Probabilistically Robust Joint Transmit Code and Receive Filter Design for Extended Targets: A Mixed-Integer Programming Approach

Geng Chen, Chunyang Wang, Xuejing Zhang, *Member, IEEE*, Jian Gong, and Ming Tan

Abstract—This paper proposes a probabilistically robust joint transmit code and receive filter design approach for extended targets detection. The goal is to maximize the probability that the radar output signal-to-interference-pulse-noise ratio exceeds a given threshold, under the target impulse response with a complex Gaussian distribution. By restricting the transmit waveform to discrete phase codes, we construct a joint transmit-receive vector with respect to the transmit code and receive filter. Based on the joint vector, the probabilistically robust joint design problem is reformulated as a Mixed-Integer Programming (MIP) problem. Subsequently, we linearize the nonlinear mixed-integer constraint using a linearization technique and derive an approximate analytical expression for the intractable probabilistic objective function. Finally, the MIP problem is solved using the bisection method, yielding high quality (sub-optimal) solutions. Numerical experiments are presented to demonstrate the effectiveness and superiority of the proposed method in extended targets detection.

Index Terms—Radar waveform design, receive filter, extended targets detection, probabilistically robust optimization, mixed-integer programming.

I. INTRODUCTION

TARGET detection is a critical requirement for radar systems, and the Signal-to-Interference-pulse-Noise Ratio (SINR) of the echo directly affects the detection probability [1]–[3]. To enhance radar detection performance in different environments, it is advantageous to leverage the degrees of freedom available in both the transmitter and receiver. This allows for flexible design of the Transmit Waveform (TW) and Receive Filter (RF), resulting in an improved output SINR of the radar [4], [5].

Over the past few decades, there has been extensive research on the joint design problem of TW and RF in various scenarios. In the presence of signal-dependent interference (clutter) scenarios, an Alternating Iterative Optimization (AIO) technique-based method was proposed in [6] to jointly design the robust TW and RF. Moreover, the AIO technique has also been widely used in other studies to obtain local optimal solutions for the problem of joint design of TW and RF, including multiple-input multiple-output space-time adaptive processing systems [7], dual-function radar-communication systems [8], [9], and so on [10]–[16]. Note that the AIO technique optimizes only a subset of variables in each iteration. This characteristic carries the risk of converging to a local optimum and may not yield high quality (sub-optimal) solutions. Additionally, it is worth mentioning that

the aforementioned methods [6]–[16] primarily investigate the joint design problem of TW and RF for point targets.

With the improvement of radar range resolution, extended targets have become commonplace in practical radar detection. In contrast to point targets, extended targets occupy multiple range bins, resulting in target echoes that are the convolution of the Target Impulse Response (TIR) and TW [17], [18]. Consequently, traditional processing methods designed for point targets may not be the best choices for extended targets. Over the past decades, numerous techniques have been developed to address the design of TW and RF specifically for extended targets detection [19]–[27]. Some studies have focused on designing TW and RF under deterministic TIR assumptions [18]–[21]. For example, the authors of [19] generalized the optimal TW to accommodate the extended targets detection. In [20], the AIO technique was employed to jointly optimize the TW and RF, with the aim of maximizing the mathematical expectation of the output SINR. Nevertheless, it is crucial to highlight that these methods require precise knowledge of the TIR information for extended targets.

In practice, acquiring precise information about the TIR before detection is impractical due to its high sensitivity to line-of-sight (LOS) [28]. As a result, the aforementioned methods [18]–[21] exhibit poor performance in scenarios with uncertain TIR. To circumvent this difficulty, several studies have proposed methods for designing robust TW and RF under uncertain TIR (see [22]–[27]). For instance, in [22], a worst-case approach was utilized to design the robust TW for extended targets. The authors of [23] proposed a robust joint design method for TW and RF in clutter environments, which used AIO techniques to optimize the TW and RF alternatively. Notice that these methods [22], [23] model the uncertain TIR as a compact convex set to optimize TW and RF, and use the worst-case performance metric as the objective function. Nevertheless, the worst-case performance metric may be too conservative and does not adequately capture the distribution structure of the output SINR.

In general, the TIR is a random vector that follows a specific distribution, such as complex Gaussian distribution. Accordingly, several studies have investigated the problem of designing TW and RF when the TIR follows a complex Gaussian distribution (see [24]–[27]). Xu *et al.* [26] proposed a probabilistically robust TW design method for extended targets detection. Specifically, the probabilistically robust method aims to optimize the TW by maximizing the probability that the output SINR exceeds a given threshold,

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2024.3471493

2

and its performance is evaluated using the Probabilistically Robust Detection (PRD) metric. For the joint design problem of TW and RF, Xu *et al.* [27] investigated the design of radar waveform-filter for extended targets based on the PRD criterion, and utilized the Block Coordinate Descent (BCD) technique to optimize TW and RF alternately. However, these methods [26], [27] do not provide an explicit analytical expression for the Probabilistic Objective Function (POF). Instead, they rely on geometric approximation and limit interpretation techniques to approximate the POF. More importantly, these methods only focus on the continuous phase for the TW.

From a practical point of view, the TW should consist of discrete phase codes [29]–[31]. In [30], a constant modulus discrete phase radar waveform design method was proposed to optimize radar detection performance. In [31], a discrete phase-coded sequence design method was introduced, aiming to minimize the sidelobe level of the ambiguity function in coherent waveform-agile radar systems. However, these discrete-phase waveform design methods incorporate approximate steps, leading to the inability to obtain high-quality solutions from a given discrete-phase alphabet.

In this paper, we address the problem of probabilistically robust joint design of transmit code (TC) and RF for extended targets detection. The goal is to maximize the probability that the radar output SINR exceeds a given threshold. By restricting the TW to discrete phase codes, we propose a Mixed-Integer Programming (MIP) approach to solve the probabilistically robust joint design problem and find high quality TC and RF. Numerical experiment results demonstrate the effectiveness and superiority of the proposed method. To summarize, Tab. I succinctly outlines the key features of this paper and the research work related to extended targets detection. The research contributions of our work are summarized as follows:

1) Probabilistically robust joint TC and RF design: Considering the TIR with a complex Gaussian distribution and the PRD metric in [26], [27], we address the problem of probabilistically robust joint design for TC and RF. Different from the studies in [27], we consider a finite number of bits in the digital waveform generators and restrict the TW to discrete phase codes.

2) An MIP approach for the joint design of TC and RF: We construct a transmit-receive joint vector with respect to the TC and the RF. The joint vector is then utilized to reformulate the probabilistically robust design problem as an MIP problem. In contrast to the AIO [23] and BCD [27] techniques, the proposed MIP approach optimizes the TC and RF simultaneously, resulting in high quality solutions.

3) A tractable approximate expression for the probabilistic objective function: For the intractable POF, we employ relaxation processing techniques to derive a tractable approximate expression. This approach differs from that presented in previous studies [26], [27], which utilize geometric interpretation and limit interpretation techniques to approximate the POF.

The rest of the paper is organized as follows. Section II presents the signal model and formulates the problem of probabilistically robust joint design for TC and RF. In Section III, we propose an MIP approach for solving this problem. Numerical experiments are used to validate the performance of the proposed method in Section IV. Finally, Section V concludes the paper.

Notations: In this paper, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\dagger}$ denote the transpose and the conjugate transpose of matrix or vector, respectively. We use boldface to represent vectors $\boldsymbol{a}$ (lower case) and matrices $\boldsymbol{A}$ (upper case), respectively. $a\,(p)$ represents the p-th element of the vector $\boldsymbol{a}$, and $A\,(p,q)$ denotes the element at the p-th row and q-th column of the matrix $\boldsymbol{A}$. $\mathrm{vec}\,(\boldsymbol{A})$

TABLE I
A BRIEFLY SUMMARY OF THE EXTENDED TARGETS DETECTION.

Ref. no.	Optimization variable	Optimization criteria	Constant modulus constraint	Discrete phase constraint	TIR model	Optimization algorithm	Contribution	Drawbacks
[20]	TW and RF	Maximize output SINR	×	×	Deterministic model	AIO	First joint optimization of TW and RF.	Sensitive to TIR error, TW is not discrete phase code.
[23]	TW and RF	Maximize worst-case performance	√	×	Compact convex set	AIO	Proposed robust joint TW and RF design.	The probability distribution of the TIR error is not considered, the AIO algorithm tends to fall into local optimum, TW is not discrete phase code.
[26]	TW	Maximize POF	√	×	Complex Gaussian distribution	Dinkelbach's algorithm	Proposed PRD criterion.	No optimized filter, TW is not discrete phase code.
[27]	TW and RF	Maximize POF	√	×	Complex Gaussian distribution	BCD	Using the PRD criterion to jointly optimize TW and RF.	The BCD algorithm tends to fall into local optimum, TW is not discrete phase code.
Proposed	TW and RF	Maximize POF	√	√	Complex Gaussian distribution	MIP	Using MIP to optimize TW and RF simultaneously.	Requires higher computational costs.
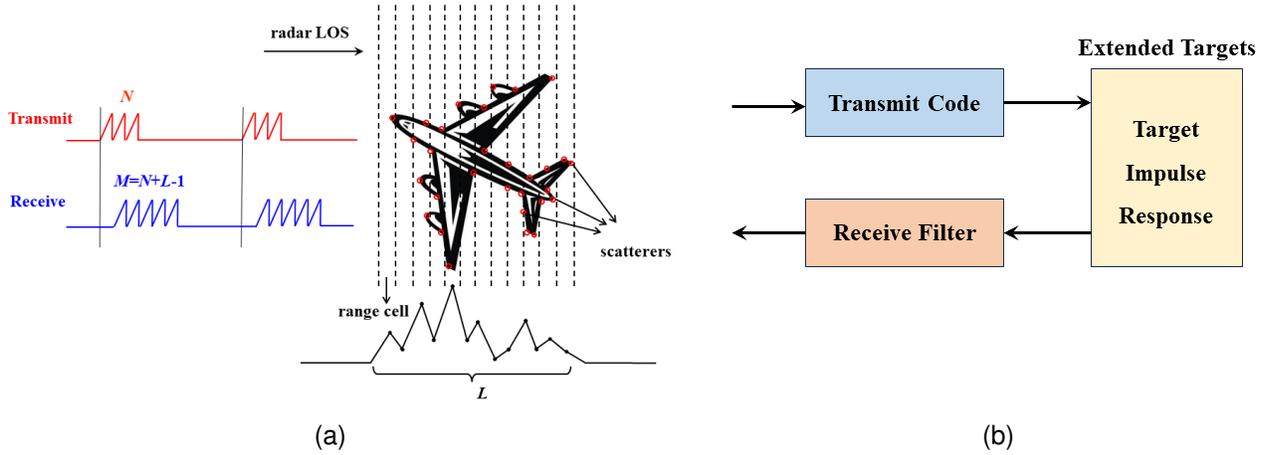
Fig. 1. The architecture and detection mechanisms of the considered extended targets. (a) The return from the extended targets scattering center are projected on the LOS. (b) Transmit-receive processing procedure.

