{M_k}^{12}\boldsymbol{M}^{21})) + \hat{c}_3. \tag{1.27}$$

Based on the above results, it can be inferred that a surrogate function of $f(\boldsymbol{S})$ is given by

$$g(\boldsymbol{S};\boldsymbol{S}_k) = g_1(\boldsymbol{S};\boldsymbol{S}_k) + g_2(\boldsymbol{S};\boldsymbol{S}_k) + g_3(\boldsymbol{S};\boldsymbol{S}_k). \tag{1.28}$$

Moreover, we can verify that $g(\boldsymbol{S}; \boldsymbol{S}_k) \leq f(\boldsymbol{S})$ and $g(\boldsymbol{S}_k; \boldsymbol{S}_k) = f(\boldsymbol{S}_k)$. Thus, $f(\boldsymbol{S})$ is minorized by $g(\boldsymbol{S}; \boldsymbol{S}_k)$.

To show that $g(\boldsymbol{S}; \boldsymbol{S}_k)$ can be written as a quadratic function of $\boldsymbol{s}$, we let $\boldsymbol{Z}_k = [\boldsymbol{Z}_{k,-P}, \dots, \boldsymbol{Z}_{k,P}]$, where $\boldsymbol{Z}_{k,p} \in \mathbb{C}^{LN_R \times N_T N_R}$, and let $\boldsymbol{F}_k = \boldsymbol{G}_{M_k}^{12} \boldsymbol{R}_{TC}^{\frac{1}{2}} = [\boldsymbol{F}_{k,-P}, \dots, \boldsymbol{F}_{k,P}]$, where $\boldsymbol{F}_{k,p} \in \mathbb{C}^{N_T N_R \times N_T N_R}$, $p = -P, \dots, P$. Partition $\boldsymbol{W}_k = \boldsymbol{R}_{TC}^{\frac{1}{2}} \boldsymbol{G}_{M_k}^{22} \boldsymbol{R}_{TC}^{\frac{1}{2}}$ as

$$\begin{bmatrix} \boldsymbol{W}_{11,k} & \boldsymbol{W}_{12,k} & \cdots & \boldsymbol{W}_{1N_c,k} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{W}_{N_c1,k} & \cdots & \cdots & \boldsymbol{W}_{N_cN_c,k} \end{bmatrix},$$

where $\boldsymbol{W}_{\bar{p}\bar{q},k} \in \mathbb{C}^{N_T N_R \times N_T N_R}$, $\bar{p} = 1, \dots, N_c$, $\bar{q} = 1, \dots, N_c$. Then according to Proposition 1 in [37],

$$g(\boldsymbol{S}; \boldsymbol{S}_k) = 2\mathrm{Re}(\boldsymbol{b}_k^\dagger \boldsymbol{s}) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s} + c_0, \tag{1.29}$$

where $\boldsymbol{b}_k = \sum_{p=-P}^{P} \widetilde{\boldsymbol{E}}_p^\top \boldsymbol{z}_{k,p}$, $\boldsymbol{z}_{k,p} = \mathrm{vec}(\boldsymbol{Z}_{k,p})$, $\widetilde{\boldsymbol{E}}_p = [\widetilde{\boldsymbol{E}}_{1,p}^\top, \dots, \widetilde{\boldsymbol{E}}_{N_T N_R,p}^\top]^\top \in \mathbb{R}^{N_T N_R^2 L \times L N_T}$, $\widetilde{\boldsymbol{E}}_{r,p} = \boldsymbol{E}_r^\top \otimes \boldsymbol{J}_p$, $\mathrm{vec}(\boldsymbol{E}_r) = \boldsymbol{e}_r$ with $\boldsymbol{e}_r \in \mathbb{R}^{N_T N_R}$ being the rth column of $\boldsymbol{I}_{N_T N_R}$ $(r = 1, \dots, N_T N_R)$, $\boldsymbol{B}_k = \boldsymbol{B}_{4,k} + \boldsymbol{B}_{32,k} + \boldsymbol{B}_{33,k} + \boldsymbol{B}_{33,k}^\dagger$, $\boldsymbol{B}_{4,k} = \sum_{p=-P}^{P} \widetilde{\boldsymbol{E}}_p^\top (\boldsymbol{R}_{c,p}^* \otimes (\boldsymbol{G}_{T_k}^{22} - \boldsymbol{R}_{0,k}^{-1}))\widetilde{\boldsymbol{E}}_p + \widetilde{\boldsymbol{E}}_0^\top (\boldsymbol{R}_h^* \otimes (\boldsymbol{G}_{T_k}^{22} - \boldsymbol{R}_n^{-1}))\widetilde{\boldsymbol{E}}_0$, $\boldsymbol{B}_{32,k} = \sum_{q=-P}^{P} \sum_{p=-P}^{P} \widetilde{\boldsymbol{E}}_q^\top (\boldsymbol{W}_{\bar{q}\bar{p},k}^* \otimes \boldsymbol{R}_n^{-1})\widetilde{\boldsymbol{E}}_p$, $\boldsymbol{B}_{33,k} = \sum_{p=-P}^{P} \widetilde{\boldsymbol{E}}_0^\top (\boldsymbol{H}_{k,p}^* \otimes \boldsymbol{R}_n^{-1})\widetilde{\boldsymbol{E}}_p$, $\bar{p} = p + P + 1$, $\bar{q} = q + P + 1$, $c_0 = c_1 + c_2 + c_3$, and $c_3 = \mathrm{tr}(\boldsymbol{G}_{M_k}^{22}) + \hat{c}_3$.

