Signal Processing 225 (2024) 109635

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Optimization algorithms for transmit waveform design in radar-centric dual-function radar-communication systems☆

Licheng Zhao [a,*], Rui Zhou [a], Qingjiang Shi [a,b]

[a] Shenzhen Research Institute of Big Data, Shenzhen 518172, China
[b] School of Software Engineering, Tongji University, Shanghai 201804, China

ARTICLE INFO

Keywords:
Transmit waveform design
Dual-function radar-communication system
Constant modulus
Descent direction combination

ABSTRACT

We study the transmit waveform optimization problem in multiple-input-multiple-output (MIMO) dual-function radar-communication (DFRC) systems. A penalty-based formulation is adopted where engineering concerns on radar and communication are integrated in a weighted manner. We transform the original problem into an unconstrained formulation and propose a novel waveform synthesis algorithm via combination of multiple algorithm updates. The combination operation experimentally reaches a better objective value than without. We develop two combination strategies and invent a shrinkage-expansion line search method for monotonicity maintenance. Numerical experiments show that the proposed algorithm suffices to generate constant modulus waveforms which achieve radar and communication purposes.

1. Introduction

Nowadays, with the severe challenge of radio-frequency (RF) spectrum congestion, the DFRC design scheme has been gaining increasing popularity. DFRC systems have a vast array of applications ranging from civilian ones, such as radar and wireless systems [1], autonomous vehicle network [2], and indoor localization [3], to military ones, like multi-function RF systems [4], unmanned aerial vehicle (UAV) communication and sensing [5], and radar-assisted low-probability-of-intercept (LPI) communication [6]. DFRC systems integrate radar and communication functions into the same device [7] and are internally coordinative: the bandwidth, power, and antenna resources are shared and properly scheduled between radar functions and communication services [8,9]. Meanwhile, the hardware cost and weight can be reduced [10]. Early DFRC systems [11,12] were confined to a single antenna, but modern systems are designed in a MIMO fashion in exploitation of higher degrees of spatial freedom, in which case radar target tracking and wireless communication services can be supported instantly and simultaneously despite scattered locations of objects and users [8,13,14].

Liu et al. [8] have categorized the studies of MIMO DFRC systems into two philosophies: information embedding and transmit beamforming. The information embedding philosophy [15–21] suggests encoding communication data streams into radar waveforms by manipulating waveform parameters. This philosophy is radar-centric and ensures the primary status of the radar system, but limits the data transmission rate at a low level. The encoding process is traditionally formulated as a beampattern matching problem subject to a number of communication constraints. In latest studies, Wen et al. work on the minimum signal-to-interference-plus-noise ratio (SINR) maximization problem for target detection and investigate multi-objective optimization on integrated mainlobe-to-sidelobe ratio (IMSR) and waveform similarity [20,21]. In contrast, the transmit beamforming philosophy [22–26] is communication-centric and regards radar sensing as a secondary function. It supports a higher transmission rate, but poses a challenge in waveform synthesis which is expected to meet the functional requirements of radar and communication systems as well as mitigate their mutual interference. In this paper, we treat radar as the primary infrastructure and focus our research on the information embedding philosophy in what follows.

1.1. Related works

MIMO Radar and DFRC System. The MIMO radar technology is widely acknowledged due to its improved capabilities in adaptive localization and detection compared with a traditional phased-array radar [27]. This technology can be applied to strengthen or weaken

☆ The work of Licheng Zhao was supported in part by the National Nature Science Foundation of China under Grant 62206182 and in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515010154. The work of Rui Zhou was supported in part by the National Nature Science Foundation of China under Grant 62201362 and in part by the Shenzhen Science and Technology Program under Grant RCBS20221008093126071.

* Corresponding author.
E-mail addresses: zhaolicheng@sribd.cn (L. Zhao), rui.zhou@sribd.cn (R. Zhou), shiqj@tongji.edu.cn (Q. Shi).

https://doi.org/10.1016/j.sigpro.2024.109635
Received 5 May 2024; Received in revised form 17 July 2024; Accepted 25 July 2024
Available online 27 July 2024

the spatial power transmitted towards a range of locations and the power levels collected from all directions form a spatial beampattern. An important branch of the MIMO radar technology is the slow-time frequency-diverse array (FDA)-MIMO technique, which can be applied to the space–time adaptive processing (STAP) radar [28]. There are several practical concerns to be considered in the MIMO radar system. Previous works [27,29,30] considered the classical beampattern matching problem, whose objective is to minimize the Euclidean distance between the designed and desired beampatterns as well as the cross-correlation between user-specified directions. The pioneer work [27] only optimized the waveform covariance matrix while the later works [29,30] optimized the transmit waveforms directly with an additional constant modulus constraint. This constraint avoids nonlinear distortion of power amplifiers so that the system energy is fully utilized [30,31]. The classical transmit waveform design problem contains three concerns: beampattern matching, cross-correlation suppression, and constant modulus.

