# Mutual Information Maximization for Waveform Design Using Minorization-Maximization

Zhexian Yang*, Linlong Wu†, and Ziping Zhao*
*School of Information Science and Technology, ShanghaiTech University, Shanghai, China
†Interdisciplinary Centre for Security, Reliability, and Trust (SnT), University of Luxembourg, Luxembourg
Email: *{zhexianyang, zipingzhao}@shanghaitech.edu.cn, †linlong.wu@uni.lu

***Abstract*—In this paper, we study the problem of MIMO waveform design for target estimation by maximizing the conditional mutual information between an extended target and the received signals. Existing methods solved this problem through a two-stage process, which is suboptimal and limited in handling practical waveform constraints. To address these issues, we formulate the problem as a nonconvex one and solve it directly through a one-stage numerical approach. Our proposed algorithm leverages the minorization-maximization (MM) method, enabling flexible and efficient handling of diverse waveform constraints. Notably, the MM algorithm is single-loop, where each iteration offers a closed-form solution. Furthermore, we reveal that the derived MM algorithm can be interpreted from the gradient projection perspective. Numerical experiments demonstrate that the proposed algorithm outperforms existing methods under various waveform constraints.**

***Index Terms*—Mutual information, multiple-input multiple-output (MIMO) radar, waveform design, target estimation, minorization-maximization.**

## I. Introduction

Target estimation is of immense value among various radar applications, such as air-traffic control [1] and air defense [2]. The target impulse response (TIR) provides a convenient representation to characterizing the target profile [3], [4]. Consequently, the key to effective target estimation is extracting TIR information from radar echoes. To enhance extraction performance, our goal is to maximize the TIR information in the echoes by leveraging the adaptability of waveform design. Therefore, waveform design plays a critical role in target estimation.

Using information theory in waveform design has a rich history with extensive research exploring this area. In the early 1950s, Woodward and Davis [5] were among the pioneers who applied information theory to radar receiver design. Following this, Bell introduced the idea of maximizing conditional mutual information (MI) between the TIR and the echo to improve target parameter estimation capabilities [3]. This approach was further developed for multiple-input multiple-output (MIMO) radar under power constraints, introducing the well-known water-filling solution [6]. The problem was formulated as a convex optimization in [6], involving a Toeplitz and Kronecker product constraint on the waveform matrix. A two-stage process was utilized: first, a waveform matrix was derived using the water-filling solution to meet power constraints, followed by an alternating optimization algorithm to enforce the Kronecker product structure [7]. However, while effective, this approach was deemed suboptimal due to limitations in efficiently considering the Toeplitz product structure and practical waveform constraints. Subsequently, in [8], the authors expanded their research to a robust design by addressing the uncertainty of target power spectral density. This information theory-based design approach has since been extensively explored in various studies. For instance, the colored noise was considered in [9]. A two-stage approach for distributed MIMO radar was proposed in [10], using MI for waveform design and echo selection. Furthermore, joint radar and communication systems were successfully designed using MI-based radar waveforms [11]. More recently, a minorization-maximization (MM) method was suggested in [12] for single-input and single-input single-output (SISO) radar waveform design, involving a double-loop process for subproblem resolution. Apart from the research above focusing on time domain waveform design, there is another line of work that explores waveform design through MI maximization in the frequency domain [13]–[15]. However, this aspect is beyond the scope of the current discussion.

In this paper, we revisit the problem of maximizing mutual information for MIMO waveform design. We formulate a nonconvex problem and introduce an efficient algorithm based on the MM framework [16], which efficiently exploits the problem structure. Our method differs from the approaches in [6] and [17] by being a single-loop algorithm that directly designs waveforms, with each iteration providing a closed-form solution. Moreover, our algorithm is adaptable to accommodate different practical waveform constraints. Additionally, we demonstrate that the proposed MM algorithm can be interpreted from a gradient projection perspective. Numerical experiments demonstrate that the proposed algorithm outperforms existing methods under various waveform constraints.