Therefore, we can formulate the minorized problem based on (1.29) at the $(k+1)$th iteration as the following quadratic programming problem:

$$\begin{aligned} \max_{\boldsymbol{S}} \quad & 2\mathrm{Re}(\boldsymbol{b}_k^\dagger \boldsymbol{s}) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s} \\ \text{s.t.} \quad & \boldsymbol{S} \in \mathcal{S}. \end{aligned} \tag{1.30}$$

1.4.2 Efficient Solutions to the Quadratic Programming Problem (1.30)

• Energy Constraint

We can formulate the quadratic programming problem under the energy constraint as

$$\begin{aligned} \max_{\boldsymbol{s}} \; & 2\mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{b}_k) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s} \\ \text{s.t.} \; & \boldsymbol{s}^\dagger \boldsymbol{s} = e_t, \end{aligned} \tag{1.31}$$

where we have used the fact that $\mathrm{tr}(\boldsymbol{S}\boldsymbol{S}^\dagger) = \boldsymbol{s}^\dagger \boldsymbol{s}$. The optimization problem (1.31) can be solved by the Lagrange multiplier method (see, e.g., [12, 27, 37] for details), and the solution is given by

$$\boldsymbol{s}^{(k+1)} = \left(\varrho^{(k+1)} \boldsymbol{I} - \boldsymbol{B}_k\right)^{-1} \boldsymbol{b}_k, \tag{1.32}$$

where $\varrho^{(k+1)} \geq \lambda_{\max}(\boldsymbol{B}_k)$ is the optimal Lagrange multiplier corresponding to the energy constraint, which can be obtained via solving the equation $\boldsymbol{b}_k^\dagger(\varrho\boldsymbol{I} - \boldsymbol{B}_k)^{-2}\boldsymbol{b}_k = e_t$.

- **PAPR Constraint**

We formulate the quadratic programming problem under the PAPR constraint as

$$\begin{aligned} &\max_{\boldsymbol{s}} \ 2\mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{b}_k) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s} \\ &\quad \text{s.t.} \ \boldsymbol{s}_m^\dagger \boldsymbol{s}_m = e_t/N_T, \mathrm{PAPR}(\boldsymbol{s}_m) \leq \rho, m = 1, \dots, N_T. \end{aligned} \tag{1.33}$$

The above quadratic programming problem can be tackled by MM. To this purpose, note that $(\boldsymbol{s} - \boldsymbol{s}^{(k,q)})^\dagger(\boldsymbol{B}_k - \lambda_{\min}(\boldsymbol{B}_k)\boldsymbol{I})(\boldsymbol{s} - \boldsymbol{s}^{(k,q)}) \geq 0$, where $\boldsymbol{s}^{(k,q)}$ denotes the solution obtained by the proposed algorithm at the (k,q)th iteration. Here, the index k denotes the outer loop and the index q denotes the inner loop. Using the above inequality and some algebraic manipulations, we can show that $\boldsymbol{s}^\dagger\boldsymbol{B}_k\boldsymbol{s}$ is minorized by $2\mathrm{Re}(\boldsymbol{s}^\dagger(\boldsymbol{B}_k - \lambda_{\min}(\boldsymbol{B}_k)\boldsymbol{I})\boldsymbol{s}^{(k,q)}) - (\boldsymbol{s}^{(k,q)})^\dagger\boldsymbol{B}_k\boldsymbol{s}^{(k,q)} + 2\lambda_{\min}(\boldsymbol{B}_k)e_t$. Thus, a linear surrogate function of $g(\boldsymbol{S};\boldsymbol{S}_k)$ is given by

$$h(\boldsymbol{s};\boldsymbol{s}^{(k,q)}) = 2\mathrm{Re}(\boldsymbol{s}^\dagger\boldsymbol{d}_{k,q}) + c_4, \tag{1.34}$$

where

$$\boldsymbol{d}_{k,q} = \boldsymbol{b}_k + (\boldsymbol{B}_k - \lambda_{\min}(\boldsymbol{B}_k)\boldsymbol{I})\boldsymbol{s}^{(k,q)}, \tag{1.35}$$

and $c_4 = c_0 + 2\lambda_{\min}(\boldsymbol{B}_k)e_t - (\boldsymbol{s}^{(k,q)})^\dagger\boldsymbol{B}_k\boldsymbol{s}^{(k,q)}$. As a consequence, the following problem will be solved at the $(k, q+1)$th iteration:

$$\begin{aligned} &\max_{\boldsymbol{s}} \ \mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{d}_{k,q}) \\ &\quad \text{s.t.} \ \boldsymbol{s}_m^\dagger \boldsymbol{s}_m = e_t/N_T, \mathrm{PAPR}(\boldsymbol{s}_m) \leq \rho, m = 1, \dots, N_T. \end{aligned} \tag{1.36}$$

Let $\boldsymbol{d}_{k,q} = [\boldsymbol{u}_1^\top, \dots, \boldsymbol{u}_{N_T}^\top]^\top$, where $\boldsymbol{u}_m \in \mathbb{C}^L$, $m = 1, \dots, N_T$. Note that $\boldsymbol{s}^\dagger\boldsymbol{d}_{k,q} = \sum_{m=1}^{N_T}\boldsymbol{s}_m^\dagger\boldsymbol{u}_m$. Thus, the optimization problem (1.36) can be split into N_T independent subproblems, the mth of which is given by

$$\begin{aligned} &\max_{\boldsymbol{s}_m} \ \mathrm{Re}(\boldsymbol{s}_m^\dagger \boldsymbol{u}_m) \\ &\quad \text{s.t.} \ \boldsymbol{s}_m^\dagger \boldsymbol{s}_m = e_t/N_T, \mathrm{PAPR}(\boldsymbol{s}_m) \leq \rho. \end{aligned} \tag{1.37}$$

Algorithm 2 in [40] can be used to solve the above problem. Particularly, if $\rho = 1$, the solution to the optimization problem (1.37) is: $s_m^{(k,q+1)}(l) = a_s \exp(\jmath \arg(u_m(l)))$, where $s_m^{(k,q+1)}(l)$ and $u_m(l)$ are the lth element of $\boldsymbol{s}_m^{(k,q+1)}$ and $\boldsymbol{u}_m$.