represents the vectorization of the matrix $\boldsymbol{A}$. The modulus of the complex number a is denoted by the $|a|$. $\|\boldsymbol{a}\|$ denotes the Euclidean norm of vector $\boldsymbol{a}$. $\mathbf{1}_N$ denotes $N \times 1$ dimensional vector with all elements of one, and $\boldsymbol{I}_N$ represents the $N \times N$ identity matrix. $\otimes$ represents Kronecker product, $\oplus$ is the convolution operator. $\boldsymbol{a} \preceq \boldsymbol{b}$ and $\boldsymbol{a} \succeq \boldsymbol{b}$ are defined as componentwise inequality between vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. This means that $a(p) \leq b(p)$ and $a(p) \geq b(p)$ for every p index. $\boldsymbol{a} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Xi})$ indicates that the random vector $\boldsymbol{a}$ follows a Gaussian distribution with mean zero and variance $\boldsymbol{\Xi}$. $\mathbb{R}^{N \times M}$ and $\mathbb{C}^{N \times M}$ refer to a real and complex valued matrix of size $N \times M$, respectively. The expectation of the random variable x is denoted as $\mathbb{E}[x]$, while its standard deviation is represented by $\mathbb{D}[x]$.

II. SIGNAL MODEL AND PROBLEM FORMULATION

A. Signal Model

Let us consider a high-resolution radar that transmits the N-dimensional fast-time signal $\boldsymbol{s} = [s(0), ..., s(N-1)]^{\mathrm{T}} \in \mathbb{C}^N$. For the given extended target with L lengths of occupied range cells[1], let $\boldsymbol{t} = [t(0), ..., t(L-1)]^{\mathrm{T}} \in \mathbb{C}^L$ denote the discrete samples of the TIR[2]. The extended targets architecture and detection mechanisms are shown in Fig. 1. In the presence of clutter, after the transmit signal is reflected by the extended target, the received discrete-time baseband signal captured by the radar can be expressed as:

$$y(i) = \sum_{k=0}^{N-1}[t(i-k) + c(i-k)]s(k) + v(i)$$
$$= [t(i) + c(i)] \oplus s(i) + v(i) \quad (1)$$

[1]The parameter L can be chosen based on the maximum target length and the radar range resolution.

[2]The TIR characterizes the scattering behavior of the target and is determined by the target physical parameters and the radar operating parameters [32].

where i is the discrete time index, $t(i) = 0$ unless $i \in \{0, ..., L-1\}$, and $v(i)$ is the complex additive white noise signal. $c(i)$ denotes the Clutter Impulse Response (CIR).

Considering N transmit codes, we collect all received signals $y(i)$ from the extended target into a vector $\boldsymbol{y} = [y(0), ..., y(M-1)]^T \in \mathbb{C}^M$. This can be expressed as:

$$\boldsymbol{y} = \boldsymbol{T}\boldsymbol{s} + \boldsymbol{C}\boldsymbol{s} + \boldsymbol{v} \quad (2)$$

where $M = N + L - 1$ denotes the number of discrete time observations, $\boldsymbol{v} = [v(0), ..., v(M-1)]^{\mathrm{T}} \in \mathbb{C}^M$ is the noise vector. We assume that $\boldsymbol{v}$ follows a complex circular white Gaussian vector with zero mean and covariance matrix $\sigma_n^2 \boldsymbol{I}$, i.e., $\boldsymbol{v} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \boldsymbol{I})$, σ_n^2 represents the variance of each filtered noise sample. $\boldsymbol{T}$ denotes the TIR matrix, defined as:

$$\boldsymbol{T} = \sum_{i=0}^{L-1} t(i)\boldsymbol{J}_i \quad (3)$$

and $\boldsymbol{C}$ represents the CIR matrix, defined as:

$$\boldsymbol{C} = \sum_{i=-N+1}^{M-1} c(i)\boldsymbol{J}_i \quad (4)$$

where $\boldsymbol{J}_i \in \mathbb{R}^{M \times N}$ stands for the shift matrix, defined by

$$\boldsymbol{J}_i(k_1, k_2) = \begin{cases} 1, \text{if } k_1 - k_2 = i \\ 0, \text{if } k_1 - k_2 \neq i \end{cases} \quad (5)$$
$$k_1 = \{1, ..., M\}, k_2 = \{1, ..., N\}$$

Since the TIR information is not precisely known before detection, we assume that $\boldsymbol{t}$ follows a complex Gaussian distribution as $\boldsymbol{t} \sim \mathcal{CN}(\boldsymbol{t}_0, \boldsymbol{G}_t)$ (similar assumption can also be found in [25]–[27]). Here, $\boldsymbol{t}_0$ denotes the prior knowledge of the TIR, and $\boldsymbol{G}_t$ represents the uncertainty error information and correlations of the elements in $\boldsymbol{t}$. In this paper, similar to [26], we assume that $\boldsymbol{t}_0$ and $\boldsymbol{G}_t$ are known, and they can be obtained from previous detections or the target database using cognitive methods [33], [34]. Under this assumption, we define

$$\boldsymbol{t} = \boldsymbol{t}_0 + \boldsymbol{\varepsilon} \quad (6)$$

where $\boldsymbol{\varepsilon}$ denotes the error vector, and it follows a complex Gaussian distribution as $\boldsymbol{\varepsilon} \sim \mathcal{CN}\left(\mathbf{0}, \boldsymbol{G}_t\right)$. Then, the TIR matrix can be reformulated as:

$$\boldsymbol{T} = \boldsymbol{T}_0 + \boldsymbol{T}_\varepsilon \tag{7}$$

where $\boldsymbol{T}_0$ and $\boldsymbol{T}_\varepsilon$ are respectively defined as:

$$\boldsymbol{T}_0 = \sum_{i=0}^{L-1} t_0(i) \boldsymbol{J}_i \tag{8}$$

and

$$\boldsymbol{T}_\varepsilon = \sum_{i=0}^{L-1} \varepsilon(i) \boldsymbol{J}_i \tag{9}$$

B. Problem Formulation

Assuming that the received vector $\boldsymbol{y}$ is filtered through $\boldsymbol{w} = [w(0), ..., w(M-1)]^{\mathrm{T}} \in \mathbb{C}^M$, the output SINR of the filter can be expressed as:

$$\gamma\left(\boldsymbol{s}, \boldsymbol{w}\right) = \frac{\left|\boldsymbol{w}^\dagger \boldsymbol{T} \boldsymbol{s}\right|^2}{\left|\boldsymbol{w}^\dagger \tilde{\boldsymbol{C}} \boldsymbol{s}\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \tag{10}$$

where $\tilde{\boldsymbol{C}}$ is the expectation information for the CIR matrix, $\tilde{\boldsymbol{C}}$ and σ_n^2 can be obtained from previous detections [26], [27].

Since the TIR is a complex Gaussian random vector, the output SINR $\gamma\left(\boldsymbol{s}, \boldsymbol{w}\right)$ is also a random variable. In this scenario, there is no guarantee that the detection performance of the radar will always remain satisfactory. To ensure the detection of a target of interest, it is generally desirable for the $\gamma\left(\boldsymbol{s}, \boldsymbol{w}\right)$ to exceed a given threshold γ_0[3]. In the stochastic cases, we can design the TW and RF to maximize the probability achieving this event. Therefore, the design of the TW and RF should adhere the following design principle (similar principle as in [26], [27]):

$$\max_{\boldsymbol{s}, \boldsymbol{w}} \Pr\left\{\gamma\left(\boldsymbol{s}, \boldsymbol{w}\right) \geq \gamma_0\right\} \tag{11}$$

Taking into account a limited number of bits in the digital waveform generators, we restrict the TW to discrete phase codes (similar constraint can be found in [29], [30]). This means that the values of $s(n)$ must belong to a discrete alphabet set. Mathematically, assuming that B denotes the number of quantization bits and defining $Q = 2^B$, the TC must satisfy the following constraint:

$$s(n) \in \mathbb{S} \tag{12}$$

where $\mathbb{S} = \left\{1, e^{j2\pi/Q}, ..., e^{j2\pi(Q-1)/Q}\right\}$ is the discrete alphabet.

By combining constraint (12), the probabilistically robust joint design problem for TC and RF can be formulated as:

$$\max_{\boldsymbol{s}, \boldsymbol{w}} \Pr\left\{\gamma(\boldsymbol{s}, \boldsymbol{w}) \geq \gamma_0\right\} \tag{13a}$$

$$\text{s.t. } s(n) \in \mathbb{S} \tag{13b}$$

$$\|\boldsymbol{w}\| \leq 1/\sqrt{N} \tag{13c}$$

---

[3]The SINR threshold γ_0 is chosen based on the desired detection probability P_0 for a given false alarm rate P_{fa}.