## II. System Model

We consider a MIMO radar system with $M$ transmit antennas and $N$ receive antennas. The response of the extended target from the $m$-th transmit antenna to the $n$-th receive

The experiments of this work were supported by the core facility Platform of Computer Science and Communication, SIST, ShanghaiTech University.

$$\mathbf{H}_{n,m} = \begin{bmatrix} h_{n,m}(t) & h_{n,m}(t-1) & \cdots & h_{n,m}(t-L_t+1) \\ h_{n,m}(t+1) & h_{n,m}(t) & \cdots & h_{n,m}(t-L_t+2) \\ \vdots & \vdots & \ddots & \vdots \\ h_{n,m}(t+L_r-1) & h_{n,m}(t+L_r-2) & \cdots & h_{n,m}(t+L_r-L_t) \end{bmatrix} \in \mathbb{C}^{L_r \times L_t} \tag{1}$$

antenna is modeled as a finite impulse response linear system in baseband with length $L$, i.e., $h_{n,m}(t),\ t = 0, \ldots, L-1$.

Define $x_m(t)$ as the transmit waveform from the $m$-th transmit element at time $t$ and $e_n(t)$ to be the additive complex Gaussian noise with mean zero and variance $\sigma^2$ received at the $n$-th receive element at time $t$. Then, the received signal of the $n$-th receive element at time $t$ is given by

$$y_n(t) = \sum_{m=1}^{M} \sum_{\tau=0}^{L-1} h_{n,m}(t-\tau) x_m(\tau) + e_n(t), \tag{2}$$

for $n = 1, \ldots, N$. Let $L_t$ be the duration of the transmitted signals and $L_r$ be the received signal length (we assume $L_t \gg L$). Denote $\mathbf{x}_m = [x_m(0), \ldots, x_m(L_t-1)]^\mathsf{T} \in \mathbb{C}^{L_t}$, $\mathbf{y}_n = [y_n(t), \ldots, y_n(t+L_r-1)]^\mathsf{T} \in \mathbb{C}^{L_r}$, and $\mathbf{e}_n = [e_n(t), \ldots, e_n(t+L_r-1)]^\mathsf{T} \in \mathbb{C}^{L_r}$. Based on (2), we have

$$\mathbf{y}_n = \sum_{m=1}^{M} \mathbf{H}_{n,m} \mathbf{x}_m + \mathbf{e}_n, \tag{3}$$

for $n = 1, \ldots, N$, where $\mathbf{H}_{n,m}$ is defined in (1) with $h_{n,m}(t) = 0$ for $t < 0$ or $t > M-1$. We further define $\mathbf{x} = [\mathbf{x}_1^\mathsf{T}, \cdots, \mathbf{x}_M^\mathsf{T}]^\mathsf{T}$, $\mathbf{y} = [\mathbf{y}_1^\mathsf{T}, \ldots, \mathbf{y}_N^\mathsf{T}]^\mathsf{T}$, $\mathbf{e} = [\mathbf{e}_1^\mathsf{T}, \ldots, \mathbf{e}_N^\mathsf{T}]^\mathsf{T}$, and $\mathbf{H} = [\mathbf{H}_1^\mathsf{T}, \ldots, \mathbf{H}_N^\mathsf{T}]^\mathsf{T}$ with $\mathbf{H}_n = [\mathbf{H}_{n,1}, \ldots, \mathbf{H}_{n,M}]$, we finally obtain the following compact model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{e}.$$

where $\mathbf{H}$ can be calculated by $\mathbf{h} = [h_{1,1}(0), \ldots, h_{N,M}(L-1)]^\top$, which is assumed be a Gaussian random vector with zero mean and covariance $\boldsymbol{\Sigma}_\mathbf{h}$. Next, we present mutual information for waveform design.

## III. Problem Formulation

For given waveform $\mathbf{x}$, the mutual information between receive signal $\mathbf{y}$ and impulse response $\mathbf{H}$ is given by [18]

$$I(\mathbf{y}; \mathbf{H}) = H(\mathbf{y}) - H(\mathbf{y} \mid \mathbf{H}) = H(\mathbf{y}) - H(\mathbf{e}),$$

where $H(\mathbf{y}) = -\int \log p(\mathbf{y}) p(\mathbf{y})\, \mathrm{d}\mathbf{y}$ is the differential entropy of $\mathbf{y}$, $p(\mathbf{y})$ is the probability density function of $\mathbf{y}$, and $H(\mathbf{e})$ is the differential entropy of $\mathbf{e}$. The MI $I(\mathbf{y}; \mathbf{H})$ measures the information about $\mathbf{H}$ provided by the received signal $\mathbf{y}$ [19].