• Similarity Constraint

We formulate the quadratic programming problem under the similarity constraint as

$$\begin{aligned}&\max_{\boldsymbol{s}}\ 2\mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{b}_k) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s}\\ &\text{ s.t. }\ \boldsymbol{s}^\dagger \boldsymbol{s} = e_t,\ \|\boldsymbol{s} - \boldsymbol{s}_0\|_2^2 \le \epsilon.\end{aligned} \tag{1.38}$$

Similarly, we use the linear function in (1.34) as a surrogate of the objective in (1.38) and the surrogate problem at the $(k, q+1)$th iteration is given by:

$$\begin{aligned}&\max_{\boldsymbol{s}}\ \mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{d}_{k,q})\\ &\text{ s.t. }\ \boldsymbol{s}^\dagger \boldsymbol{s} = e_t,\ \|\boldsymbol{s} - \boldsymbol{s}_0\|_2^2 \le \epsilon.\end{aligned} \tag{1.39}$$

The solution to this optimization problem is given by [26] (see also in [11, 27, 41] for similar solutions)

$$\boldsymbol{s}^{(k,q+1)} = \frac{\mu_2 \boldsymbol{s}_0 + \boldsymbol{d}_{k,q}}{\mu_1}, \tag{1.40}$$

where $\mu_1 = e_t^{-\frac{1}{2}} \|\mu_2 \boldsymbol{s}_0 + \boldsymbol{d}_{k,q}\|_2$, and μ_2 can be obtained by solving (details can be found in [[37], Appendix B]):

$$\begin{aligned}&\max_{\mu_2}\ -2e_t^{\frac{1}{2}} \|\mu_2 \boldsymbol{s}_0 + \boldsymbol{d}_{k,q}\|_2 + \mu_2(2e_t - \epsilon)\\ &\text{ s.t. }\ \mu_2 \ge 0.\end{aligned}$$

• CMS Constraints

Note that the constraint in (1.14) is tantamount to

$$|s(l)| = a_s, \varphi_0^-(l) \le \arg(s(l)) \le \varphi_0^+(l), l = 1, \dots, LN_T,$$

where $\varphi_0^-(l) = \varphi_0(l) - \delta_c$, $\varphi_0^+(l) = \varphi_0(l) + \delta_c$, $\varphi_0(l) = \arg(s_0(l))$, $s_0(l)$ denotes the lth element of $\boldsymbol{s}_0$, and $\delta_c = \arccos(1 - (\epsilon_\infty/a_s)^2/2)$. Thus, the quadratic programming problem under the CMS constraints is stated as:

$$\begin{aligned}&\max_{\boldsymbol{s}}\ 2\mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{b}_k) + \boldsymbol{s}^\dagger \boldsymbol{B}_k \boldsymbol{s}\\ &\text{ s.t. }\ |s(l)| = a_s, \varphi_0^-(l) \le \arg(s(l)) \le \varphi_0^+(l), l = 1, \dots, LN_T.\end{aligned} \tag{1.41}$$

By using the minorizer derived in (1.34) again, the surrogate problem at the $(k, q+1)$th iteration is given by:

$$\begin{aligned}&\max_{\boldsymbol{s}}\ \mathrm{Re}(\boldsymbol{s}^\dagger \boldsymbol{d}_{k,q})\\ &\text{ s.t. }\ |s(l)| = a_s, \varphi_0^-(l) \le \arg(s(l)) \le \varphi_0^+(l), l = 1, \dots, LN_T.\end{aligned} \tag{1.42}$$

The solution to (1.42) can be written as

$$s^{(k,q+1)}(l) = a_s \exp(\jmath \varphi_{s,k}(l)), \tag{1.43}$$

where

$$\varphi_{s,k}(l) = \begin{cases} \varphi_d(l), & \text{if } \varphi_0^-(l) \le \varphi_d(l) \le \varphi_0^+(l), \\ \varphi_0^-(l), & \text{if } \cos(\varphi_d^+(l)) < \cos(\varphi_d^-(l)), \\ \varphi_0^+(l), & \text{if } \cos(\varphi_d^-(l)) < \cos(\varphi_d^+(l)), \end{cases}$$

$\varphi_d(l) = \arg(d_{k,q}(l))$, $\varphi_d^-(l) = \varphi_d(l) - \varphi_0^-(l)$, and $\varphi_d^+(l) = \varphi_d(l) + \varphi_0^+(l)$.

1.4.3 Algorithm Summary and Properties

The proposed MIMO radar waveform design algorithm is summarized in Algorithm 1.1, where $f(\boldsymbol{s}) \triangleq f(\boldsymbol{S})$, $g(\boldsymbol{s}; \boldsymbol{s}^{(k)}) \triangleq g(\boldsymbol{S}; \boldsymbol{S}_k)$ $(\boldsymbol{s}^{(k)} = \text{vec}(\boldsymbol{S}_k))$, and $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ are predefined small values. The squared iterative methods (SQUAREM) [42] can be used to speed up the convergence of Algorithm 1.1.

Algorithm 1.1: Waveform design algorithm for MIMO radar target detection in range-spread clutter.