In this problem, constraint (13b) ensures that the TC is selected only from the given discrete alphabet $\mathbb{S}$. Additionally, for ease of subsequent processing, we add the constraint (13c) to limit the Euclidean norm of $\boldsymbol{w}$. It is worth noting that the constant $1/\sqrt{N}$ in constraint (13c) can be assigned any positive number without impacting the optimal solutions of problem (13). In the next section, we devise an MIP approach to tackle problem (13).

III. PROBABILISTICALLY ROBUST JOINT TC AND RF DESIGN

In this section, we begin by reformulating problem (13) as an MIP problem. To solve this problem, we proceed to linearize the nonlinear mixed-integer constraint and convert it into a set of linear constraints. Additionally, we provide a tractable approximate expression for the POF. Finally, the reformulated MIP problem is solved using the bisection method, resulting in high quality TC and RF.

To begin with, we define the vector $\boldsymbol{p}$ to represent the given discrete alphabet $\mathbb{S}$, given by

$$\boldsymbol{p} = \left[1, e^{j2\pi/Q}, ..., e^{j2\pi(Q-1)/Q}\right]^{\mathrm{T}} \in \mathbb{C}^Q \tag{14}$$

Then, by introducing an $N \times Q$ binary matrix $\boldsymbol{D}$ to determine the TC, we can express the TC $\boldsymbol{s}$ as:

$$\boldsymbol{s} = \boldsymbol{D}\boldsymbol{p} \tag{15}$$

Note that the binary matrix $\boldsymbol{D}$ must satisfy the following constraint:

$$\boldsymbol{D}\mathbf{1}_Q = \mathbf{1}_N \tag{16}$$

which ensures that exactly one element of each row in matrix $\boldsymbol{D}$ is one. In this way, the discrete phase codes constraint (13b) is satisfied. Now, the design for TC $\boldsymbol{s}$ can be transformed into the design of binary matrix $\boldsymbol{D}$, since the vector $\boldsymbol{p}$ is preassigned.

By combining the constraints (15) and (16), we can reformulate the problem (13) as follows:

$$\max_{\boldsymbol{s}, \boldsymbol{w}, \boldsymbol{D}} \Pr\left\{\frac{\left|\boldsymbol{w}^\dagger \boldsymbol{T} \boldsymbol{s}\right|^2}{\left|\boldsymbol{w}^\dagger \tilde{\boldsymbol{C}} \boldsymbol{s}\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \geq \gamma_0\right\} \tag{17a}$$

$$\text{s.t. } \boldsymbol{s} = \boldsymbol{D}\boldsymbol{p} \tag{17b}$$

$$\boldsymbol{D}\mathbf{1}_Q = \mathbf{1}_N \tag{17c}$$

$$\|\boldsymbol{w}\| \leq 1/\sqrt{N} \tag{17d}$$

$$\boldsymbol{D} \in \{0,1\}^{N \times Q} \tag{17e}$$

Notice that in problem (17), the matrix $\boldsymbol{D}$ is an $N \times Q$ binary matrix, meaning its elements can only take the values of zero or one as indicated in (17e). In fact, the problem (17) is an MIP problem, where some variables (such as matrix $\boldsymbol{D}$) can only taken on integer values, while others (such as vector $\boldsymbol{w}$) can taken on continuous values. For the above MIP problem, it can be solved using universal MIP solvers such as GUROBI [35] and CPLEX [36]. However, problem (17) involves an intractable expression of the POF. Unlike existing AIO [23] and BCD [27] techniques, we will propose an MIP approach to find sub-optimal solutions for problem (17) in this section.

A. Joint Design of TC and RF Using Mixed-Integer Programming

In this subsection, we construct a transmit–receive joint vector with respect to the TC and RF. This joint vector is subsequently utilized to reformulate problem (17) as a new MIP problem.

By utilizing equation (15) and introducing an identity matrix $\boldsymbol{I}_N$, the target echo power $\left|\boldsymbol{w}^\dagger \boldsymbol{T} \boldsymbol{s}\right|^2$ can be rewritten as

$$\left|\boldsymbol{w}^\dagger \boldsymbol{T} \boldsymbol{s}\right|^2 = \left|\boldsymbol{w}^\dagger \boldsymbol{T} \boldsymbol{I}_N \boldsymbol{D} \boldsymbol{p}\right|^2 \tag{18a}$$

$$= \left|\boldsymbol{w}^\dagger \boldsymbol{T} \left(\boldsymbol{p}^\mathrm{T} \otimes \boldsymbol{I}_N\right) \mathrm{vec}\left(\boldsymbol{D}\right)\right|^2 \tag{18b}$$

$$= \left|\boldsymbol{w}^\dagger \boldsymbol{H}_t \mathrm{vec}\left(\boldsymbol{D}\right)\right|^2 \tag{18c}$$

$$= \left|\left(\mathrm{vec}\left(\boldsymbol{D}\right)^\mathrm{T} \otimes \boldsymbol{w}^\dagger\right) \mathrm{vec}\left(\boldsymbol{H}_t\right)\right|^2 \tag{18d}$$

$$= \left|\boldsymbol{f}^\dagger \boldsymbol{h}_t\right|^2 \tag{18e}$$

where

$$\boldsymbol{f} = \mathrm{vec}\left(\boldsymbol{D}\right) \otimes \boldsymbol{w} \in \mathbb{C}^{NQM} \tag{19a}$$

$$\boldsymbol{h}_t = \mathrm{vec}\left(\boldsymbol{H}_t\right) \in \mathbb{C}^{NQM} \tag{19b}$$

$$\boldsymbol{H}_t = \boldsymbol{T} \left(\boldsymbol{p}^\mathrm{T} \otimes \boldsymbol{I}_N\right) \in \mathbb{C}^{M \times NQ} \tag{19c}$$

The equations (18b) and (18d) are derived from the properties of matrix multiplication, i.e., for any matrix $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$, it holds that $\mathrm{vec}(\boldsymbol{A}_1 \boldsymbol{A}_2 \boldsymbol{A}_3) = \left(\boldsymbol{A}_3^\mathrm{T} \otimes \boldsymbol{A}_1\right) \mathrm{vec}\left(\boldsymbol{A}_2\right)$.

Similar to equation (18), the clutter power $\left|\boldsymbol{w}^\dagger \tilde{\boldsymbol{C}} \boldsymbol{s}\right|^2$ can be rewritten as

$$\left|\boldsymbol{w}^\dagger \tilde{\boldsymbol{C}} \boldsymbol{s}\right|^2 = \left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 \tag{20}$$

where $\boldsymbol{h}_c = \mathrm{vec}\left(\boldsymbol{H}_c\right) \in \mathbb{C}^{NQM}$ and $\boldsymbol{H}_c = \tilde{\boldsymbol{C}} \left(\boldsymbol{p}^\mathrm{T} \otimes \boldsymbol{I}_N\right) \in \mathbb{C}^{M \times NQ}$. Meanwhile, we observe that the TC $\boldsymbol{s}$ and the RF $\boldsymbol{w}$ are constructed as the transmit–receive joint vector $\boldsymbol{f}$. This implies that $\boldsymbol{s}$ and $\boldsymbol{w}$ can be optimized simultaneously by optimizing $\boldsymbol{f}$, as long as $\boldsymbol{f}$ is ensured to have the decomposition form of (19a).

By combining the equations (18e) and (20), we can reformulate the output SINR γ as:

$$\gamma = \frac{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_t\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \tag{21}$$

As described earlier in equation (7), there is an error $\boldsymbol{T}_\varepsilon$ in the TIR matrix $\boldsymbol{T}$. Based on the definitions of $\boldsymbol{h}_t$, it can be observed that $\boldsymbol{h}_t$ also has a random error vector $\boldsymbol{\delta}$. In this context, we define

$$\boldsymbol{h}_t = \bar{\boldsymbol{h}}_t + \boldsymbol{\delta} \tag{22}$$

where $\bar{\boldsymbol{h}}_t = \mathrm{vec}\left(\bar{\boldsymbol{H}}_t\right)$ and $\bar{\boldsymbol{H}}_t = \boldsymbol{T}_0 \left(\boldsymbol{p}^\mathrm{T} \otimes \boldsymbol{I}_N\right)$.

Similar to the calculation of $\bar{\boldsymbol{h}}_t$, the error vector $\boldsymbol{\delta}$ can be calculated by

$$\boldsymbol{\delta} = \mathrm{vec}\left(\boldsymbol{H}_\varepsilon\right) \tag{23}$$

where $\boldsymbol{H}_\varepsilon = \boldsymbol{T}_\varepsilon \left(\boldsymbol{p}^T \otimes \boldsymbol{I}_N\right)$.

For this consideration, the output SINR is reformulated as:

$$\gamma = \frac{\left|\boldsymbol{f}^\dagger \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta}\right)\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \tag{24}$$

Now, the problem (17) can be recast as:

$$\max_{\boldsymbol{f}, \boldsymbol{D}, \boldsymbol{w}} \Pr\left\{\frac{\left|\boldsymbol{f}^\dagger \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta}\right)\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \geq \gamma_0\right\} \tag{25a}$$

$$\text{s.t.} \quad \boldsymbol{f} = \mathrm{vec}\left(\boldsymbol{D}\right) \otimes \boldsymbol{w} \tag{25b}$$

$$\boldsymbol{D} \boldsymbol{1}_Q = \boldsymbol{1}_N \tag{25c}$$

$$\boldsymbol{D} \in \{0, 1\}^{N \times Q} \tag{25d}$$

$$\|\boldsymbol{f}\| \leq 1 \tag{25e}$$

Compared to problem (17), the optimization now focuses on the variable $\boldsymbol{f}$ rather than $\boldsymbol{s}$, and the new objective function is not directly related to $\boldsymbol{s}$. As a result, the constraint (17b) is removed, and two new constraints (25b) and (25e) are added. The constraint (25b) is added to ensure that the values of the variable $\boldsymbol{f}$ are associated with both $\boldsymbol{D}$ and $\boldsymbol{w}$. The constraint (25e) is derived based on the constraints (25b)-(25d) and (17d), and it is used to restrict the values of $\boldsymbol{f}$, similar to constraint (17d). It should be noted that constraint (17d) is also omitted in problem (25). As a matter of fact, the omitted constraint (17d) is equivalent to the constraint (25e). If constraints (25b)-(25e) are satisfied, it ensures that $\boldsymbol{w}$ automatically satisfies constraint (17d). Hence, the constraint $\|\boldsymbol{w}\| \leq 1/\sqrt{N}$ is implicitly satisfied by the constraints of problem (25) and can be omitted.