Given the knowledge of $\mathbf{x}$, the received signal $\mathbf{y}$ obeys the Gaussian random process with mean zero and covariance matrix $(\mathbf{I} \otimes \mathbf{x})^\mathsf{H} \boldsymbol{\Sigma} (\mathbf{I} \otimes \mathbf{x}) + \sigma^2 \mathbf{I}$, where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{1,1} & \boldsymbol{\Sigma}_{1,2} & \ldots & \boldsymbol{\Sigma}_{1,L_rN} \\ \boldsymbol{\Sigma}_{2,1} & \boldsymbol{\Sigma}_{2,2} & \ldots & \boldsymbol{\Sigma}_{2,L_rN} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{L_rN,1} & \boldsymbol{\Sigma}_{L_rN,2} & \ldots & \boldsymbol{\Sigma}_{L_rN,L_rN} \end{bmatrix}, \tag{4}$$

with $\boldsymbol{\Sigma}_{i,j} \in \mathbb{C}^{L_tM \times L_tM}$ being the covariance matrix of the $i$-th and the $j$-th rows of $\mathbf{H}$. Then, the MI is computed as follows:

$$I(\mathbf{y}; \mathbf{H}) = \log\det\left(\mathbf{I} + \frac{1}{\sigma^2} (\mathbf{I} \otimes \mathbf{x})^\mathsf{H} \boldsymbol{\Sigma} (\mathbf{I} \otimes \mathbf{x})\right). \tag{5}$$

Hence, the MI maximization problem is

$$\underset{\mathbf{x} \in \mathcal{C}}{\text{maximize}} \quad \log\det\left(\mathbf{I} + \frac{1}{\sigma^2} (\mathbf{I} \otimes \mathbf{x})^\mathsf{H} \boldsymbol{\Sigma} (\mathbf{I} \otimes \mathbf{x})\right), \tag{6}$$

where $\mathcal{C}$ generally denotes the waveform constraints. In practice, we can use $\mathcal{C}$ to constrain the power of the transmitting waveforms. In this case, we define the power constraint:

$$\mathcal{C} = \{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq P_0\}.$$

The power constraints can be specifically considered for each antenna, in which case we define the per-antenna power constraint [20], [21]

$$\mathcal{C} = \left\{\mathbf{x} \mid \|\mathbf{x}_m\|_2^2 \leq P_0/M, m = 1, \ldots, M\right\}.$$

Besides, for each antenna, we consider the peak-to-average ratio (PAR) constraint [22], [23] defined as

$$\mathcal{C} = \left\{\mathbf{x} \mid \frac{\max_t |x_m(t)|^2}{\|\mathbf{x}_m\|_2^2 / L_t} \leq \rho, \|\mathbf{x}_m\|_2^2 \leq P_0/M\right\},$$

where $\rho \in [1, L_t]$.

Furthermore, in some cases, we may require the designed waveforms to be similar to one that has attained good properties, which leads to the similarity constraint [24], [25]:

$$\mathcal{C} = \left\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}^{\text{ref}}\|_\infty \leq \epsilon\right\}.$$

where $\mathbf{x}^{\text{ref}}$ is a reference waveform and $\epsilon$ denotes the similarity level. A more restrictive waveform constraint is the constant modulus constraint [26]:

$$\mathcal{C} = \{\mathbf{x} \mid |x_m(t)| = c\},$$

which requires the waveforms to have the same amplitudes. In practice, the phase shifter only operators on finite phases [27], [28], and hence the constant modulus with phase alphabet constraint is commonly considered

$$\mathcal{C} = \left\{\mathbf{x} \mid x_m(t) = c e^{j \frac{2\pi k}{K}}, k = 0, \ldots, K-1\right\},$$

where $K$ is the total number of the alphabet.