Input: $\{\boldsymbol{R}_{\text{hc},p}\}_{p=-P}^{P}$, $\boldsymbol{R}_{\text{n}}$, e_t.
Output: $\boldsymbol{s}^\star$.
Initialize: $k = 0$, $\boldsymbol{s}^{(0)}$.
repeat
  Compute $\boldsymbol{B}_k$ and $\boldsymbol{b}_k$;
  `// a) Energy constraint`
  Update $\boldsymbol{s}^{(k+1)}$ using (1.32);
  $q = 0$; $\boldsymbol{s}^{(k,q)} = \boldsymbol{s}^{(k)}$;
  repeat
    Compute $\boldsymbol{d}_{k,q}$ using (1.35);
    `// b) PAPR constraint`
    Update $\boldsymbol{s}^{(k,q+1)}$ by solving (1.36);
    `// c) Similarity constraint`
    Update $\boldsymbol{s}^{(k,q+1)}$ by (1.40);
    `// d) CMS constraints`
    Update $\boldsymbol{s}^{(k,q+1)}$ by (1.43);
    $q = q + 1$;
  until $|g(\boldsymbol{s}^{(k,q+1)}; \boldsymbol{s}^{(k)}) - g(\boldsymbol{s}^{(k,q)}; \boldsymbol{s}^{(k)})| / g(\boldsymbol{s}^{(k,q+1)}; \boldsymbol{s}^{(k)}) < \varepsilon_2$;
  $\boldsymbol{s}^{(k+1)} = \boldsymbol{s}^{(k,q)}$;
  $k = k + 1$;
until $|f(\boldsymbol{s}^{(k+1)}) - f(\boldsymbol{s}^{(k)})| / f(\boldsymbol{s}^{(k)}) < \varepsilon_1$;
$\boldsymbol{s}^\star = \boldsymbol{s}^{(k)}$.

(for more details, see [27]). Note that for simplicity the correlations between the clutter across different range cells have been ignored. Nevertheless, it is not difficult to extend the proposed waveform design algorithm to deal with correlated clutter.

Next we analyze the convergence property of the proposed algorithm. To demonstrate that the objective values rise progressively with each iteration, we first prove that $g(\boldsymbol{s}^{(k+1)}; \boldsymbol{s}^{(k)}) \geq g(\boldsymbol{s}^{(k)}; \boldsymbol{s}^{(k)})$. This is evident for the energy constraint because $\boldsymbol{s}^{(k+1)}$ is the optimal solution for given $\boldsymbol{s}^{(k)}$. For the other constraints, it can be checked that

$$\begin{aligned} g(\boldsymbol{s}^{(k,q+1)}; \boldsymbol{s}^{(k)}) &\geq h(\boldsymbol{s}^{(k,q+1)}; \boldsymbol{s}^{(k,q)}) \\ &\geq h(\boldsymbol{s}^{(k,q)}; \boldsymbol{s}^{(k,q)}) \\ &= g(\boldsymbol{s}^{(k,q)}; \boldsymbol{s}^{(k)}), \end{aligned} \tag{1.44}$$

where the first line and the third line are results of the properties of the constructed surrogate function, and the second line holds since $\boldsymbol{s}^{(k,q+1)}$ is the optimal solution when $\boldsymbol{s}^{(k,q)}$ is given. Thus, we can verify that

$$g(\boldsymbol{s}^{(k)}; \boldsymbol{s}^{(k)}) = g(\boldsymbol{s}^{(k,0)}; \boldsymbol{s}^{(k)}) \leq g(\boldsymbol{s}^{(k,N_{\text{iter}})}; \boldsymbol{s}^{(k)}) = g(\boldsymbol{s}^{(k+1)}; \boldsymbol{s}^{(k)}), \tag{1.45}$$

where N_{iter} denotes the required number of iterations for the problem (1.30) to achieve convergence. Note that $g(\boldsymbol{s}; \boldsymbol{s}^{(k)})$ is a minorizer of $f(\boldsymbol{s})$. As a consequence, $f(\boldsymbol{s}^{(k)}) = g(\boldsymbol{s}^{(k)}; \boldsymbol{s}^{(k)})$ and $f(\boldsymbol{s}^{(k+1)}) \geq g(\boldsymbol{s}^{(k+1)}; \boldsymbol{s}^{(k)})$. Combining with the result in (1.45), we have $f(\boldsymbol{s}^{(k+1)}) \geq f(\boldsymbol{s}^{(k)})$.

Remark 1.2 It is worth remarking that the MM algorithm derived in this section is suitable for a target covariance matrix with an arbitrary rank. Note that for a point-like target, which has a rank-one covariance matrix, the maximization of relative entropy is equivalent to the maximization of SINR. To prove this claim, we note that if $\boldsymbol{R}_h$ is of rank one, it can be factorized as $\boldsymbol{R}_h = \boldsymbol{h}_0\boldsymbol{h}_0^\dagger$. Then, $\log\det(\boldsymbol{R}_0^{-1}\boldsymbol{R}_1) = \log\det(\boldsymbol{I} + \boldsymbol{R}_0^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0\boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}^\dagger) = \log(1 + \text{SINR})$, where $\text{SINR} = \boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_0^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0$. In addition, $\text{tr}(\boldsymbol{R}_1^{-1}\boldsymbol{R}_0) = \text{tr}\,(\boldsymbol{I} + \boldsymbol{R}_0^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0\boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}^\dagger)^{-1} = LN_R + (1 + \text{SINR})^{-1} - 1$. As a result, the relative entropy corresponding to a point-like target is $D(P_0||P_1) = \log(1 + \text{SINR}) + (1 + \text{SINR})^{-1} - 1$. Note that $\log(1 + x) + (1 + x)^{-1}$ is a monotonically increasing function of $x \geq 0$ [25]. Therefore, for a point-like target, maximization of SINR is equivalent to maximization of relative entropy. In [[12], Lemma 1], it was proved that $\text{tr}(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^\dagger)$ is jointly convex with respect to $\boldsymbol{A}$ and $\boldsymbol{B}$ for $\boldsymbol{B} \succeq \boldsymbol{0}$. Thus, a minorizer of $\text{SINR} = \boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_0^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0$ can be found by using the following inequality:

$$\boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}^\dagger\boldsymbol{R}_0^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0 \geq 2\text{Re}\left(\boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}_k^\dagger\boldsymbol{R}_{0,k}^{-1}\widetilde{\boldsymbol{S}}\boldsymbol{h}_0\right) - \text{tr}\left(\boldsymbol{R}_{0,k}^{-1}\widetilde{\boldsymbol{S}}_k\boldsymbol{h}_0\boldsymbol{h}_0^\dagger\widetilde{\boldsymbol{S}}_k^\dagger\boldsymbol{R}_0\right). \tag{1.46}$$

As a result, we can use MM to tackle the SINR maximization problem. Moreover, it is not difficult to show that the minorizer on the right-hand side of (1.46) is a quadratic function of $\boldsymbol{s}$. Thus, we need to solve a quadratic programming problem at each iteration of the MM algorithm, similar to that in (1.30).