In DFRC systems, the requirements of beampattern matching and constant modulus are reserved and communication symbols are encoded on the power levels of some specified communication directions. This system is radar-centric and the communication service is only intended for sending short but secured instructional messages. In the existing the literature, [15,16] took ASK-based DFRC approaches where power amplitude is utilized for communication services. In contrast, [17–19] additionally exploited phasic freedom and adopted a quadrature-amplitude-modulation (QAM)-based scheme. For simplicity, we only consider the ASK-based DFRC techniques and control the power levels in the communication directions. Wherever the communication service is active, the power level in that direction is designated to some prescribed value in the codebook [15,16]. To summarize, radar-centric MIMO DFRC systems need to consider beampattern matching, constant modulus, and communication power level fitting.

Algorithmic Scheme. Back in 2007, Stoica et al. [27] formulated the beampattern matching problem as a SemiDefinite Programming (SDP) and the solution algorithm requested an off-the-shelf solver like SeDuMi [32]. When the constant modulus constraint is imposed, Wang et al. [29] reformulated the original problem into an unconstrained one by treating phase angles as optimization variables and solved the unconstrained problem with a limited memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS)-based algorithm; Cheng et al. [30] applied the Alternating Direction Method of Multipliers (ADMM) framework and solved the problem in an iterative primal–dual manner; earlier works [33,34] proposed a series of Majorization-Minimization (MM)-based algorithms and achieved monotonic objective decrease. However, the LBFGS-based algorithm can hardly be extended to constrained optimization problems and the numerical performance is not stable enough for nonconvex problems. As for ADMM, the convergence criteria are quite restrictive for general nonconvex problems [35,36] and only empirical convergence is guaranteed in spite of reasonable initialization and proper choice of penalty factors in the augmented Lagrangian. The convergence speed of the MM-based algorithms is not satisfactory without an acceleration scheme, due to the conservative step sizes derived from the inverse of matrix spectral norm. Moreover, the obtained descent directions coincide with gradient on many occasions [33]. Liu et al. [22] studied the classical MIMO radar beampattern matching problem as well as the beamforming strategies for separated and shared deployment. The classical problem was handled in the same way as [27]. The problem for separated deployment was solved via the semidefinite relaxation (SDR) technique while the one for shared deployment was transformed into a penalty formulation and tackled with Riemannian manifold optimization.

In radar-centric MIMO DFRC systems, the beampattern design problems [15–19] are mostly convex and solved with an off-the-shelf optimization toolbox like CVX, but fail to consider the constant modulus property in the waveform. In our paper, we intend to model the concerns of beampattern matching and communication power level fitting as two objective functions and integrate them in a weighted manner. The constant modulus property is treated as a constraint. Hence, the optimization problem to be studied is a penalty-based formulation. If the penalty-based formulation can be recast as an equivalent unconstrained problem, many advanced algorithmic frameworks are readily applicable, such as [37–39].[1] For nonconvex problems with numerous saddle points, applying a single algorithmic framework (like gradient descent) does not necessarily bring us satisfactory optimization outputs [41,42], so a desirable solution would be combining multiple potential descent directions in an effort to avoid saddle points [43–46]. Keskar and Socher [43] developed a feasible combination method which switches from algorithm Adam [38] to gradient descent. This method is by essence a violent one-shot move upon meeting the switching criterion. So far, it remains an open problem how to decently combine multiple algorithmic frameworks and relevant literature is still scarce.

1.2. Contribution

The major contributions of this paper are listed as follows:

(1) We transform the constrained optimization problem into an unconstrained one after a change of variable. We propose a novel shrinkage-expansion line search method to maintain objective monotonicity over iterations. The proposed line search method allows a customized initial step and adopts a step size expansion scheme. It enables a larger step size than initialized and helps promote computational efficiency.

(2) We develop a novel and provably convergent algorithm to solve the reformulated unconstrained problem, which combines two descent directions provided by gradient and Amsgrad [39]. The proposed algorithm experimentally achieves a better objective value than simply applying gradient or Amsgrad.

(3) We put forward two combination strategies: binary and probabilistic combination. The binary combination strategy maintains the monotonicity property in objective and attains stationarity convergence. The probabilistic combination strategy empirically reaches an even lower converged objective than binary combination and a convergence rate of $\mathcal{O}\left(\frac{\log T}{\sqrt{T}}\right)$ (T is the number of iterations) can be verified under certain assumptions. Numerical simulations show that the proposed algorithm is able to synthesize constant modulus waveforms satisfying radar and communication demands.