## IV. Proposed Algorithm

### A. Minorization-Maximization

We briefly introduce the minorization-maximization (MM) algorithmic framework [16], [29]. For a general optimization problem, minimization of $f(\mathbf{x})$ subject to the constraint set

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{A}\left(\mathbf{I}+\mathbf{X}^{\mathsf{H}}\mathbf{B}\mathbf{X}\right)^{-1}\right) \le{}& \operatorname{tr}\left(\mathbf{X}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\mathbf{A}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\mathbf{X}\right) \\
&-2\Re\left(\operatorname{tr}\left(\mathbf{A}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\mathbf{X}\right)\right)+\operatorname{tr}(\mathbf{A})+\operatorname{tr}\left(\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\mathbf{A}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)
\end{aligned} \tag{8}
$$

$\mathcal{C}$, an MM algorithm iteratively solves a sequence of subproblems with a surrogate function over $\mathcal{C}$. The surrogate function, denoted as $\underline{f}(\mathbf{x}, \mathbf{x}^{(k)})$, of $f(\mathbf{x})$ at the iterate $\mathbf{x}^{(k)}$, for $k = 0, 1, \ldots$, satisfies

$$
\begin{aligned}
\underline{f}(\mathbf{x}^{(k)}, \mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}), \qquad & \forall \mathbf{x}^{(k)} \in \mathcal{C}, \\
\underline{f}(\mathbf{x}, \mathbf{x}^{(k)}) \le f(\mathbf{x}), \qquad & \forall \mathbf{x}, \mathbf{x}^{(k)} \in \mathcal{C}.
\end{aligned}
$$

Thus, at the $k$-th iteration, the update rule is

$$
\mathbf{x}^{(k+1)} \in \arg\max_{\mathbf{x}\in\mathcal{C}} \underline{f}(\mathbf{x}, \mathbf{x}^{(k)}).
$$

It can be verified that within the MM scheme, the objective function value is monotonically non-increasing over iterations, i.e., $f(\mathbf{x}^{(k+1)}) \ge \underline{f}(\mathbf{x}^{(k+1)}, \mathbf{x}^{(k)}) \ge \underline{f}(\mathbf{x}^{(k)}, \mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)})$.

### *B. Algorithm Derivation*

To deploy the MM framework, the key point is to find the lower bound of the objective function in problem (6) such that the subproblems are easy to solve.

**Lemma 1.** *For a positive definite variable* $\mathbf{X}$*, at iterate* $\underline{\mathbf{X}}$*, it follows that*[1]

$$
\log\det(\mathbf{X}) \ge \log\det(\underline{\mathbf{X}}) + \operatorname{tr}\left(\mathbf{X}^{-1}(\mathbf{X}-\underline{\mathbf{X}})\right),
$$

*where the equality is attained at* $\mathbf{X} = \underline{\mathbf{X}}$.

For notational simplicity, we define $\mathbf{X} = \mathbf{I}_{L_r N} \otimes \mathbf{x} \in \mathbb{C}^{L_r N L_t M \times L_r N}$. Applying Lemma 1 to the objective of problem (6) yields

$$
\begin{aligned}
\log\det\left(\mathbf{I}+\frac{1}{\sigma^2}\mathbf{X}^{\mathsf{H}}\mathbf{\Sigma}\mathbf{X}\right) \ge{}& \log\det\left(\mathbf{I}+\frac{1}{\sigma^2}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{\Sigma}\underline{\mathbf{X}}\right) \\
&+L_r N - \operatorname{tr}\left(\left(\mathbf{I}+\frac{1}{\sigma^2}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{\Sigma}\underline{\mathbf{X}}\right)\left(\mathbf{I}+\frac{1}{\sigma^2}\mathbf{X}^{\mathsf{H}}\mathbf{\Sigma}\mathbf{X}\right)^{-1}\right).
\end{aligned} \tag{7}
$$

We then apply a further minorization step based on the following lemma.

**Lemma 2.** *Given a positive definite* $\mathbf{A}$ *and a positive semi-definite* $\mathbf{B}$*, we have* (8)*, where the equality is attained at* $\mathbf{X} = \underline{\mathbf{X}}$.