1.5 Performance Assessment

In this section, we assess the performance of the synthesized waveforms for a colocated MIMO radar in the presence of range-spread clutter. The MIMO radar has $N_T = 4$ transmit antennas and $N_R = 4$ receive antennas. Both the antenna arrays on the transmit and receive end are uniform and linear, with inter-element spacings of $d_T = 2\lambda$ and $d_R = \lambda/2$, respectively, where λ denotes the wavelength. The transmit waveforms have a code length of $L = 16$. The target and clutter distribution over the angle-range region is illustrated in Figure 1.4. Specifically, we consider two targets with directions of 10° and 11°, respectively. We denote the range cell where the targets are present as cell 0. Both targets have an average power of $\sigma_{t,k}^2 = \mathbb{E}(|a_{t,k}|^2) = 1$. The clutter is either in the same range cell as the targets (i.e., cell 0) or in the nearest cells (i.e., cell -1 or 1). The clutter distributions in each range cell are presented in Table 1.2. All the clutter patches have the same average power, i.e., $\sigma_{c,p,j}^2 = \mathbb{E}(|a_{c,p,j}|^2) = 100$. The disturbance $\boldsymbol{n}$ follows an AR(1) process with the correlation coefficient equal to 0.4. Quasi-orthogonal waveforms, which are randomly generated and have constant modulus, are used to initialize the proposed waveform design algorithm. For the proposed iterative algorithm, we set $\varepsilon_1 = 10^{-3}$ and $\varepsilon_2 = 10^{-10}$ to terminate the outer and the inner loop.

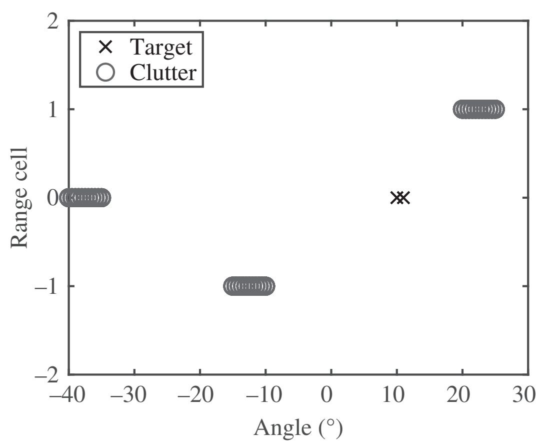

Figure 1.4 Target and clutter spatial distributions.

Table 1.2 Clutter distributions.

Range cell number	Maximum angle	Minimum angle	Grid step
−1	−15°	−10°	0.5°
0	−40°	−35°	0.5°
1	20°	25°	0.5°

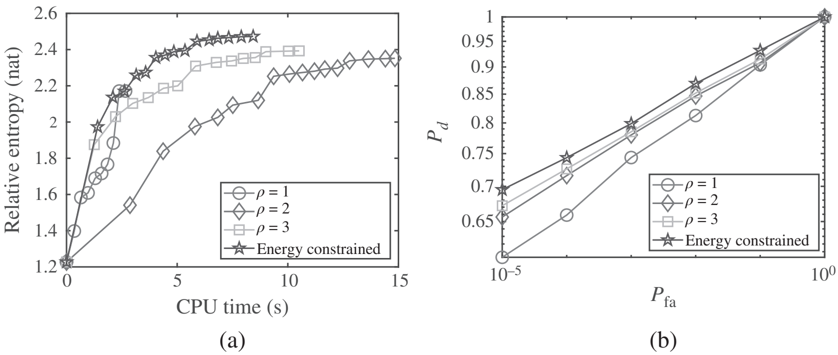

Figure 1.5 (a) Relative entropy of the waveforms synthesized under the PAPR constraint versus the iteration time. (b) P_d versus P_{fa} for the waveforms in (a). $e_t = 1$.

In the detection performance analysis of the synthesized waveforms versus the transmit energy, the probability of false alarm is fixed to be $P_{fa} = 10^{-3}$. Moreover, $100/P_{fa}$ Monte Carlo trials are performed to determine the detection threshold and calculate the detection probability, respectively. Finally, all the computations are conducted on a standard laptop with Intel Core i7-8550U CPU and 16 GB RAM.

Figure 1.5(a) analyzes the relative entropy convergence of the waveforms synthesized under the PAPR constraint, where the transmit waveforms have an energy of $e_t = 1$. The maximum allowed PAPR values (i.e., ρ) of the three sets of waveforms are 1, 2, and 3, respectively. To demonstrate the performance of the synthesized waveforms, we include the relative entropy of the energy-constrained waveforms (i.e., $\rho = L$) as a benchmark. It is seen that the proposed algorithm increases the objective value of the synthesized waveforms monotonically until convergence. Moreover, a larger PAPR value raises the relative entropy. Figure 1.5(b) shows the detection probability of the waveforms in Figure 1.5(a) versus the probability of false alarm. The results show that a larger PAPR value not only raises the relative entropy but also enhances detection performance.