1.3. Notation

The following notation is adopted. Boldface upper-case letters represent matrices, boldface lower-case letters denote column vectors, and standard lower-case or upper-case letters stand for scalars. $\mathbb{R}$ ($\mathbb{C}$) denotes the real (complex) field. $\exp(\cdot)$ stands for the natural exponential function. $|\cdot|$ denotes the absolute value for the real case and modulus for the complex case. $\mathrm{Re}[\cdot]$ denotes the real part of a complex number. $\mathrm{sgn}(x) = 1$ if $x > 0$; $\mathrm{sgn}(x) = 0$ if $x = 0$; $\mathrm{sgn}(x) = -1$ if $x < 0$. $\mathbb{1}\{x \in \mathcal{X}\} = 1$ if $x \in \mathcal{X}$; $\mathbb{1}\{x \in \mathcal{X}\} = 0$ if $x \notin \mathcal{X}$. $\|\cdot\|_p$ denotes the ℓ_p-norm of a vector. $\nabla(\cdot)$ represents the gradient of a multivariate function (the way to derive the complex-valued gradient follows Euclidean gradient in [47]). $\odot$ stands for the Hadamard product. $\mathbf{X}^T$, $\mathbf{X}^*$, $\mathbf{X}^H$, $\mathrm{Tr}(\mathbf{X})$, and $\mathrm{vec}(\mathbf{X})$ denote the transpose, complex conjugate, conjugate transpose, trace, and stacking vectorization of $\mathbf{X}$, respectively. $\|\mathbf{X}\|_F$ is the Frobenius norm of $\mathbf{X}$. $\langle \mathbf{X}, \mathbf{Y} \rangle$ stands for the inner product of $\mathbf{X}$ and $\mathbf{Y}$, which is equivalent to $\mathrm{Tr}\left(\mathbf{X}^T\mathbf{Y}\right)$ for the real case and $\mathrm{Re}\left[\mathrm{Tr}\left(\mathbf{X}^H\mathbf{Y}\right)\right]$ for the complex case. The superscript $\star$ represents the optimal solution of some optimization problem. Whenever arithmetic operators ($\sqrt{\cdot}$, $\cdot/\cdot$, $\cdot^2$, $|\cdot|$, max, etc.) are applied to vectors or matrices, they are performed elementwisely.

[1] Nesterov's fast gradient method [40] is restricted to convex programming problems.

2. Problem statement

We consider a MIMO DFRC system with M transmit antennas. Each antenna emits a discrete-time complex-number waveform $x_m(n)$ with $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$ (N is the temporal length of the probing signal). We compactly stack all $x_m(n)$'s into a space–time waveform matrix $\mathbf{X} \in \mathbb{C}^{M\times N}$ with $\mathbf{X}_{mn} = x_m(n)$. The transmit steering vector is expressed as $\mathbf{a}(\theta) = \left[1, \exp(j\pi \sin\theta), \ldots, \exp(j\pi(M-1)\sin\theta)\right]^T$ with θ being the detection angle, under the assumption of half-wavelength spacing [27,29]. To perform waveform design, we integrate the concerns of beampattern matching, communication power level fitting, and constant modulus. The former two are modeled as loss functions and the last one is imposed as a constraint. Mathematically, we present the problem formulation as

$$\begin{aligned} \underset{q,\mathbf{X}}{\text{minimize}} \quad & f_{\text{BM}}(q,\mathbf{X}) + \tau f_{\text{CP}}(\mathbf{X}) \\ \text{subject to} \quad & |\mathbf{X}_{mn}| = \frac{1}{\sqrt{MN}},\ \forall m, n. \end{aligned} \tag{1}$$

where $f_{\text{BM}}(q,\mathbf{X})$ is the beampattern matching loss, expressed as

$$f_{\text{BM}}(q,\mathbf{X}) = \frac{1}{I}\sum_{i=1}^{I}\left(qP_d(\theta_i) - \left\|\mathbf{a}^H(\theta_i)\mathbf{X}\right\|_2^2\right)^2, \tag{2}$$

$f_{\text{CP}}(\mathbf{X})$ is the communication power level fitting loss, expressed as

$$f_{\text{CP}}(\mathbf{X}) = \frac{1}{L}\sum_{l=1}^{L}\left(\log\left(\left\|\mathbf{a}^H(\theta_l)\mathbf{X}\right\|_2^2\right) - \log(\Delta_l)\right)^2, \tag{3}$$

and $\tau > 0$ is the integration parameter. Relevant notational and modeling explanations are elaborated below. $P_d(\theta_i)$ is the desired binary or normalized beampattern at θ_i, q represents the actual amplitude of $P_d(\theta_i)$'s, the set $\Theta_I = \{\theta_i\}_{i=1}^I$ contains all the fine-grid angular directions of interest. The communication symbols are modulated into the ASK scheme. The set $\Theta_L = \{\theta_l\}_{l=1}^L \subset \Theta_I$ consists of all the communication directions and the communication power at θ_l is $\left\|\mathbf{a}^H(\theta_l)\mathbf{X}\right\|_2^2$. Δ_l is the predetermined communication power level at θ_l and is selected from the power level set $\{\Delta_{\text{High}}, \Delta_{\text{Low}}\}$ with $\Delta_{\text{High}} > \Delta_{\text{Low}}$. In engineering practice, Δ_l's are very closed to zero and a linear fitting may fail due to limited numerical precision. Thus, we fit $\left\|\mathbf{a}^H(\theta_l)\mathbf{X}\right\|_2^2$ to Δ_l on a logarithmic scale. The waveform matrix $\mathbf{X}$ has a unit power budget and is elementwisely constant modulus due to power and hardware concerns [48]. Due to the primary status of radar infrastructure, the communication concern should be down-weighted, so the choice of τ actually lies within $(0, 1]$. The DFRC scheme is achieved by weighted-sum minimization of the beampattern matching loss and the communication power level fitting loss.