From problem (25), it is evident that the optimal solutions of the binary matrix $\boldsymbol{D}$ and RF $\boldsymbol{w}$ can be directly obtained by solving for the transmit-receive joint vector $\boldsymbol{f}$, which is different from the existing AIO [23] and BCD [27] techniques. However, there are two main challenges in solving problem (25). The first difficulty arises from the nonlinear constraint (25b). The optimal solutions for the variables $\boldsymbol{f}$, $\boldsymbol{D}$, and $\boldsymbol{w}$ must satisfy this constraint. The second difficulty lies in the intractable POF (25a). To address these challenges, the following subsections discuss how to handle the nonlinear constraint and how to relax the POF into a tractable expression.

Remark 1: Based on equation (9), we know that $\boldsymbol{T}_\varepsilon$ is constructed based on the error vector $\boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}$ follows a complex Gaussian distribution $\boldsymbol{\varepsilon} \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_t\right)$. Since linear transformations do not change the probability distribution characteristics, the error vector $\boldsymbol{\delta}$ also follows a complex Gaussian distribution. We assume that the error vector $\boldsymbol{\delta} \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_\delta\right)$. In practice applications, we can generate a large number of error vectors $\boldsymbol{\varepsilon}$ using the distribution $\mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_t\right)$ and compute $\boldsymbol{\delta}$ using equation (23). By collecting a significant number of samples of $\boldsymbol{\delta}$, we can calculate the covariance matrix $\boldsymbol{G}_\delta$.

B. Mixed-Integer Constraint Linearization

To address the nonlinear constraint (25b), we first introduce an important lemma related to MIP problems.

Lemma 1 [37]: For two real variables $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$, where $x_1 \in \{0, 1\}$ is a binary variable and $x_2 \in [a, b]$, the

nonlinear constraint $x = x_1 x_2$ can be linearized as:

$$x \leq x_2 + K(1 - x_1) \tag{26a}$$

$$x \geq x_2 - K(1 - x_1) \tag{26b}$$

$$ax_1 \leq x \leq bx_1 \tag{26c}$$

$$a \leq x_2 \leq b \tag{26d}$$

$$x_1 \in \{0, 1\} \tag{26e}$$

where K is a sufficiently large positive number. To facilitate further discussions, we provide a brief proof of *Lemma 1*.

Proof: We can observe that the constraint (26a) restricts x to be less than x_2, and the constraint (26b) restricts x to be greater than x_2 when $x_1 = 1$. Moreover, the constraint (26c) limits the values of x within the interval $[a, b]$. As a result, if $x_1 = 1$, x can only take the value x_2. On the other hand, it is not hard to observe that the constraints (26a) and (26b) fail when $x_1 = 0$, since K is a sufficiently large positive number. Meanwhile, the constraint (26c) enforces x to be zero. With these observations, we can conclude that constraint (26) is equivalent to the constraint $x = x_1 x_2$.

Next, considering that the nonlinear constraint is expressed as the Kronecker product of two vectors, we can utilize *Lemma 1* to derive the following corollary.

Corollary 1: For the vectors $\boldsymbol{g}_1 \in \mathbb{R}^N$ and $\boldsymbol{g}_2 \in \mathbb{R}^M$, where $\boldsymbol{g}_1$ is a binary vector, i.e., $\boldsymbol{g}_1 \in \{0, 1\}^N$, and $g_2(m) \in [a, b]$, the nonlinear constraint $\boldsymbol{g} = \boldsymbol{g}_1 \otimes \boldsymbol{g}_2$ can be linearized as

$$\boldsymbol{g} \preceq \mathbf{1}_N \otimes \boldsymbol{g}_2 + K(\mathbf{1}_{NM} - \boldsymbol{g}_1 \otimes \mathbf{1}_M) \tag{27a}$$

$$\boldsymbol{g} \succeq \mathbf{1}_N \otimes \boldsymbol{g}_2 - K(\mathbf{1}_{NM} - \boldsymbol{g}_1 \otimes \mathbf{1}_M) \tag{27b}$$

$$a(\boldsymbol{g}_1 \otimes \mathbf{1}_M) \preceq \boldsymbol{g} \preceq b(\boldsymbol{g}_1 \otimes \mathbf{1}_M) \tag{27c}$$

$$a\mathbf{1}_M \preceq \boldsymbol{g}_2 \preceq b\mathbf{1}_M \tag{27d}$$

$$\boldsymbol{g}_1 \in \{0, 1\}^N \tag{27e}$$

where K also is a sufficiently large positive number. The proof of *Corollary 1* follows a similar approach to that of *Lemma 1*, and for the sake of brevity, we omit the detailed proof here.

Now, let us address the intractable nonlinear constraint (25b). Notice that $\text{vec}(\boldsymbol{D})$ is a binary vector that corresponds to $\boldsymbol{g}_1$, while $\boldsymbol{w}$ is a complex vector that does not directly correspond to $\boldsymbol{g}_2$ in *Corollary 1*. However, it is observed that the real and imaginary parts of $\boldsymbol{w}$ can be considered independently. For this, the constraint (25b) is written as:

$$\text{Re}\{\boldsymbol{f}\} = \text{vec}(\boldsymbol{D}) \otimes \text{Re}\{\boldsymbol{w}\} \tag{28a}$$

$$\text{Im}\{\boldsymbol{f}\} = \text{vec}(\boldsymbol{D}) \otimes \text{Im}\{\boldsymbol{w}\} \tag{28b}$$

Based on the implicit constraint $\|\boldsymbol{w}\| \leq 1/\sqrt{N}$ in problem (25), it is clear that both the real and imaginary parts of $\boldsymbol{w}$ must be in the interval $[-1, 1]$. Consequently, we assign $a = -1$, $b = 1$ in the constraints (27c) and (27d). Then, constraints (28a) and (28b) (i.e., constraint (25b)) can be linearized individually using *Corollary 1*, and problem (25) can be recast into problem (29) (shown at the top of the next page).

Compared to problem (25), in problem (29), the nonlinear constraint (25b) is now transformed into a set of tractable linear constraints (29b)-(29g). It is worth noting that there are no separate constraints imposed on the real and imaginary

parts of $\boldsymbol{w}$ in problem (29), similar to constraint (27d). This is because the fact that the implicit constraint $\|\boldsymbol{w}\| \leq 1/\sqrt{N}$ already imposes limitations on the feasible range of $\boldsymbol{w}$. For this reason, there is no need to introduce additional constraints to restrict the real and imaginary parts of $\boldsymbol{w}$ in problem (29). Finally, we set $K = 2$ in problem (29) (This will be explained in *Remark 2*).

Remark 2: Observing constraints (26a) and (26b), we know that the value of x must lie within the interval $[a, b]$. Therefore, for any $\forall x_2 \in [a, b]$, the value of K need to satisfy the following inequalities:

$$x_2 - K \leq a \tag{30a}$$

$$x_2 + K \geq b \tag{30b}$$

Based on (30), it is not hard to derive that K must take a positive number satisfying the inequality $K \geq b - a$. This conclusion can be applied to problem (29), so we set $K = 2$.

C. A Tractable Approximate Expression for the Probabilistic Objective Function

To obtain a tractable analytical expression, it is necessary to perform relaxation processing for the POF (29a). The relaxation starts with the inequality in the objective function (29a), and it is easy to observe the following inequality:

$$\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \geq \frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2/N} \tag{31}$$

The reason for the inequality is attributed to the implicit constraint $\|\boldsymbol{w}\| \leq 1/\sqrt{N}$ in problem (29). Utilizing the properties of probability, we can obtain the following inequality:

$$\Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{w}\|^2} \geq \gamma_0\right\} \geq \Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2/N} \geq \gamma_0\right\} \tag{32}$$

Now, the problem (29) can be further relaxed as:

$$\max_{\boldsymbol{f}, \boldsymbol{D}, \boldsymbol{w}} \Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2/N} \geq \gamma_0\right\} \tag{33a}$$

$$\text{s.t. } (29\text{b}) - (29\text{j}) \tag{33b}$$

To solve problem (33), we introduce the following problem:

$$\max_{\boldsymbol{f}, \boldsymbol{D}, \boldsymbol{w}} \Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{f}\|^2/N} \geq \gamma_0\right\} \tag{34a}$$

$$\text{s.t. } (29\text{b}) - (29\text{j}) \tag{34b}$$

In fact, both problems are equivalent, i.e., they share the same optimal solutions. The proof can be found in Appendix A of [38].