In Figure 1.6(a), the relative entropy of the waveforms synthesized under the constant-modulus constraint (i.e., $\rho = 1$) is plotted for different transmit energy,

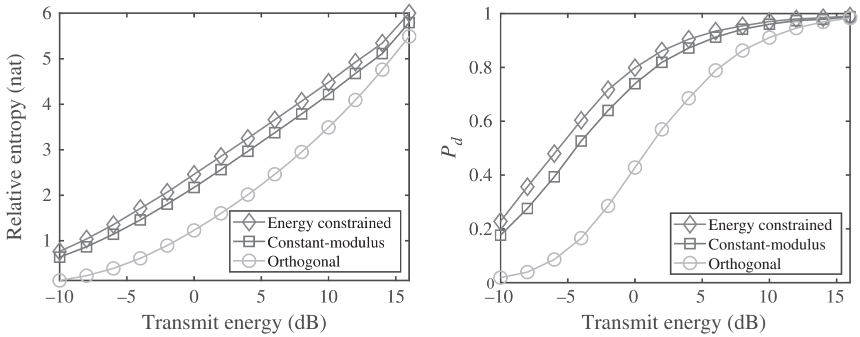

Figure 1.6 Performance assessment of the constant-modulus waveforms synthesized by the proposed algorithm. (a) Relative entropy versus transmit energy. (b) P_d versus transmit energy.

where the transmit energy of the waveforms is varied from −10 to 16 dB. Note that, as expected, the relative entropy rises as the transmit energy increases. Moreover, compared with the orthogonal waveforms, the performance of the waveforms synthesized under the constant-modulus constraint is closer to that of the energy-constrained waveforms. The detection performance associated with the waveforms in Figure 1.6(a) is shown in Figure 1.6(b). Observe that the result in Figure 1.6(b) is consistent with that in Figure 1.6(a), showing that the synthesized waveforms with a larger relative entropy achieve improved target detection performance.

Next, the performance of the waveforms synthesized under the similarity constraint is evaluated. Figure 1.7(a) plots the relative entropy convergence of the waveforms synthesized under the similarity constraint, where $e_t = 1$, and four different similarity parameters (i.e., ϵ) are considered. We use orthogonal waveforms as a reference as well as an initial point of the proposed algorithm. Figure 1.7(b) analyzes the detection probability of the waveforms in Figure 1.7(a) versus the probability of false alarm. Once again, the convergence of the proposed algorithm can be observed for all the four cases. For a smaller similarity parameter, the proposed algorithm converge slower. Because the feasibility region increases with the similarity parameter, the performance of the synthesized waveforms boosts as the similarity parameter increases.

Figure 1.8(a) shows the relative entropy of the waveforms synthesized under the similarity constraint, where the similarity parameter is equal to $0.5e_t$, and the transmit energy ranges from −10 to 16 dB. The detection probability of the synthesized waveforms in Figure 1.8(a) is analyzed in Figure 1.8(b). We can see that

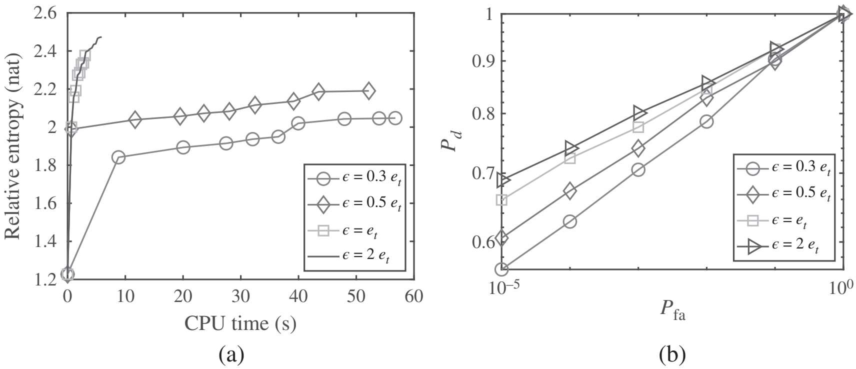

Figure 1.7 (a) Relative entropy of the waveforms synthesized under the similarity constraint. (b) P_d versus P_{fa} for the waveforms in (a). $e_t = 1$.

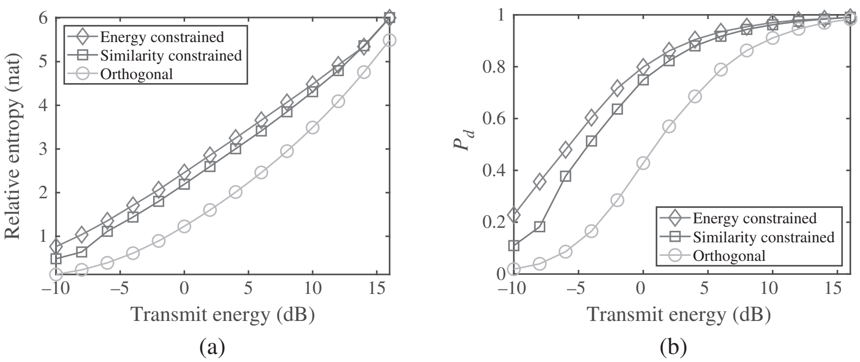

Figure 1.8 Performance assessment of the waveforms synthesized under the similarity constraint. (a) Relative entropy versus transmit energy. (b) P_d versus transmit energy. $\epsilon = 0.5e_t$.

the synthesized waveforms achieve a better performance as the transmit energy increases and demonstrate significant superiority over the orthogonal waveforms.