3. Multiobjective beampattern matching design

3.1. Reformulation to an unconstrained problem

To solve problem (1), we first minimize the objective with respect to q. For any given $\mathbf{X}$, the optimal solution to q is

$$q^\star(\mathbf{X}) = \frac{\sum_{i=1}^I P_d(\theta_i)\left\|\mathbf{a}^H(\theta_i)\mathbf{X}\right\|_2^2}{\sum_{i=1}^I P_d^2(\theta_i)}. \tag{4}$$

Substituting $q^\star(\mathbf{X})$ into the beampattern matching loss $f_{\text{BM}}(q,\mathbf{X})$, we obtain

$$\begin{aligned} f_{\text{BM}}(q^\star(\mathbf{X}),\mathbf{X}) = {} & \frac{1}{I}\sum_{i=1}^{I}\left\|\mathbf{a}^H(\theta_i)\mathbf{X}\right\|_2^4 \\ & - \frac{\left(\sum_{i=1}^I P_d(\theta_i)\left\|\mathbf{a}^H(\theta_i)\mathbf{X}\right\|_2^2\right)^2}{I\sum_{i=1}^I P_d^2(\theta_i)} \triangleq \bar{f}_{\text{BM}}(\mathbf{X}). \end{aligned} \tag{5}$$

To handle the constant modulus constraint, we consider a variable transform on $\mathbf{X}$: $\mathbf{X} = \frac{1}{\sqrt{MN}}\frac{\mathbf{Y}}{|\mathbf{Y}|}$. This transform removes the constraint but causes a nonsmoothness issue in $\{\mathbf{Y}|\mathbf{Y}_{mn} = 0, \forall m, n\}$. For the sake of objective differentiability, we slightly relax the constant modulus constraint and suggest a smooth surrogate: $\mathbf{X} = \frac{1}{\sqrt{MN}}\frac{\mathbf{Y}}{|\mathbf{Y}|+\epsilon}$ with $\epsilon > 0$ being sufficiently small. A similar smooth transform has already appeared in [38,49], and this smooth transform belongs to a family of smoothing techniques which are widely acknowledged to be effective in nonsmooth optimization with convergence guarantee [50,51]. Thus, problem (1) becomes unconstrained:

$$\underset{\mathbf{Y}}{\text{minimize}} \quad \bar{f}_{\text{BM}}\left(\frac{1}{\sqrt{MN}}\frac{\mathbf{Y}}{|\mathbf{Y}|+\epsilon}\right) + \tau f_{\text{CP}}\left(\frac{1}{\sqrt{MN}}\frac{\mathbf{Y}}{|\mathbf{Y}|+\epsilon}\right) \triangleq f(\mathbf{Y}). \tag{6}$$

Problem (6) opens up the possibility for advanced algorithm design and integration of multiple frameworks. The most commonly used optimization algorithm is gradient descent and the update equation is

$$\mathbf{Y}_{t+1} = \mathbf{Y}_t - \alpha_t \mathbf{G}_t, \tag{7}$$

where the subscripts t and $t+1$ represent numbers of iterations, $\mathbf{G}_t = \nabla f(\mathbf{Y}_t)$, and α_t is the step size at iteration t. Reliable numerical gradients are available from the automatic differentiation engines in machine learning platforms, so we do not have to express $\mathbf{G}_t$ manually. The gradient descent framework computes the negative gradient as the descent direction. In the following, we introduce the framework Amsgrad which offers a different direction for the descending purpose.

Remark 1. In theory, the smooth surrogate is almost equivalent to the original term when $\epsilon \to 0$. Numerically, the effect of an extra ϵ on the denominator is negligible when ϵ is down to 10^{-8}. A constant modulus rounding could be executed on the converged result in case of a strict feasibility request.

3.2. Preliminaries: The Amsgrad framework

The Amsgrad algorithm [39] provides a different descent direction from the negative gradient and demonstrates a fast empirical convergence speed in deep learning applications like Computer Vision and Natural Language Processing. The Amsgrad algorithm adopts the concepts of (1) momentum, which is a weighted moving average of the current state and the update term, and applied on the first- and second-order moments of gradient $\mathbf{G}_t$:

$$\mathbf{M}_t = \beta_1\mathbf{M}_{t-1} + (1-\beta_1)\mathbf{G}_t \quad \text{and} \quad \mathbf{V}_t = \beta_2\mathbf{V}_{t-1} + (1-\beta_2)(\mathbf{G}_t^* \odot \mathbf{G}_t); \tag{8}$$