Notice that $\boldsymbol{h}_c \boldsymbol{h}_c^\dagger + \sigma_n^2/N$ is a positive definite matrix and can be decomposed as $\boldsymbol{L}\boldsymbol{L}^\dagger$ using the Cholesky decomposition. Then, the following equation holds true.

$$\Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\left|\boldsymbol{f}^\dagger \boldsymbol{h}_c\right|^2 + \sigma_n^2 \|\boldsymbol{f}\|^2/N} \geq \gamma_0\right\} = \Pr\left\{\frac{\left|\boldsymbol{f}^\dagger(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta})\right|^2}{\|\boldsymbol{f}^\dagger \boldsymbol{L}\|^2} \geq \gamma_0\right\} \tag{35}$$

$$\max_{\boldsymbol{f},\boldsymbol{D},\boldsymbol{w}} \Pr \left\{ \frac{\left| \boldsymbol{f}^{\dagger} \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta} \right) \right|^2}{\left| \boldsymbol{f}^{\dagger} \boldsymbol{h}_c \right|^2 + \sigma_n^2 \left\| \boldsymbol{w} \right\|^2} \geq \gamma_0 \right\} \tag{29a}$$

$$\text{s.t. } \text{Re}\left\{ \boldsymbol{f} \right\} \preceq \boldsymbol{1}_{NQ} \otimes \text{Re}\left\{ \boldsymbol{w} \right\} + 2 \left(\boldsymbol{1}_{NQM} - \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \right) \tag{29b}$$

$$\text{Re}\left\{ \boldsymbol{f} \right\} \succeq \boldsymbol{1}_{NQ} \otimes \text{Re}\left\{ \boldsymbol{w} \right\} - 2 \left(\boldsymbol{1}_{NQM} - \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \right) \tag{29c}$$

$$\text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \preceq \text{Re}\left\{ \boldsymbol{f} \right\} \preceq \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \tag{29d}$$

$$\text{Im}\left\{ \boldsymbol{f} \right\} \preceq \boldsymbol{1}_{NQ} \otimes \text{Im}\left\{ \boldsymbol{w} \right\} + 2 \left(\boldsymbol{1}_{NQM} - \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \right) \tag{29e}$$

$$\text{Im}\left\{ \boldsymbol{f} \right\} \succeq \boldsymbol{1}_{NQ} \otimes \text{Im}\left\{ \boldsymbol{w} \right\} - 2 \left(\boldsymbol{1}_{NQM} - \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \right) \tag{29f}$$

$$\text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \preceq \text{Im}\left\{ \boldsymbol{f} \right\} \preceq \text{vec}\left(\boldsymbol{D} \right) \otimes \boldsymbol{1}_M \tag{29g}$$

$$\boldsymbol{D}\boldsymbol{1}_Q = \boldsymbol{1}_N \tag{29h}$$

$$\boldsymbol{D} \in \{0,1\}^{N \times Q} \tag{29i}$$

$$\left\| \boldsymbol{f} \right\| \leq 1 \tag{29j}$$

For the right-hand side of equation (35), we observe that

$$\Pr \left\{ \left| \boldsymbol{f}^{\dagger} \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta} \right) \right|^2 \geq \gamma_0 \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\|^2 \right\} \tag{36a}$$

$$= \Pr \left\{ \left| \boldsymbol{f}^{\dagger} \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta} \right) \right| \geq \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \right\} \tag{36b}$$

$$\geq \Pr \left\{ \left| \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right| - \left| \boldsymbol{f}^{\dagger} \boldsymbol{\delta} \right| \geq \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \right\} \tag{36c}$$

$$= \Pr \left\{ \left| \boldsymbol{f}^{\dagger} \boldsymbol{\delta} \right| \leq \left| \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right| - \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \right\} \tag{36d}$$

$$= 1 - \exp \left(- \frac{\left(\left| \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right| - \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \right)^2}{\left\| \boldsymbol{G}_{\delta}^{1/2} \boldsymbol{f} \right\|^2} \right) \tag{36e}$$

where (36b) is relaxed to (36c) due to the properties of probability and the following inequality:

$$\left| \boldsymbol{f}^{\dagger} \left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta} \right) \right| \geq \left| \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right| - \left| \boldsymbol{f}^{\dagger} \boldsymbol{\delta} \right| \tag{37}$$

The inequality (37) is derived from the triangle inequality theorem. For (36d), since the random variable $\boldsymbol{\delta}$ follows a complex Gaussian distribution, i.e., $\boldsymbol{\delta} \sim \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_{\delta} \right)$, it has been shown in [39] that the real and imaginary parts of $\boldsymbol{f}^{\dagger} \boldsymbol{\delta}$ are real independent identically distributed (i.i.d.) Gaussian, and the modulus value of the random variable $\boldsymbol{f}^{\dagger} \boldsymbol{\delta}$ further follows a Rayleigh random distribution. Consequently, based on the Cumulative Distribution Function (CDF) of the Rayleigh distribution, we can derive the equation (36e).

In formulation (36), we relax the objective function of problem (34) for a tractable expression (36e). Then, by utilizing the properties of probability and the monotonicity of the exponential function, the problem (34) can be relaxed as:

$$\max_{\boldsymbol{f},\boldsymbol{D},\boldsymbol{w}} \frac{\left| \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right| - \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\|}{\left\| \boldsymbol{G}_{\delta}^{1/2} \boldsymbol{f} \right\|} \tag{38a}$$

$$\text{s.t. } (29b) - (29j) \tag{38b}$$

Compared to problem (29), the problem (38) effectively tackles the intractable POF. Before proceeding further, it should be noted that problem (38) is equivalent to the following problem:

$$\max_{\boldsymbol{f},\boldsymbol{D},\boldsymbol{w}} \frac{\text{Re}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} - \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\|}{\left\| \boldsymbol{G}_{\delta}^{1/2} \boldsymbol{f} \right\|} \tag{39a}$$

$$\text{s.t. } (29b) - (29j) \tag{39b}$$

$$\text{Im}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} = 0 \tag{39c}$$

This is because the objective value of problem (38) is unchanged when $\boldsymbol{f}$ undergoes an arbitrary phase rotation [40]. This implies that, without affecting the value of the objective function, we can rotate $\boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t$ to a real number by adding a phase to $\boldsymbol{f}$. Consequently, we assume that $\boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t$ is real, and add the constraint (39c), transforming problem (38) into problem (39).

However, it is observed that the objective function of problem (39) is non-convex, and the problem (39) cannot be solved optimally in polynomial time. To solve problem (39), we reformulate it as:

$$\max_{\boldsymbol{f},\boldsymbol{D},\boldsymbol{w}} \tau \tag{40a}$$

$$\text{s.t. } \tau \left\| \boldsymbol{G}_{\delta}^{1/2} \boldsymbol{f} \right\| + \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \leq \text{Re}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} \tag{40b}$$

$$(29b) - (29j) \tag{40c}$$

$$\text{Im}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} = 0 \tag{40d}$$

It can be found that, for a fixed τ, the following MIP problem (41) is solvable in polynomial time.

$$\text{find } \boldsymbol{f}, \boldsymbol{D}, \boldsymbol{w} \tag{41a}$$

$$\text{s.t. } \tau \left\| \boldsymbol{G}_{\delta}^{1/2} \boldsymbol{f} \right\| + \sqrt{\gamma_0} \left\| \boldsymbol{f}^{\dagger} \boldsymbol{L} \right\| \leq \text{Re}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} \tag{41b}$$

$$(29b) - (29j) \tag{41c}$$

$$\text{Im}\left\{ \boldsymbol{f}^{\dagger} \bar{\boldsymbol{h}}_t \right\} = 0 \tag{41d}$$

Evidently, if τ has a definite range of values, such as $\tau \in [\tau_{min}, \tau_{max}]$ (it will be discussed subsequently), then the solutions to problem (39) can be obtained by using the bisection method [41]. With this observation, assuming that the variable τ has a definite range of values, the procedure for

finding the optimal solutions of problem (39) is summarized in Algorithm 1. Once the optimal binary matrix $\boldsymbol{D}^{\star}$ is obtained, the optimal TC can be calculated by equation (15).

Last but not least, our proposed approach, while not capable of obtaining globally optimal solutions, is shown to provide superior solutions compared to the existing AIO [23] and BCD [27] techniques (This is corroborated by our simulation results). The main reason for this is that our proposed MIP method optimizes all variables simultaneously.

Algorithm 1 Bisection search for finding optimal solutions to problem (39).

Input: τ_{min}, τ_{max} and set a solution accuracy $\epsilon > 0$.
Output: The optimal solutions $\boldsymbol{f}^{\star}$, $\boldsymbol{D}^{\star}$ and $\boldsymbol{w}^{\star}$.
1: **repeat**
2: $\quad \tau = (\tau_{min} + \tau_{max})/2$;
3: $\quad$ If the problem (41) is feasible, update $\tau_{min} = \tau$, $\boldsymbol{f}^{\star} = \boldsymbol{f}$, $\boldsymbol{D}^{\star} = \boldsymbol{D}$ and $\boldsymbol{w}^{\star} = \boldsymbol{w}$; otherwise update $\tau_{max} = \tau$.
4: **until** $\tau_{max} - \tau_{min} < \epsilon$

Remark 3: Recalling equation (35), we can easily obtain the following inequality:

$$0 \leq \Pr\left\{ \left| \boldsymbol{f}^{\dagger}\left(\bar{\boldsymbol{h}}_t + \boldsymbol{\delta}\right) \right|^2 \geq \gamma_0 \left\| \boldsymbol{f}^{\dagger}\boldsymbol{L} \right\|^2 \right\} \leq 1 \quad (42)$$

Based on the above inequality and referring to formulation (36), it is straightforward to derive the following inequality:

$$0 \leq \exp\left(-\tau^2\right) \leq 1. \quad (43)$$

The monotonicity of the exponential function implies that $\tau \in [0, +\infty)$, indicating that there is no strict upper bound for τ. In practical applications, it is evident that setting τ to an excessively large number is unnecessary to achieve a probability value extremely close to 1. Thus, in this paper, we set $\tau_{max} = 10$, which yields a value of $\exp\left(-\tau^2\right)$ that is very close to 1.

Remark 4: The research content in this paper can also be addressed using the traversal method, specifically by fixing TC and solving RF. However, the traversal method may not be highly practical as it entails multiple rounds of two-step optimization solutions, and the computational cost is very high when the discrete sequence is long. The proposed method establishes a new joint solution framework that avoids multiple rounds of two-step solution. This provides a new perspective and solution approach for the joint optimization problem of discrete and continuous variables, with potential applicability to complex scenarios.

IV. NUMERICAL RESULTS

In this section, we present numerical experiments to evaluate the performance of the proposed method. For comparison purpose, the robust joint design method (constant modulus code model) in [23] is tested[4].