Finally, the performance of the waveforms synthesized under the CMS constraints is assessed. Figure 1.9(a) shows the relative entropy of the waveforms synthesized under the CMS constraints for $e_t = 1$ and four different values of ϵ_∞. We use orthogonal waveforms as a reference as well as an initial point. Figure 1.9(b) displays the detection probability of the waveforms in Figure 1.9(a) versus the probability of false alarm. It can be seen that the convergence of the proposed algorithm is ensured in all four cases. We also observe that the performance of the synthesized waveforms for the case of $\epsilon_\infty = a_s$ is very close to that for the case of $\epsilon_\infty = 2a_s$. Figure 1.10 shows the performance of the waveforms synthesized under

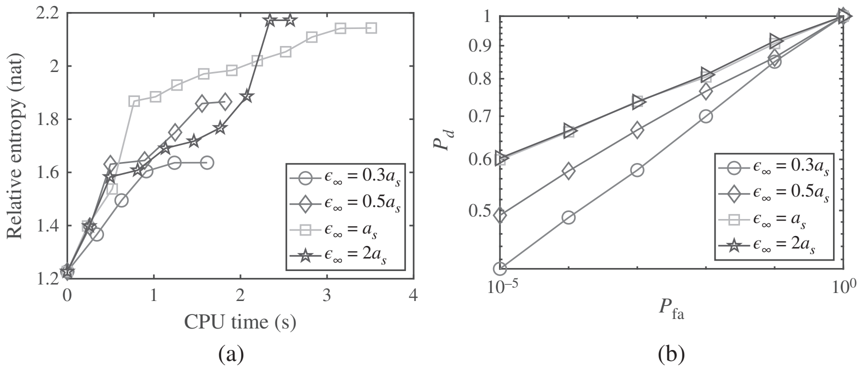

Figure 1.9 (a) Relative entropy of the waveforms synthesized under the CMS constraints. (b) P_d versus P_{fa} for the waveforms in (a). $e_t = 1$.

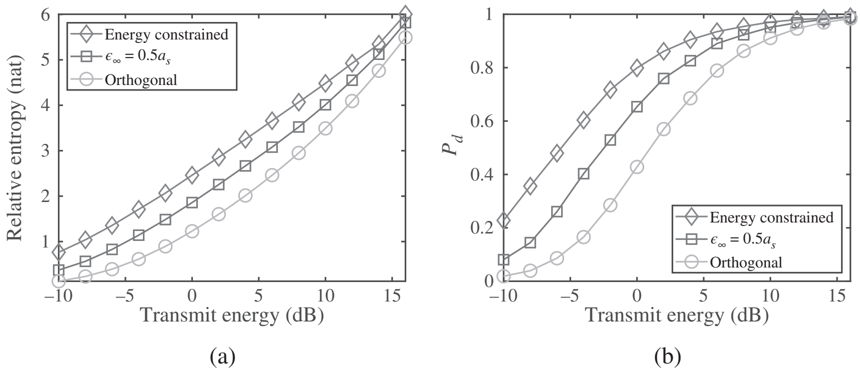

Figure 1.10 Performance assessment of the waveforms synthesized under the CMS constraints. (a) Relative entropy versus transmit energy. (b) P_d versus transmit energy. $\epsilon_\infty = 0.5a_s$.

the CMS constraints, where $\epsilon_\infty = 0.5a_s$, and the transmit energy ranges from −10 to 16 dB. It is evident that the waveforms synthesized under the CMS constraints outperform the orthogonal waveforms.

1.6 Conclusion

To improve the detection performance of MIMO radar in the presence of disturbance, we have addressed the problem of designing waveforms based on maximizing the relative entropy. We showed that for the case of a distributed MIMO radar in the absence of clutter, the optimal waveform matrix can be obtained in closed form. Moreover, the left singular vectors of the optimal waveform matrix should be aligned with the eigenvectors of the interference covariance matrix, and the right singular vectors should be aligned with the eigenvectors of the target covariance matrix. For the more general case of a MIMO radar operating in range-spread clutter, the proposed MM algorithm can synthesize waveforms under several practical constraints, and it has guaranteed local convergence. Numerical results showed that the waveforms synthesized by the proposed algorithm can achieve better detection performance than orthogonal waveforms.

Acknowledgments

The work of Bo Tang was supported in part by the National Natural Science Foundation of China under Grant 62171450, the Anhui Provincial Natural Science Foundation under Grant 2108085J30, and the Research Plan of National University of Defense Technology under Grant 23-ZZCX-JDZ-42. The work of Petre Stoica was supported in part by the Swedish Research Council (VR Grants 2021-05022, 2017-04610, and 2016-06079).

References

1 Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27(3): 379–423.

2 Cover, T.M. and Thomas, J.A. (2012). *Elements of Information Theory*. Wiley.

3 Woodward, P. (1951). Information theory and the design of radar receivers. *Proceedings of the IRE* 39(12): 1521–1524.

4 Woodward, P. (1953). Theory of radar information. *Transactions of the IRE Professional Group on Information Theory* 1(1): 108–113.

5 Woodward, P. (1953). *Probability and Information Theory, with Applications to Radar*. London: Pergamon.

6 Richmond, C.D. (1996). PDF's, confidence regions, and relevant statistics for a class of sample covariance-based array processors. *IEEE Transactions on Signal Processing* 44(7): 1779–1793.

7 Bell, M.R. (1993). Information theory and radar waveform design. *IEEE Transactions on Information Theory* 39(5): 1578–1597.

8 Romero, R. and Goodman, N. (2009). Waveform design in signal-dependent interference and application to target recognition with multiple transmissions. *IET Radar, Sonar and Navigation* 3(4): 328–340.

9 Romero, R.A., Bae, J., and Goodman, N.A. (2011). Theory and application of SNR and mutual information matched illumination waveforms. *IEEE Transactions on Aerospace and Electronic Systems* 47(2): 912–927.

10 Rowe, W., Stoica, P., and Li, J. (2014). Spectrally constrained waveform design. *IEEE Signal Processing Magazine* 31(3): 157–162.