(2) bias correction, which adjusts $\mathbf{M}_t$ and $\mathbf{V}_t$ with a scalar:

$$\hat{\mathbf{M}}_t = \mathbf{M}_t/(1-\beta_1^t) \quad \text{and} \quad \hat{\mathbf{V}}_t = \mathbf{V}_t/(1-\beta_2^t); \tag{9}$$

and (3) per-coordinate adaptation, which elementwisely divides $-\hat{\mathbf{M}}_t$ by $\sqrt{\hat{\mathbf{V}}_t}$ as a potential update direction. Furthermore, Amsgrad incorporates a long-term memory for elementwise maximum of all historical $\hat{\mathbf{V}}_t$'s to avoid theoretical divergence. To sum up, the update equations of Amsgrad are

$$\begin{aligned} &\mathbf{M}_t = \beta_1\mathbf{M}_{t-1} + (1-\beta_1)\mathbf{G}_t, \\ &\hat{\mathbf{M}}_t = \mathbf{M}_t/(1-\beta_1^t); \end{aligned} \tag{10}$$

$$\begin{aligned} &\mathbf{V}_t = \beta_2\mathbf{V}_{t-1} + (1-\beta_2)(\mathbf{G}_t^* \odot \mathbf{G}_t), \\ &\hat{\mathbf{V}}_t = \max\left(\hat{\mathbf{V}}_{t-1}, \mathbf{V}_t/(1-\beta_2^t)\right); \end{aligned} \tag{11}$$

$$\mathbf{Y}_{t+1} = \mathbf{Y}_t - \alpha_t\hat{\mathbf{M}}_t/\sqrt{\hat{\mathbf{V}}_t}. \tag{12}$$

Amsgrad also provides an empirical descent direction $-\hat{\mathbf{M}}_t/\sqrt{\hat{\mathbf{V}}_t}$. $\mathbf{M}_0$, $\mathbf{V}_0$, and $\hat{\mathbf{V}}_0$ are initialized with all-zero matrices. The good default

settings of α, β_1, and β_2 follow [38, Algorithm 1], which have been tested on many engineering problems. Both $\mathbf{M}_t$ and $\mathbf{V}_t$ have closed-form expressions:

$$\mathbf{M}_t = (1-\beta_1)\sum_{s=1}^{t}\beta_1^{t-s}\mathbf{G}_s \quad \text{and} \quad \mathbf{V}_t = (1-\beta_2)\sum_{s=1}^{t}\beta_2^{t-s}\left(\mathbf{G}_s^* \odot \mathbf{G}_s\right). \tag{13}$$

3.3. Descent direction combination

Both gradient descent and Amsgrad algorithms can provide descent directions, given as $-\mathbf{G}_t$ and $-\hat{\mathbf{M}}^{(t)}/\sqrt{\hat{\mathbf{V}}^{(t)}}$, respectively. For simple notation, we denote them as $\Delta\mathbf{Y}_{t,\mathrm{GD}}$ and $\Delta\mathbf{Y}_{t,\mathrm{Amsgrad}}$. The corresponding step sizes are denoted as $\alpha_{t,\mathrm{GD}}$ and $\alpha_{t,\mathrm{Amsgrad}}$. Our proposed algorithmic scheme is to derive a convex combination of the two directions rescaled by their step sizes, i.e.,

$$\mathbf{Y}_{t+1} = \mathbf{Y}_t + w_t\cdot\left(\alpha_{t,\mathrm{GD}}\Delta\mathbf{Y}_{t,\mathrm{GD}}\right) + \left(1-w_t\right)\cdot\left(\alpha_{t,\mathrm{Amsgrad}}\Delta\mathbf{Y}_{t,\mathrm{Amsgrad}}\right) \tag{14}$$

with $w_t \in [0,1]$. We are going to propose two combination strategies based on this update scheme.

3.3.1. Binary combination strategy

The motivation of binary combination is to select one of the two descent directions and attain as large an objective decrease as possible. We propose a shrinkage-expansion line search method to obtain the step sizes $\alpha_{t,\mathrm{GD}}$ and $\alpha_{t,\mathrm{Amsgrad}}$. The detailed implementation is elaborated in Algorithm 1. The intuition of this line search method is that, if the actual objective decrease D_A is less than the expected decrease D_E, the step size shrinks iteratively and the output is smaller than the initial step size α_0; otherwise the step size keeps expanding until the actual decrease fails to exceed the expected one and the output is larger than initial. Compared with the traditional Armijo step size rule, the shrinkage-expansion line search method has the following advantages: (1) a customized initial step size can reduce the number of attempts in the while loop; and (2) the expansion mechanism enables a more aggressive step size than initialized. These advantages are expected to promote computational efficiency for maintaining monotonicity. Upon obtaining $(\alpha_{t,\mathrm{GD}},\ D_{t,\mathrm{GD}})$ and $(\alpha_{t,\mathrm{Amsgrad}},\ D_{t,\mathrm{Amsgrad}})$, we choose the one that yields a larger decrease as the iteration update, i.e.,

$$w_t = \mathbb{1}\left\{D_{t,\mathrm{GD}} \geq D_{t,\mathrm{Amsgrad}}\right\}. \tag{15}$$

This strategy guarantees monotonic objective decrease.