---

[4]It should be noted that the BCD technique utilized in [27] is essentially an alternating optimization method (similar to [23]), which does not enhance the output SINR. Instead, the method in [27] only changes the distribution structure of the output SINR. Therefore, we solely simulated the method in [23].

Unless otherwise specified, we consider a radar system in S-band with operating frequency $f_0 = 3$ GHz, and rectangular subpulse duration of 10 ns (corresponding to 1.5 m range resolution). The extended target is the typical fighter with 15 meters long, occupying $L = 10$ range cells. For the TC, $N = 7$ and $B = 1$ are chosen, so $Q = 2$ and $M = 16$. The covariance matrix $\boldsymbol{G}_t = \sigma_{\varepsilon}^2 \boldsymbol{I}$ of the error $\boldsymbol{\varepsilon}$ is set. The covariance matrix $\boldsymbol{G}_{\delta}$ of $\boldsymbol{\delta}$ is calculated by generating the 5000 TIRs (Monte Carlo trials) from $\mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_t\right)$. The TIR is defined as

$$t_0\left(i\right) = \sigma_t^2\left(\cos\left(0.6\pi i - 1.8\pi\right) + 1\right)e^{-|i-2|+j\pi(i-4)/6} \quad (44)$$

and the CIR information is defined as

$$\tilde{C}\left(k_1, k_2\right) = \sigma_c^2\left(1.2e^{-(k_1-k_2)^2} + 0.2e^{j2(k_1-k_2)}\right) \quad (45)$$

where σ_t^2 and σ_c^2 denote the power of TIR and CIR, respectively, with 10 dB and 30 dB. For problem (41), we set the desired SINR $\gamma_0 = 15$ dB to achieve a relatively high detection probability. Moreover, regarding to the stop criterion of the Algorithm 1, we set $\epsilon = 10^{-2}$ and $\tau \in [0, 10]$. In our simulations, the proposed problem (41) is solved by the high performance MIP solver GUROBI [35], and all the numerical simulations are performed based on MATLAB 2022b on a PC with CPU Intel(R) Xeon(R) Gold 6246 CPU @ 3.30 GHz 3.29 GHz (two processors) and 256 GB RAM.

A. The Output SINR under Different Nosie Power

In this subsection, we evaluate the performance of radar output SINR for different noise powers. Specifically, we consider setting the noise power σ_n^2 to -20 dB, -10 dB and 0 dB respectively, while setting the error power $\sigma_{\varepsilon}^2 = 0$ dB.

Fig. 2 shows the radar output SINR of the different methods for different noise powers. The results are obtained by conducting 5000 Monte Carlo trials, where TIRs are randomly generated from $\mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{G}_t\right)$. The optimal solutions for the TC and RF, obtained through the proposed method, are given in Tab. II and Tab. III, respectively. Here, we present the output SINR of the robust method [23] for continuous phase and 1-bit quantization, respectively. It can be observed that the proposed method achieves a significant improvement in output SINR compared to the robust method. More specifically, our method exhibits an approximately 10 dB enhancement in the output SINR when compared to the continuous phase case. The substantial improvement in output SINR can be attributed to two key factors. Firstly, our proposed method directly solves for the joint vector of the TC and RF, resulting in high quality solutions for both components. In contrast, the robust method employs an AIO technique that optimizes only a subset of variables in each iteration, which often leads to convergence to a local optimum without achieving high performance solutions. Secondly, our proposed method enforces the constraint that the TC must be selected from a given discrete alphabet. This ensures that the solved TC and RF are sub-optimal solutions for the given number of quantization bits. In contrast, the robust method, when using the constant modulus constraint, does not restrict the TC to be chosen from the given discrete alphabet. As a result, we
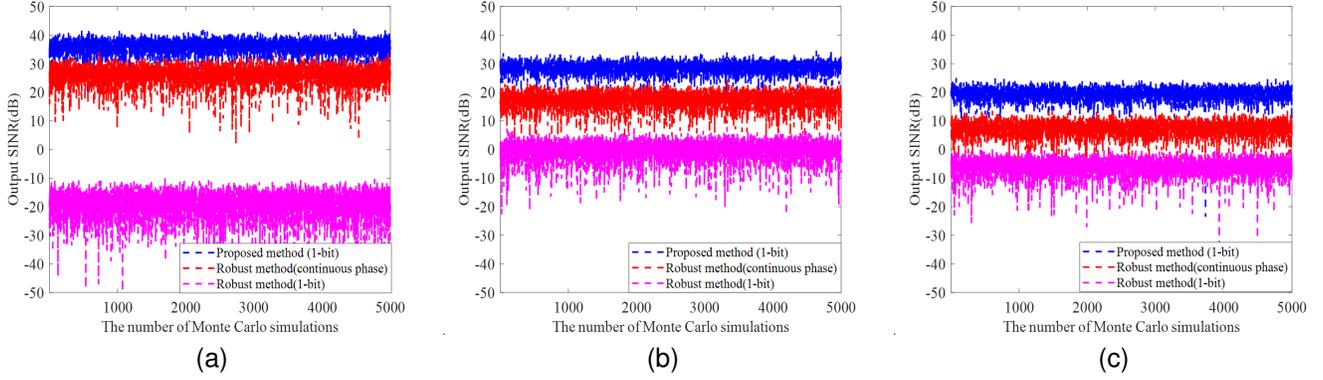
Fig. 2. The radar output SINR for different noise powers. (a) $\sigma_n^2 = -20$ dB. (b) $\sigma_n^2 = -10$ dB. (c) $\sigma_n^2 = 0$ dB.

observe a degradation in the output SINR when operating with 1-bit quantization.

Moreover, the distribution histogram of 5000 Monte Carlo SINRs for different noise powers are shown in Fig. 3. The mean $\mathbb{E}[\gamma]$ and the standard deviation $\mathbb{D}[\gamma]$ of the output SINR are given in Tab. IV. It can be seen that proposed method achieves a significantly smaller standard deviation while improving the output SINR, resulting in a much higher value of $\frac{\mathbb{E}[\gamma] - \gamma_0}{\mathbb{D}[\gamma]}$. These observations align with the concept of probabilistically robust TW design presented in [26], which enhances detection performance by reducing the standard deviation of the output SINR. This indicates that the structure of the output SINR distribution obtained from our proposed method corresponds to the PRD metric, providing evidence for probabilistic robustness of the proposed method.

TABLE II
PARAMETERS OF THE TC BY THE PROPOSED METHOD FOR DIFFERENT NOISE POWERS

σ_n^2 \ n	0	1	2	3	4	5	6
0 dB	$e^{j\pi}$	$e^{j\pi}$	$e^{j\pi}$	$e^{j\pi}$	$e^{j\pi}$	e^{j0}	e^{j0}
-10 dB	$e^{j\pi}$	$e^{j\pi}$	$e^{j\pi}$	e^{j0}	e^{j0}	$e^{j\pi}$	$e^{j\pi}$
-20 dB	$e^{j\pi}$	$e^{j\pi}$	e^{j0}	e^{j0}	$e^{j\pi}$	$e^{j\pi}$	e^{j0}

TABLE III
PARAMETERS OF THE RF BY THE PROPOSED METHOD FOR DIFFERENT NOISE POWERS

m \ σ_n^2	0 dB	-10 dB	-20 dB
0	$0.0751e^{-j0.3108}$	$0.0378e^{+j0.1420}$	$0.0219e^{+j1.5567}$
1	$0.0648e^{+j1.0600}$	$0.0941e^{+j1.8137}$	$0.1953e^{+j2.0859}$
2	$0.0992e^{+j2.3100}$	$0.1673e^{+j2.2907}$	$0.1963e^{+j2.3233}$
3	$0.0980e^{+j2.5474}$	$0.1424e^{+j2.4637}$	$0.0583e^{-j1.4812}$
4	$0.0899e^{+j2.5447}$	$0.0475e^{-j1.9633}$	$0.0861e^{-j0.8163}$
5	$0.1882e^{+j2.4599}$	$0.1395e^{-j0.7829}$	$0.0227e^{+j0.7835}$
6	$0.0524e^{-j3.0996}$	$0.0322e^{+j0.4519}$	$0.0707e^{+j2.3374}$
7	$0.1488e^{-j0.7437}$	$0.0660e^{+j2.3029}$	$0.0163e^{-j2.2766}$
8	$0.1179e^{-j0.5064}$	$0.0356e^{+j2.8855}$	$0.0651e^{-j0.6812}$
9	$0.0361e^{-j0.9823}$	$0.0404e^{-j0.9955}$	$0.0159e^{-j0.7400}$
10	$0.0223e^{-j1.1266}$	$0.0384e^{-j0.6455}$	$0.0067e^{-j2.6661}$
11	$0.0124e^{+j2.5803}$	$0.0034e^{-j0.9751}$	$0.0454e^{-j1.0364}$
12	$0.0255e^{+j0.3859}$	$0.0344e^{-j0.4880}$	$0.0273e^{+j2.1898}$
13	$0.0282e^{+j0.4567}$	$0.0358e^{-j0.9218}$	$0.0193e^{+j1.7762}$
14	$0.0451e^{-j0.3670}$	$0.0268e^{-j0.6150}$	$0.0589e^{-j0.7242}$
15	$0.0340e^{-j1.1037}$	$0.0071e^{+j2.2355}$	$0.0978e^{-j1.0060}$
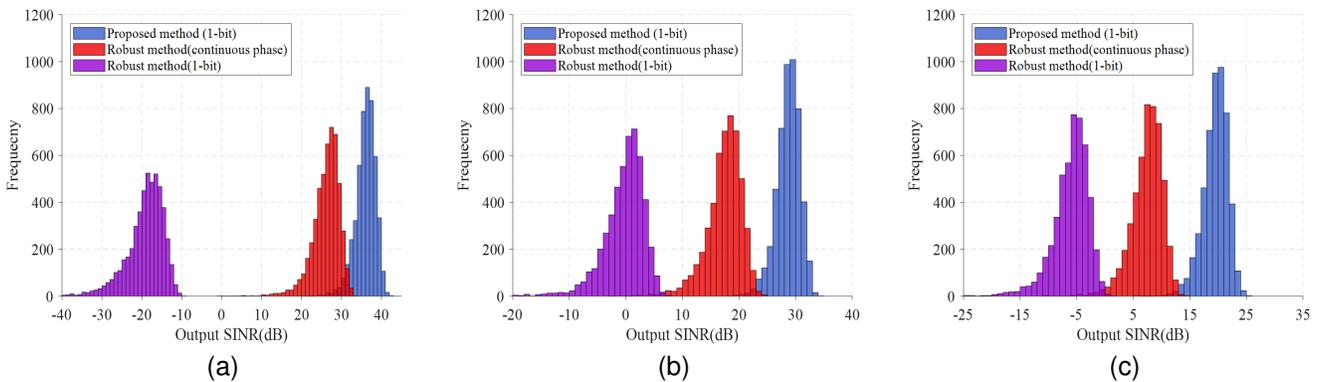
Fig. 3. The distribution of radar output SINR for different noise powers. (a) $\sigma_n^2 = -20$ dB. (b) $\sigma_n^2 = -10$ dB. (c) $\sigma_n^2 = 0$ dB.