Denote $\mathbf{A} = \mathbf{I} + \frac{1}{\sigma^2}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{\Sigma}\underline{\mathbf{X}}$ and $\mathbf{B} = \frac{1}{\sigma^2}\mathbf{\Sigma}$. Based on Lemma 2, a further minorization step is applied to the third term in the RHS of the inequality (7), leading to the problem:

$$
\underset{\mathbf{x}\in\mathcal{C}}{\text{maximize}} \quad 2\Re(\mathbf{w}^{\mathsf{H}}\mathbf{x}) - \mathbf{x}^{\mathsf{H}}\mathbf{M}\mathbf{x}, \tag{9}
$$

where $\mathbf{w} \in \mathbb{C}^{N_t M}$ and $\mathbf{M} \in \mathbb{C}^{N_t M \times N_t M}$ are defined in the following way. Define $\mathbf{V} = \mathbf{B}\underline{\mathbf{X}}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\mathbf{A}$, and $\mathbf{U} = \mathbf{B}\underline{\mathbf{X}}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\mathbf{A}\left(\mathbf{I}+\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}\underline{\mathbf{X}}\right)^{-1}\underline{\mathbf{X}}^{\mathsf{H}}\mathbf{B}$. We have $\mathbf{w} = \mathbf{v}_1 + \ldots + \mathbf{v}_{L_r N}$ where $\mathbf{v}_i \in \mathbb{C}^{L_t M}$, for $i = 1, \ldots, L_r N$, is defined as a block vector taking its rows from $(i-1)L_t M + 1$ to $iL_t M$ in the $i$th column of $\mathbf{V}$, and

$$
\mathbf{M} = \mathbf{U}_{1,1} + \ldots + \mathbf{U}_{L_r N, L_r N}
$$

where $\mathbf{U}_{i,i} \in \mathbb{C}^{L_t M \times L_t M}$, for $i = 1, \ldots, L_r N$, is a block matrix of $\mathbf{U}$ defined by its rows from $(i-1)L_t M + 1$ to $iL_t M$ and columns from $(i-1)L_t M + 1$ to $iL_t M$. We apply a further minorization based on the following lemma to evoke an easy solution.

**Lemma 3.** *For any Hermitian matrix* $\mathbf{M}$ *and* $\lambda \ge \lambda_{\max}(\mathbf{M})$*, at iterate* $\underline{\mathbf{x}}$*, we have*

$$
\mathbf{x}^{\mathsf{H}}\mathbf{M}\mathbf{x} \le \lambda\mathbf{x}^{\mathsf{H}}\mathbf{x} + 2\Re\left(\mathbf{x}^{\mathsf{H}}(\mathbf{M}-\lambda\mathbf{I})\underline{\mathbf{x}}\right) + \underline{\mathbf{x}}^{\mathsf{H}}(\lambda\mathbf{I}-\mathbf{M})\underline{\mathbf{x}},
$$

*where the equality is attained at* $\mathbf{x} = \underline{\mathbf{x}}$.

Based on Lemma 3, a minorized problem is given by

$$
\underset{\mathbf{x}\in\mathcal{C}}{\text{maximize}} \quad 2\Re(\mathbf{z}^{\mathsf{H}}\mathbf{x}) - \lambda\mathbf{x}^{\mathsf{H}}\mathbf{x}, \tag{10}
$$

where $\mathbf{z} = \mathbf{w} - (\mathbf{M}-\lambda\mathbf{I})\underline{\mathbf{x}}$. Solutions to problem (10) for different waveform constraints are summarized in the following lemma.

**Lemma 4.** *When* $\mathcal{C}$ *refers to the power constraint,*

$$
\mathbf{x}^\star = \min\left\{\lambda^{-1}, \sqrt{P_0}/\|\mathbf{z}\|_2\right\}\mathbf{z}.
$$

*When* $\mathcal{C}$ *refers to PAR constraint, the problem* (10) *has a closed-form solution by analyzing the KKT conditions [30]. When* $\mathcal{C}$ *refers to similarity constraint,*

$$
x_m(t)^\star = x_m^{\text{ref}}(t) + \min\left\{1, \epsilon/\left|z_m^{\text{ref}}(t)\right|\right\} z_m^{\text{ref}}(t),
$$