Algorithm 1 Shrinkage-Expansion Line Search Method.

Require: descent parameter $p_{\mathrm{des}} \in (0, 0.5)$, shrinkage parameter $p_{\mathrm{shk}} \in (0,1)$, expansion parameter $p_{\exp} \in (1, +\infty)$, initial step size α_0;
Set step size $\alpha = \alpha_0$ and $\left[D_A, D_E\right] = \left[f(\mathbf{Y}) - f(\mathbf{Y}+\alpha\cdot\Delta\mathbf{Y}), -p_{\mathrm{des}}\cdot\alpha\cdot\langle\mathbf{G},\Delta\mathbf{Y}\rangle\right]$;
$\mathrm{sign}_{\mathrm{init}} = \mathrm{sgn}\left(D_A - D_E\right)$; $\mathrm{sign} = \mathrm{sign}_{\mathrm{init}}$;
$u = p_{\mathrm{shk}} \times\!\!\!\!\!\!\diagup\ \{\mathrm{sign} < 0\} + p_{\exp} \times\!\!\!\!\!\!\diagup\ \{\mathrm{sign} \geq 0\}$;
while $\mathrm{sign} = \mathrm{sign}_{\mathrm{init}}$ **do**
  $\alpha = u\alpha$; $\left[D_A, D_E\right] = \left[f(\mathbf{Y}) - f(\mathbf{Y}+\alpha\cdot\Delta\mathbf{Y}), -p_{\mathrm{des}}\cdot\alpha\cdot\langle\mathbf{G},\Delta\mathbf{Y}\rangle\right]$;
  $\mathrm{sign} = \mathrm{sgn}\left(D_A - D_E\right)$;
end while
if $\mathrm{sign}_{\mathrm{init}} \geq 0$ **then**
  $\alpha = \alpha/u$; $D_A = f(\mathbf{Y}) - f(\mathbf{Y}+\alpha\cdot\Delta\mathbf{Y})$;
end if
Obtain step size α and actual decrease D_A as outputs.

3.3.2. Probabilistic combination strategy

Although binary selection of descent directions improves the abrupt algorithm switch, this strategy discards one descent direction in every iteration. In order to take full advantage of the two computed directions, we develop a heuristic but useful combination strategy which dynamically mixes the two directions with reference to the potential objective improvements. $\left(\alpha_{t,\mathrm{GD}}, D_{t,\mathrm{GD}}\right)$ and $\left(\alpha_{t,\mathrm{Amsgrad}}, D_{t,\mathrm{Amsgrad}}\right)$ are still obtained from Algorithm 1. The combination weight is obtained from the softmax operation of the objective decreases, i.e.,

$$w_t = \frac{\exp\left(D_{t,\mathrm{GD}}\right)}{\exp\left(D_{t,\mathrm{GD}}\right) + \exp\left(D_{t,\mathrm{Amsgrad}}\right)}. \tag{16}$$

The softmax operation has a probabilistic interpretation in the field of machine learning and that is what the term "probabilistic combination" is for. The essence of this strategy is that, the larger the objective decrease, the larger the combination weight that should be assigned in that particular direction. This strategy does not guarantee monotonicity but empirically achieves an even lower converged objective than its binary counterpart. To summarize, the complete algorithm is presented in Algorithm 2.

Table 1
Per-iteration computational cost of the proposed algorithms in different stages.

Descent direction computation	Gradient computation, $\mathcal{O}(MN(I+L))$ Elementwise arithmetic operations, $\mathcal{O}(MN)$
Step size derivation	Inner product, $\mathcal{O}(MN)$ Objective computation, $\mathcal{O}(MN(I+L))$
Combination weight calculation	Scalar calculation, $\mathcal{O}(1)$
Variable update	Elementwise arithmetic operations, $\mathcal{O}(MN)$

Algorithm 2 Descent Direction Combination with Adaptive Step (DDCAS).