TABLE IV
STATISTICAL CHARACTERISTICS COMPARISON OF DIFFERENT METHODS AT DIFFERENT NOISE POWERS.

σ_n^2	Robust method (continuous phase)			Robust method (1-bit)			Proposed method (1-bit)		
	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$
0 dB	7.50	2.71	-2.77	-6.02	3.16	-6.71	19.49	2.23	2.01
-10 dB	17.45	3.03	0.81	-0.29	3.56	-4.29	28.71	2.14	6.41
-20 dB	26.29	3.40	3.32	-19.21	4.92	-6.95	35.98	2.46	8.53
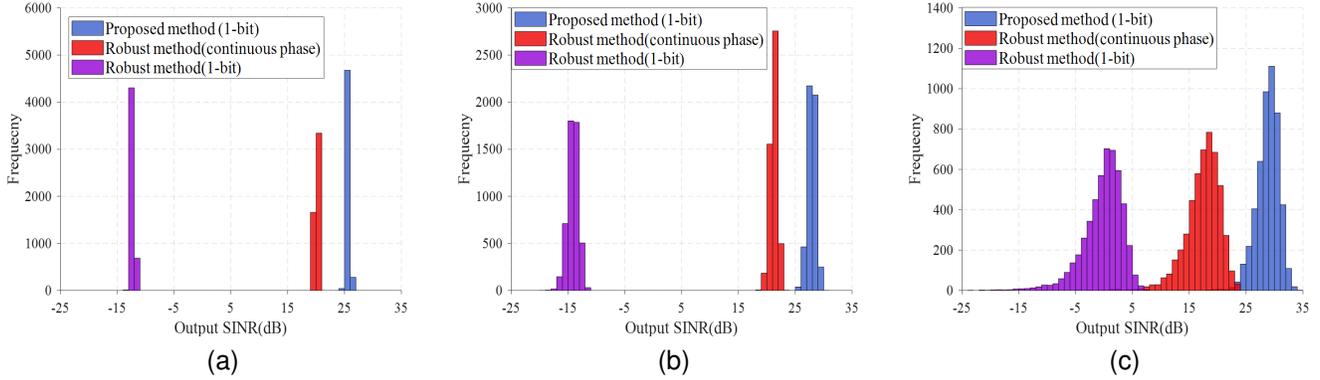
Fig. 4. The distribution of radar output SINR at different error power. (a) $\sigma_\varepsilon^2 = -20$ dB. (b) $\sigma_\varepsilon^2 = -10$ dB. (c) $\sigma_\varepsilon^2 = 0$ dB.

TABLE V
STATISTICAL CHARACTERISTICS COMPARISON OF DIFFERENT METHODS AT DIFFERENT ERROR POWER.

σ_ε^2	Robust method (continuous phase)			Robust method (1-bit)			Proposed method (1-bit)		
	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$	$\mathbb{E}\left[\gamma\right]/\text{dB}$	$\mathbb{D}\left[\gamma\right]/\text{dB}$	$\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$
0 dB	17.45	3.03	0.81	-0.29	3.56	-4.29	28.71	2.14	6.41
-10 dB	21.23	0.65	9.58	-14.11	0.96	-30.32	27.93	0.7	18.47
-20 dB	20.08	0.2	25.40	-12.27	0.24	-113.63	25.62	0.25	42.48

B. The Output SINR under Different Error Power

In this subsection, we compare the performance of radar output SINR for different error powers. We consider setting the error power σ_ε^2 to -20 dB, -10 dB and 0 dB respectively, while fixing the noise power $\sigma_n^2 = $ -10 dB.

Fig. 4 displays the output SINR distribution histograms for different error powers, generated from 5000 randomly TIRs. The mean $\mathbb{E}\left[\gamma\right]$ and the standard deviation $\mathbb{D}\left[\gamma\right]$ of the output SINR are given in Tab. V. It can be found from Fig. 4 that the shape of the histogram varies as the error power increases. When the error power is -20 dB or -10 dB, the standard deviation of our proposed method is almost the same as that of the robust method [23], while exhibiting significantly improved in terms of the output SINR. However, when the error power is 0 dB, the proposed method significantly reduces the standard deviation $\mathbb{D}\left[\gamma\right]$ of the output SINR while improving the output SINR, compared to the robust method. Additionally, our proposed method achieves a much higher value of $\frac{\mathbb{E}[\gamma]-\gamma_0}{\mathbb{D}[\gamma]}$. It is indicated that when the error power is relatively low, utilizing the PRD metric for the design of TW and RF does not alter the standard deviation of the output SINR. Conversely, as the error power increases, our proposed method reduces the standard deviation of the output SINR, demonstrating strong probabilistic robustness.

C. Probability of Output SINR Greater than Threshold

In this subsection, we firstly compare the value of $\Pr\left\{\gamma\left(\boldsymbol{w}, \boldsymbol{s}\right) \geq \gamma_0\right\}$ for different values of γ_0. Here, we set $\sigma_n^2 = $ -10 dB and $\sigma_\varepsilon^2 = 0$ dB. Fig. 5 illustrates the optimal probability values (optimal solution for problem (39)) as well as the probability values obtained from Monte Carlo simulations when different γ_0 are considered. It is evident that the probability obtained from Monte Carlo simulations, using the solved optimal TC and RF, outperforms the optimal solution of problem (39). This difference arises due to the relaxation processing in (32) and (36). Furthermore, when γ_0 exceeds a specific threshold, both probability values become zero. This occurs because, given the input signal, clutter and noise power, it is not possible to find the TC and RF in problem (39) that yields an output SINR greater than the specific threshold. Consequently, the probability value becomes zero.

Secondly, we evaluated the detection probability performance of the proposed method and compared it with the robust method. Using the evaluation criteria presented in [26], the stable detection probability is defined as

$$P_{sd} = \frac{N_r}{N_m} \qquad (46)$$
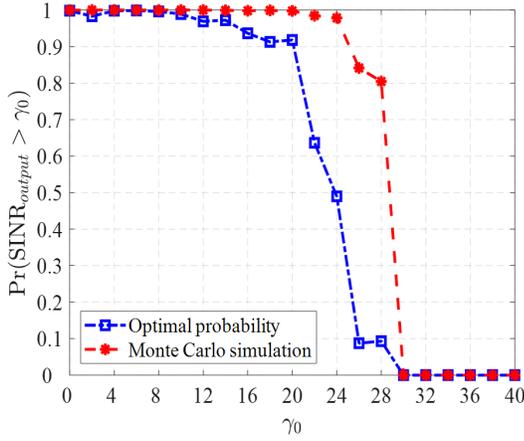
Fig. 5. Comparison of probability values.
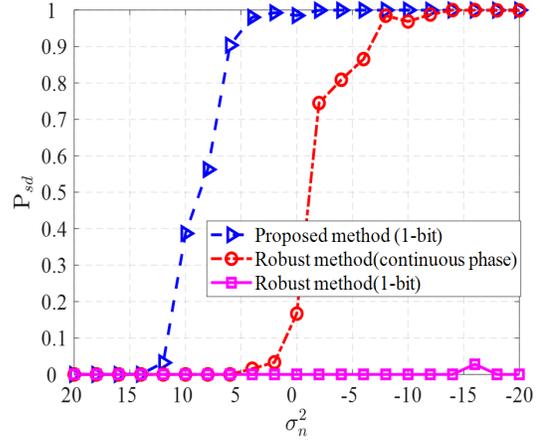
Fig. 6. Stable detection probability comparison.

TABLE VI
PERFORMANCE COMPARISON OF DIFFERENT QUANTIZATION BITS.

	Quantization bit	Robust method $\mathbb{E}\left[\gamma\right]$ /dB	Proposed method $\mathbb{E}\left[\gamma\right]$ /dB
$N=3, L=5,$ $M=7$	1-bit	-2.12	14.48
	2-bit	1.86	14.39
	3-bit	1.84	15.03
	continuous phase	6.93	/
$N=5, L=5,$ $M=9$	1-bit	-13.42	16.02
	2-bit	-2.15	17.04
	continuous phase	8.64	/

where N_m denotes the total number of Monte Carlo simulations, and N_r represents the count of occurrences of the event $\{\gamma\left(\boldsymbol{s}, \boldsymbol{w}\right) \geq \gamma_0\}$ during the Monte Carlo trials. Fig. 6 shows the stable detection probability P_{sd} for various values of σ_n^2, with $\sigma_\varepsilon^2 = 0$ dB and $\gamma_0 = 10$ dB. It can be seen that the proposed method achieves a higher P_{sd} value, particularly in scenarios with high noise power, demonstrating excellent performance.