*where* $z_m^{\text{ref}}(t) = \lambda^{-1}z(t) - x_m^{\text{ref}}(t)$*, and* $t = 0, \ldots, L_t - 1$*. When* $\mathcal{C}$ *refers to constant modulus constraint,*

$$
\mathbf{x}^\star = c \cdot e^{j\arg(\mathbf{z})}.
$$

*When* $\mathcal{C}$ *refers to constant modulus with phase alphabet constraint,*

$$
\mathbf{x}^\star = c \cdot e^{j\frac{2\pi}{K}\mathbf{k}^\star},
$$

*where* $\mathbf{k}^\star = \operatorname{round}\left(\frac{K}{2\pi}\arg(\mathbf{z})\right)$*, and the function* $\operatorname{round}(\mathbf{x})$ *denotes the element-wise rounding operation.*

Based on the MM method, to solve the original problem (6), a series of subproblems are solved with a closed-form solution at each iteration. We summarize the above algorithm in Algorithm 1.

[1] Throughout this paper, underlined variables denote those whose values are given as constants.

**Algorithm 1** MM Algorithm for Solving Problem (6)

**Input:** Covariance matrix $\boldsymbol{\Sigma}$,

Initialize $k \leftarrow 0$, $\mathbf{x}^{(0)}$.

**while** not converge **do**

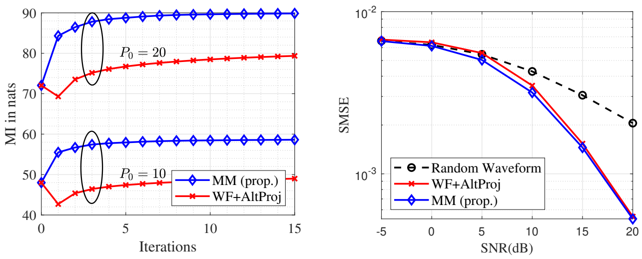
  Compute $\mathbf{V}, \mathbf{U}, \mathbf{w}, \mathbf{M}$.

  Compute $\mathbf{x}^{(k+1)}$ by Lemma 4.

  $k \leftarrow k + 1$.

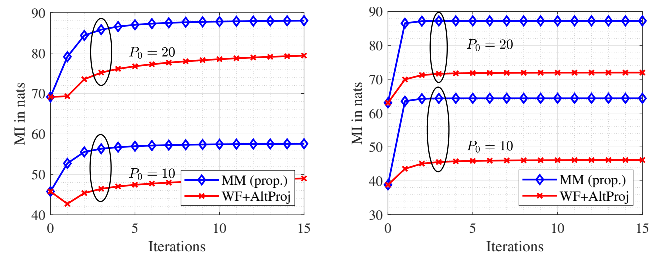
**end while**

**Output:** $\mathbf{x}^{(k)}$

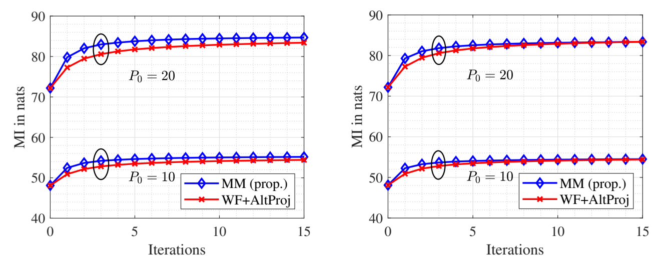



(a) power constraint. (b) sample mean square error.

Fig. 1: Convergence behavior with power constraint and the corresponding sample mean square error (averaged over 1000 Monte Carlo simulations).




(a) PAR ($\rho = 2$). (b) similarity ($\epsilon = 0.1$).




(c) constant modulus. (d) phase alphabet ($K = 256$).

Fig. 2: Comparison of the convergence behavior of the proposed MM method and the WF+AltProj method under various waveform constraints (averaged over 1000 Monte Carlo simulations).