Require: $\mathbf{Y}_1$ with initialization, $\mathbf{M}_0$, $\mathbf{V}_0$, and $\hat{\mathbf{V}}_0$ initialized as $\mathbf{0}$, $t = 1$;
repeat
  $\Delta\mathbf{Y}_{t,\mathrm{GD}} = -\mathbf{G}_t$ where $\mathbf{G}_t = \nabla f\left(\mathbf{Y}_t\right)$;
  Compute $\hat{\mathbf{M}}_t$ and $\hat{\mathbf{V}}_t$ using (10) and (11), respectively;
  $\Delta\mathbf{Y}_{t,\mathrm{Amsgrad}} = -\hat{\mathbf{M}}_t/\sqrt{\hat{\mathbf{V}}_t}$;
  Compute $(\alpha_{t,\mathrm{GD}},\ D_{t,\mathrm{GD}})$ and $(\alpha_{t,\mathrm{Amsgrad}},\ D_{t,\mathrm{Amsgrad}})$ with the shrinkage-expansion line search method (Algorithm 1);
  $w_t = \begin{cases} \mathbb{1}\left\{D_{t,\mathrm{GD}} \geq D_{t,\mathrm{Amsgrad}}\right\} & \textbf{Binary Combination} \\ \frac{\exp(D_{t,\mathrm{GD}})}{\exp(D_{t,\mathrm{GD}})+\exp(D_{t,\mathrm{Amsgrad}})} & \textbf{Probabilistic Combination} \end{cases}$;
  $\mathbf{Y}_{t+1} = \mathbf{Y}_t + w_t\cdot\left(\alpha_{t,\mathrm{GD}}\Delta\mathbf{Y}_{t,\mathrm{GD}}\right) + \left(1-w_t\right)\cdot\left(\alpha_{t,\mathrm{Amsgrad}}\Delta\mathbf{Y}_{t,\mathrm{Amsgrad}}\right)$;
  $t = t+1$;
until convergence

3.4. Computational complexity

Computational complexity analysis is carried out on a per-iteration basis. Algorithm DDCAS follows a four-stage paradigm: (1) descent direction computation, (2) step size derivation, (3) combination weight calculation, and (4) variable update. The detailed computational cost of the proposed algorithm is listed in Table 1. The processes of gradient and objective computation can be efficiently implemented in machine learning platforms, whose complexities are both $\mathcal{O}(MN(I+L))$. The inner product of gradient and the descent direction costs $\mathcal{O}(MN)$ if proper implementation is carried out: $\langle\mathbf{G},\Delta\mathbf{Y}\rangle = \mathrm{Re}\left[\mathbf{1}^T(\mathbf{G}\odot\Delta\mathbf{Y})\mathbf{1}\right]$. Therefore, the per-iteration complexity of Algorithm DDCAS amounts to $\mathcal{O}(MN(I+L))$.

4. Convergence analysis

The convergence analysis of Algorithm DDCAS is given as follows with respect to two combination strategies.

4.1. Binary combination

Theorem 2. *The iterates generated by Algorithm DDCAS with the binary combination strategy converge to a stationary point of problem (6).*

Proof. The stationarity convergence is derived from objective monotonicity. Please refer to Appendix A in the supplementary document. □

4.2. Probabilistic combination

The probabilistic combination strategy merely enjoys empirical convergence on general occasions, but a convergence rate of $\mathcal{O}\left(\frac{\log T}{\sqrt{T}}\right)$ can be obtained under the prerequisites of bounded gradients and objective Lipschitz continuity. The two assumptions are explained in the following:

- Upper bounded gradient: for any $t \geq 1$ and any m, n, $\left|\mathbf{G}_{t,mn}\right| \leq G$;
- Lower bounded gradient (the algorithm does not converge at the initial point): for any element m, n, $\left|\mathbf{G}_{1,mn}\right| \geq c$;
- Lipschitz continuous objective: for any $\mathbf{X}_1$ and $\mathbf{X}_2$, there exists a constant $L > 0$ that satisfies $f\left(\mathbf{X}_2\right) \leq f\left(\mathbf{X}_1\right) + \left\langle \nabla f\left(\mathbf{X}_1\right), \mathbf{X}_2 - \mathbf{X}_1 \right\rangle + \frac{L}{2}\left\|\mathbf{X}_2 - \mathbf{X}_1\right\|_F^2$, or equivalently $\left\|\nabla f\left(\mathbf{X}_1\right) - \nabla f\left(\mathbf{X}_2\right)\right\|_F \leq L\left\|\mathbf{X}_1 - \mathbf{X}_2\right\|_F$.

The bounded gradient assumption is inherited from [38,52]. Most objective functions satisfy Lipschitz continuity, including $f(\mathbf{Y})$. The bounded gradient assumption leads to the boundedness of $\mathbf{M}_t$, $\hat{\mathbf{M}}_t$, $\mathbf{V}_t$, and $\hat{\mathbf{V}}_t$, as is shown in the following lemma.

Lemma 3. *When gradient elements are upper and lower bounded, the bounded properties of $\mathbf{M}_t$, $\hat{\mathbf{M}}_t$, $\mathbf{V}_t$, and $\hat{\mathbf{V}}_t$ are given as: for any $t \geq 1$ and any m, n,*

$$\begin{aligned} &\left|\mathbf{M}_{t,mn}\right| \leq \left|\hat{\mathbf{M}}_{t,mn}\right| \leq G, \\ &\mathbf{V}_{t,mn} \leq G^2, \quad \textit{and} \\ &c^2 \leq \hat{\mathbf{V}}_{t,mn} \leq G^2. \end{aligned} \tag{17}$$

Proof. Please refer to Appendix B in the supplementary document. □

The convergence rate of Algorithm DDCAS under a constant combination weight is given in Theorem 4.