D. Performance under Different Quantization Bit

In this subsection, we evaluate the performance of the proposed method under different quantization bits. Here, we set $\sigma_n^2 = 0$ dB, and $\gamma_0 = 5$ dB. Considering the high operational complexity of the proposed algorithm, we select relatively small values for N and L. Tab. VI presents the performance of the proposed algorithm for different numbers of bits. It is observed that when $N = 3$, increasing the number of bits from 1 to 2 does not significantly improve the output SINR due to the relatively short TC. This phenomenon stems from two factors: the short length of the TC and the high quality of the solutions obtained through the proposed method. At $N = 3$, the proposed method already generates high-quality solutions when the number of bits is 1. Due to the short length of N, transitioning from 1 to 2 bits might not significantly enhance the output SINR of the proposed method. In this case, substantial improvements in radar output SINR can only be achieved through further increases in the number of bits (e.g., 3bits) or in the length of N (e.g., $N = 5$).

In contrast, since robust methods do not yield high-quality solutions, a significant increase in the radar output SINR is observed with an increase in the number of bits. However,

once the upper performance limit of the robust method is reached, further increasing the number of bits does not lead to an improvement in the output SINR. For instance, when $N = 3$, increasing the number of bits from 1 to 2 results in a significant improvement in the output SINR, whereas there is no significant change from 2 to 3.

E. Computational complexity

In this subsection, we attempt to analyze the complexity of solving problem (41) using the following way. The integer variable $\boldsymbol{D}$, constrained by (29h) and (29i), encompasses $\mathcal{O}(Q^N)$ states. For each state, the computational complexity of solving for variable $\boldsymbol{w}$ is $\mathcal{O}(M^{3.5})$. Employing the exhaustive method, the computational cost of solving problem (41) at each step of the bisection search does not exceed $\mathcal{O}(Q^N M^{3.5})$.

In addition, we evaluate the operational time of the proposed algorithm under different quantization bits. Since the solution accuracy ϵ influences the number of iterations in Algorithm 1, we focus on the computational time to solve the problem (41) once. In other words, the Algorithm 1 executes only one loop. Herein, we set $N = 3, L = 5$. The operation time of the proposed algorithm for different quantization bits is presented in Tab. VII. It can be observed that as the number of quantization bits increases, the operation time of the proposed method also increases. This is because an increase in the number of quantization bits leads to a higher-dimensional solution vector $\boldsymbol{f}$, resulting in a longer operation time. This is also a drawback of our proposed method.

V. CONCLUSION

In this paper, we have addressed the problem of probabilistically robust joint design of TC and RF for extended

TABLE VII
OPERATION TIME OF THE PROPOSED METHOD FOR DIFFERENT
QUANTIZATION BITS.

$N = 3, L = 5,$	1-bit	2-bit	3-bit	4-bit
$M = 7$	11.25s	60.01s	796.35s	14567s

targets detection. Considering a limited number of bits in the digital waveform generators, the TW was restricted to discrete phase codes. Subsequently, we proposed an MIP approach to optimize TC and RF simultaneously, resulting in high quality solutions. Simulation experiments demonstrated the effectiveness and superiority of the proposed algorithm. It is worth noting that that the proposed method can be applied to many joint optimization models with continuous and discrete variables. As a future work, we shall consider how to obtain high-performance TC and RF when the TIR probability distribution model is uncertain, and study how to reduce the high computational complexities arising from the solving the optimization problem.

REFERENCES

[1] B. Friedlander, "Waveform design for MIMO radars," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 1227–1238, 2007.

[2] L. Xu, J. Li, and P. Stoica, "Target detection and parameter estimation for MIMO radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 927–939, 2008.

[3] J. Xu, J. Yu, Y.-N. Peng, and X.-G. Xia, "Radon-fourier transform for radar target detection, I: Generalized doppler filter bank," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1186–1202, 2011.

[4] H. Xu, J. Zhang, W. Liu, S. Wang, and C. Li, "High-resolution radar waveform design based on target information maximization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 5, pp. 3577–3587, 2020.

[5] Z. Tang, Q. Bao, J. Pan, H. Dai, and W. Jiang, "Noncooperative bistatic radar countermeasures based on the joint design of radar waveforms and mismatched filters," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[6] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. De Maio, and A. Aubry, "A doppler robust design of transmit sequence and receive filter in the presence of signal-dependent interference," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 772–785, 2014.

[7] Z. Li, J. Shi, W. Liu, J. Pan, and B. Li, "Robust joint design of transmit waveform and receive filter for MIMO-STAP radar under target and clutter uncertainties," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 1156–1171, 2022.

[8] C. G. Tsinos, A. Arora, S. Chatzinotas, and B. Ottersten, "Joint transmit waveform and receive filter design for dual-function radar-communication systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1378–1392, 2021.

[9] C. Huang, Z. Huang, Q. Zhou, Z. Li, Z. Yang, and J. Zhang, "Joint transmit waveform and receive filter design for the dfrc system with imperfect csi," *IEEE Transactions on Radar Systems*, vol. 2, pp. 288–302, 2024.

[10] A. Aubry, A. De Maio, and M. M. Naghsh, "Optimizing radar waveform and doppler filter bank via generalized fractional programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1387–1399, 2015.

[11] W. Zhu and J. Tang, "Robust design of transmit waveform and receive filter for colocated MIMO radar," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 2112–2116, 2015.

[12] A. Aubry, A. De Maio, Y. Huang, and M. Piezzo, "Robust design of radar doppler filters," *IEEE Transactions on Signal Processing*, vol. 64, no. 22, pp. 5848–5860, 2016.

[13] 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, 2018.

[14] F. Wang, X.-G. Xia, C. Pang, X. Cheng, Y. Li, and X. Wang, "Joint design methods of unimodular sequences and receiving filters with good correlation properties and doppler tolerance," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.

[15] L. Wu, X. Cheng, H. Huang, D. Ciuonzo, B. Shankar, and B. Ottersten, "Constant-modulus waveform design with polarization-adaptive power allocation in polarimetric radar," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2146–2161, 2023.

[16] Y. Yao, Z. Li, H. Liu, P. Miao, and L. Wu, "Robust transceiver optimization against echo eclipsing via majorization-minimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 3, pp. 2464–2479, 2023.

[17] Z.-J. Wu, C.-X. Wang, Y.-C. Li, and Z.-Q. Zhou, "Extended target estimation and recognition based on multimodel approach and waveform diversity for cognitive radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[18] A. Leshem, O. Naparstek, and A. Nehorai, "Information theoretic adaptive radar waveform design for multiple extended targets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 42–55, 2007.

[19] M. Bell, "Information theory and radar waveform design," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1578–1597, 1993.

[20] C.-Y. Chen and P. 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.

[21] P. Chen, C. Qi, L. Wu, and X. Wang, "Estimation of extended targets based on compressed sensing in cognitive radar system," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 941–951, 2017.

[22] B. Tang and J. Tang, "Robust waveform design of wideband cognitive radar for extended target detection," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3096–3100.

[23] S. M. Karbasi, A. Aubry, A. De Maio, and M. H. Bastani, "Robust transmit code and receive filter design for extended targets in clutter," *IEEE Transactions on Signal Processing*, vol. 63, no. 8, pp. 1965–1976, 2015.

[24] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 330–343, 2007.

[25] B. Tang, J. Liu, H. Wang, and Y. Hu, "Constrained radar waveform design for range profiling," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1924–1937, 2021.

[26] Z. Xu, Z. Xie, C. Fan, and X. Huang, "Probabilistically robust radar waveform design for extended target detection," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4212–4224, 2022.

[27] Z. Xu, B. Tang, Z. Xie, and X. Huang, "Probabilistically robust design of radar waveform-filter for extended target detection in the presence of clutter," *IEEE Transactions on Signal Processing*, pp. 1–15, 2023.

[28] M. N. Cohen, "Variability of ultrahigh-range-resolution radar profiles and some implications for target recognition," *Proceedings of Spie the International Society for Optical Engineering*, 1992.

[29] A. De Maio, S. De Nicola, Y. Huang, Z.-Q. Luo, and S. Zhang, "Design of phase codes for radar performance optimization with a similarity constraint," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 610–621, 2009.

[30] J. Yang, A. Aubry, A. De Maio, X. Yu, and G. Cui, "Design of constant modulus discrete phase radar waveforms subject to multi-spectral constraints," *IEEE Signal Processing Letters*, vol. 27, pp. 875–879, 2020.

[31] J. Zhang and N. Xu, "Discrete phase coded sequence set design for waveform-agile radar based on alternating direction method of multipliers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4238–4252, 2020.

[32] Q. Li, E. Rothwell, K.-M. Chen, and D. Nyquist, "Scattering center analysis of radar targets using fitting scheme and genetic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 2, pp. 198–207, 1996.

[33] S. Haykin, "Cognitive radar: a way of the future," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 30–40, 2006.

[34] A. Aubry, A. DeMaio, A. Farina, and M. Wicks, "Knowledge-aided (potentially cognitive) transmit signal and receive filter design in signal-dependent clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 1, pp. 93–117, 2013.

[35] Gurobi, *Gurobi Optimizer Reference Manual*. Gurobi Optimization Incorporated, 2020.

[36] CPLEX, *CPLEX User's Manual*. IBMILOG CPLEX Optimization Studio, 2007.

[37] L. A. Wolsey, *Integer programming*. USA: John Wiley & Sons, 2020.

[38] A. Aubry, A. De Maio, and M. M. Naghsh, "Optimizing radar waveform and doppler filter bank via generalized fractional programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1387–1399, 2015.

[39] S. A. Vorobyov, H. Chen, and A. B. Gershman, "On the relationship between robust minimum variance beamformers with probabilistic and worst-case distortionless response constraints," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5719–5724, 2008.

[40] Y. Zhuang, X. Zhang, and Z. He, "Beampattern synthesis for phased array with dual-phase-shifter structure," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 10, pp. 7053–7058, 2021.

[41] X. Zhang, Z. He, X. Zhang, and W. Peng, "High-performance beampattern synthesis via linear fractional semidefinite relaxation and quasi-convex optimization," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 7, pp. 3421–3431, 2018.