11 Tang, B. and Liang, J. (2019). Efficient algorithms for synthesizing probing waveforms with desired spectral shapes. *IEEE Transactions on Aerospace and Electronic Systems* 55(3): 1174–1189.

12 Tang, B., Liu, J., Wang, H., and Hu, Y. (2021). Constrained radar waveform design for range profiling. *IEEE Transactions on Signal Processing* 69: 1924–1937.

13 Kay, S. (2007). Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation. *IEEE Journal of Selected Topics in Signal Processing* 1(1): 31–41.

14 Kay, S. (2009). Waveform design for multistatic radar detection. *IEEE Transactions on Aerospace and Electronic Systems* 45(3): 1153–1166.

15 Zhu, Z., Kay, S., and Raghavan, R.S. (2017). Information-theoretic optimal radar waveform design. *IEEE Signal Processing Letters* 24(3): 274–278.

16 Garren, D., Osborn, M., Odom, A. et al. (2001). Enhanced target detection and identification via optimised radar transmission pulse shape. *IEE Proceedings: Radar, Sonar and Navigation* 148(3): 130–138.

17 Li, J. and Stoica, P. (2009). *MIMO Radar Signal Processing*. Hoboken, NJ: Wiley-IEEE Press.

18 Li, J. and Stoica, P. (2007). MIMO radar with colocated antennas. *IEEE Signal Processing Magazine* 24(5): 106–114.

19 Haimovich, A.M., Blum, R.S., and Cimini, L.J. (2008). MIMO radar with widely separated antennas. *IEEE Signal Processing Magazine* 25(1): 116–129.

20 Yang, Y. and Blum, R.S. (2007). MIMO radar waveform design based on mutual information and minimum mean-square error estimation. *IEEE Transactions on Aerospace and Electronic Systems* 43(1): 330–343.

21 De Maio, A. and Lops, M. (2007). Design principles of MIMO radar detectors. *IEEE Transactions on Aerospace and Electronic Systems* 43(3): 886–898.

22 Aubry, A., Lops, M., Tulino, A.M., and Venturino, L. (2010). On MIMO detection under non-Gaussian target scattering. *IEEE Transactions on Information Theory* 56(11): 5822–5838.

23 Grossi, E. and Lops, M. (2012). Space-time code design for MIMO detection based on Kullback-Leibler divergence. *IEEE Transactions on Information Theory* 58(6): 3989–4004.

24 Naghsh, M.M., Modarres-Hashemi, M., Kerahroodi, M.A., and Alian, E.H.M. (2017). An information theoretic approach to robust constrained code design for MIMO radars. *IEEE Transactions on Signal Processing* 65(14): 3647–3661.

25 Tang, B., Tang, J., and Peng, Y. (2010). MIMO radar waveform design in colored noise based on information theory. *IEEE Transactions on Signal Processing* 58(9): 4684–4697.

26 Tang, B., Naghsh, M.M., and Tang, J. (2015). Relative entropy-based waveform design for MIMO radar detection in the presence of clutter and interference. *IEEE Transactions on Signal Processing* 63(14): 3783–3796.

27 Tang, B., Zhang, Y., and Tang, J. (2018). An efficient minorization maximization approach for MIMO radar waveform optimization via relative entropy. *IEEE Transactions on Signal Processing* 66(2): 400–411.

28 Tang, B. and Li, J. (2019). Spectrally constrained MIMO radar waveform design based on mutual information. *IEEE Transactions on Signal Processing* 67(3): 821–834.

29 Grossi, E., Lops, M., and Venturino, L. (2011). Robust waveform design for MIMO radars. *IEEE Transactions on Signal Processing* 59(7): 3262–3271.

30 Kay, S. (1998). *Fundamentals of Statistical Signal Processing-Volume II: Detection Theory*. New Jersey: Prentice Hall.

31 Huleihel, W., Tabrikian, J., and Shavit, R. (2013). Optimal adaptive waveform design for cognitive MIMO radar. *IEEE Transactions on Signal Processing* 61(20): 5075–5089.

32 Herbert, S., Hopgood, J.R., and Mulgrew, B. (2018). MMSE adaptive waveform design for active sensing with applications to MIMO radar. *IEEE Transactions on Signal Processing* 66(5): 1361–1373.

33 Gurbuz, S.Z., Griffiths, H.D., Charlish, A. et al. (2019). An overview of cognitive radar: past, present, and future. *IEEE Aerospace and Electronic Systems Magazine* 34(12): 6–18.

34 Carson, W.R., Chen, M., Rodrigues, M.R.D. et al. (2012). Communications-inspired projection design with application to compressive sensing. *SIAM Journal on Imaging Sciences* 5(4): 1185–1212.

35 Tang, B., Zhang, N., Zhang, S., and Huang, Z. (2019). MMSE-based waveform design for the distributed MIMO radar in spectrally crowded environments. *The Journal of Engineering.* 2019(20): 6603–6607.

36 Hunter, D.R. and Lange, K. (2004). A tutorial on MM algorithms. *American Statistician* 58(1): 30–37.

37 Tang, B. and Stoica, P. (2021). Information-theoretic waveform design for MIMO radar detection in range-spread clutter. *Signal Processing* 182: 107961.

38 Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge: Cambridge University Press.

39 Hjorungnes, A. and Gesbert, D. (2007). Complex-valued matrix differentiation: techniques and key results. *IEEE Transactions on Signal Processing* 55(6): 2740–2746.

40 Tropp, J.A., Dhillon, I.S., Heath, R.W., and Strohmer, T. (2005). Designing structured tight frames via an alternating projection method. *IEEE Transactions on Information Theory* 51(1): 188–209.

41 Tang, B., Li, J., and Liang, J. (2018). Alternating direction method of multipliers for radar waveform design in spectrally crowded environments. *Signal Processing* 142: 398–402.

42 Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any EM algorithm. *Scandinavian Journal of Statistics* 35(2): 335–353.