### *C. A Gradient Projection Perspective on the MM Algorithm*

Problem (10) to be solved in each iteration of MM can be equivalently written as

$$\underset{\mathbf{x}\in\mathcal{C}}{\text{minimize}} \quad \left\|\mathbf{x} - \left(\underline{\mathbf{x}} - \lambda^{-1}\left(\mathbf{M}\underline{\mathbf{x}} - \mathbf{w}\right)\right)\right\|_2^2,$$

which is the orthogonal projection of $\underline{\mathbf{x}} - \lambda^{-1}\left(\mathbf{M}\underline{\mathbf{x}} - \mathbf{w}\right)$ onto the constraint set $\mathcal{C}$. At the $k$-th iteration, the update of $\mathbf{x}^{(k+1)}$ is given by

$$\mathbf{x}^{(k+1)} = \Pi_{\mathcal{C}}\left(\mathbf{x}^{(k)} + \lambda^{-1}\left(\mathbf{w} - \mathbf{M}\mathbf{x}^{(k)}\right)\right),$$

where $\Pi_{\mathcal{C}}(\mathbf{x}) = \arg\min_{\mathbf{s}\in\mathcal{C}} \|\mathbf{x} - \mathbf{s}\|_2$. Therefore, the proposed MM method can be understood as a gradient projection method for problem (6) with an adaptive step size $\lambda^{-1}$ and the ascent direction $\mathbf{w} - \mathbf{M}\mathbf{x}^{(k)}$.

## V. Numerical Experiments

We consider an MIMO radar system comprising $M = 4$ transmit antennas and $N = 4$ receive antennas. The transmitted and received signal lengths are $L_t = 20$ and $L_r = 20$, respectively. The impulse response length is $L = 5$. The covariance matrix of the impulse response is randomly generated from $\boldsymbol{\Sigma}_{\mathbf{h}} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{H}}$, where $\mathbf{U}$ is a unitary matrix and $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}_{ii} \sim U(0,1)$ for $i = 1, \ldots, L_t M$. The noise is assumed to obey $\mathcal{CN}(0,1)$ with $\sigma^2 = 1$. We compare the performance of the proposed MM algorithm and WF+AltProj algorithm [7] for the MI maximization problem.

In Fig. 1(a), we compare the convergence behavior with the power constraint. Fig. 1(a) indicates that the proposed MM algorithm achieves better MI than the AltProj under total power constraint with $P_0 = 10$ and $P_0 = 20$. In Fig. 1(b) we compare the algorithm performance in terms of sample mean square error (SMSE) defined as

$$\text{SMSE} = \frac{1}{N_s NML} \sum_{i=1}^{N_s} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{t=0}^{L-1} \left| h_{n,m}^{(i)}(t) - \hat{h}_{n,m}^{(i)}(t) \right|^2,$$

where $h_{n,m}^{(i)}(t)$, for $n = 1, \ldots, N$, $m = 1, \ldots, M$, $i = 1, \ldots, N_s$ is the target impulse response samples generated by the complex Gaussian distribution, with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{h}}$. $\hat{h}_{n,m}^{(i)}(t)$ is estimated by the Bayes estimator [6] with given waveform $\mathbf{x}$, received signal $\mathbf{y}$, and covariance $\boldsymbol{\Sigma}_{\mathbf{h}}$. The SNR $= 10\log_{10}\left(P_0/\sigma^2\right)$ is set between $-5\,\text{dB}$ and $20\,\text{dB}$. Fig. 1(b) indicates that the waveform optimized using MI as the optimization criterion has a lower SMSE compared to the random waveform, and the proposed MM algorithm achieves a lower SMSE compared to WF+AlProj. This is because MM takes into account the known Toeplitz structure and Kronecker product.

Fig. 2 shows the convergence performance of the proposed MM algorithm and WF+AltProj under similarity constraint (the reference waveform $\mathbf{x}^{\text{ref}}$ is generated by a random vector with total power $P_0$), PAR constraint, constant modulus constraint ($c$ is set to $\sqrt{P_0/(ML_t)}$), and phase alphabet constraint, respectively. In this case, we can see that the MM converges faster and achieves higher MI than WF+AltProj.

## VI. Conclusion

In this paper, we have investigated the problem of waveform design by maximizing the mutual information metric for MIMO radar systems, while having accounted for various practical waveform constraints. We have formulated the waveform design challenge as a nonconvex optimization problem and have developed an MM-based iterative algorithm to address it. Through numerical experiments, we have demonstrated that the proposed algorithm outperforms existing approaches.

## References


[1] X. Li and Y. Bar-Shalom, "Design of an Interacting Multiple Model Algorithm for Air Traffic Control Tracking," *IEEE Transactions on Control Systems Technology*, vol. 1, no. 3, pp. 186–194, 1993.

[2] G. Sun, Z. He, J. Tong, X. Yu, and S. Shi, "Mutual Information-Based Waveform Design for MIMO Radar Space-Time Adaptive Processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2909–2921, 2021.

[3] M. R. Bell, "Information Theory and Radar Waveform Design," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1578–1597, 1993.

[4] C. Smith and P. M. Goggans, "Radar Target Identification," *IEEE Antennas and Propagation Magazine*, vol. 35, no. 2, pp. 27–38, 1993.

[5] P. M. Woodward and I. Davies, "XCII. A Theory of Radar Information," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 41, no. 321, pp. 1001–1017, 1950.

[6] Y. Yang and R. 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.

[7] Y. Yang, R. S. Blum, Z. He, and D. R. Fuhrmann, "MIMO Radar Waveform Design via Alternating Projection," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1440–1445, 2010.

[8] Y. Yang and R. S. Blum, "Minimax Robust MIMO Radar Waveform Design," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 147–155, 2007.

[9] B. Tang, J. Tang, and Y. Peng, "MIMO Radar Waveform Design in Colored Noise Based on Information Theory," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4684–4697, 2010.

[10] Y. Chen, Y. Nijsure, C. Yuen, Y. H. Chew, Z. Ding, and S. Boussakta, "Adaptive Distributed MIMO Radar Waveform Optimization Based on Mutual Information," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 1374–1385, 2013.

[11] M. Bica, K.-W. Huang, V. Koivunen, and U. Mitra, "Mutual Information Based Radar Waveform Design for Joint Radar and Cellular Communication Systems," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3671–3675, 2016.

[12] 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.

[13] R. A. Romero, J. Bae, and N. A. Goodman, "Theory and Application of SNR and Mutual Information Matched Illumination Waveforms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 912–927, 2011.

[14] Z. Idriss, R. G. Raj, and R. M. Narayanan, "Waveform Optimization for Multistatic Radar Imaging Using Mutual Information," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2410–2425, 2021.

[15] 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.

[16] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2017.

[17] Tang, Bo, Tang, Jun, and Peng, Yingning, "MIMO Radar Waveform Design in Colored Noise Based on Information Theory," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4684–4697, 2010.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[19] M. R. Bell, *Information Theory and Radar: Mutual Information and the Design and Analysis of Radar Waveforms and Systems*. PhD Thesis, California Institute of Technology, 1988.

[20] E. Chen and M. Tao, "ADMM-Based Fast Algorithm for Multi-Group Multicast Beamforming in Large-Scale Wireless Systems," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2685–2698, 2017.

[21] Z. Zhao and D. P. Palomar, "MIMO Transmit Beampattern Matching Under Waveform Constraints," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3281–3285, 2018.

[22] A. De Maio, Y. Huang, M. Piezzo, S. Zhang, and A. Farina, "Design of Optimized Radar Codes With a Peak to Average Power Ratio Constraint," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2683–2697, 2011.

[23] 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.

[24] A. De Maio, S. De Nicola, Y. Huang, S. Zhang, and A. Farina, "Code Design to Optimize Radar Detection Performance Under Accuracy and Similarity Constraints," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5618–5629, 2008.

[25] 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.

[26] Y. Wang, X. Wang, H. Liu, and Z. Luo, "On the Design of Constant Modulus Probing Signals for MIMO Radar," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4432–4438, 2012.

[27] 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.

[28] H. Huang, M. F. Keskin, H. Wymeersch, X. Cai, L. Wu, J. Thunberg, and F. Tufvesson, "Hybrid Precoder Design for Angle-of-Departure Estimation with Limited-Resolution Phase Shifters," *IEEE Transactions on Communications*, pp. 1–1, 2024.

[29] D. R. Hunter and K. Lange, "A Tutorial on MM Algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[30] J. A. Tropp, I. S. Dhillon, R. W. Heath, and T. Strohmer, "Designing Structured Tight Frames via an Alternating Projection Method," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 188–209, 2005.