Theorem 4. *Under the assumptions of bounded gradients and a Lipschitz continuous objective, when the combination weight remains constant and the step size adopts a square-root-decaying fashion ($\alpha_t = \frac{\alpha}{\sqrt{t}}$ for $t = 1, 2, \ldots, T$), the convergence rate of Algorithm DDCAS is given as*

$$\min_{t=1,2,\ldots,T} \left\|\mathbf{G}_t\right\|_F^2 = \mathcal{O}\left(\frac{\log T}{\sqrt{T}}\right) \tag{18}$$

regarding problem (6) with $\epsilon \to 0$.

Proof. Please refer to Appendix C in the supplementary document. □

5. Numerical simulations

In this section, we present numerical results of multiobjective waveform design. All simulations are performed on a PC with a 2.90 GHz i7-10700 CPU and 16.0 GB RAM.

Default Experiment Settings. We consider a MIMO DFRC system with $M = 32$ transmit antennas. The probing signal length N equals to 64. The angular range set Θ_I is $(-90°, 90°)$ with a uniform spacing of 1°. Define $\Theta_J = \{-40°, 0°, 40°\}$ and the desired binary beampattern is

$$P_d(\theta) = \begin{cases} 1 & \theta \in \cup_{\theta_j \in \Theta_J}\left[\theta_j - 10°, \theta_j + 10°\right], \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

The communication direction set is $\Theta_L = \{-80°, 80°\}$ and the power level set is $\{-20\ \text{dB}, -50\ \text{dB}\}$. To demonstrate the primary infrastructure status of radar, we set the tuning parameters τ to be 0.1. Each element of variable $\mathbf{Y}$ is independently initialized with a complex Gaussian distribution $\mathcal{CN}(0, 0.1)$. The ϵ parameter in the smooth surrogate is fixed as 10^{-8}. In Algorithm 1, parameters p_{des}, p_{shk}, p_{exp}, and α_0 are set to be 0.3, 0.9, 1.1, and 5×10^{-3}. The aforementioned settings remain unchanged unless otherwise specified.

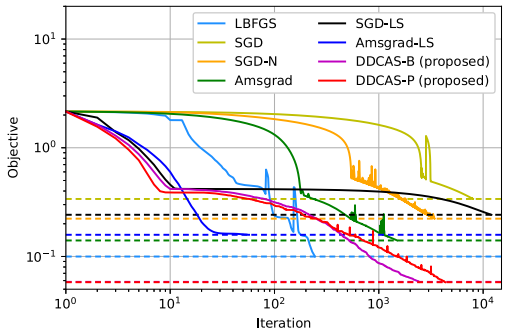


Fig. 1. Objective convergence plot: objective versus iteration. Dashed lines represent the converged objective levels.

5.1. Objective convergence

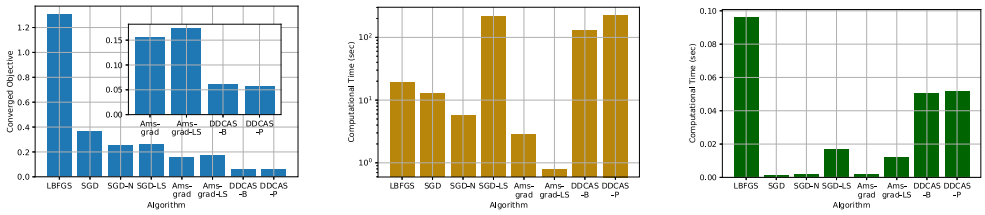
We include four existing general algorithm frameworks as benchmarks, namely, LBFGS, SGD, SGD with Nesterov acceleration (SGD-N in short), and Amsgrad. For further exploration, we incorporate the proposed line search method into two benchmark algorithms: SGD and Amsgrad (short for SGD-LS and Amsgrad-LS). The two combination strategies in the proposed algorithm are investigated separately, so the proposed algorithm is abbreviated as DDCAS-B (binary combination) and DDCAS-P (probabilistic combination).

5.1.1. Single realization performance

Objective convergence performances of different algorithms are presented in Fig. 1. The converged objectives of DDCAS-B and DDCAS-P are lower than those of benchmarks: in detail, DDCAS-P achieves the lowest objective upon termination (0.05799) but its convergence trajectory is not monotonic, possibly with a few spikes; DDCAS-B achieves the second lowest converged objective (0.05841) but the function values decrease smoothly with iterations. This can be ascribed to the proposed shrinkage-expansion line search method as well as combination of two descent directions. The line search method guarantees an objective decrease and mitigates oscillation in convergence trajectory (cf. [DDCAS-B, DDCAS-P, SGD-LS, Amsgrad-LS] versus [LBFGS, SGD, SGD-N, Amsgrad]). With descent direction combination, the update dynamics may result in a relatively lower converged objective (cf. [DDCAS-B, DDCAS-P] versus [SGD, SGD-LS, Amsgrad, Amsgrad-LS]).

5.1.2. Multiple realization performance

For performance generalization, we have tested the proposed algorithm and the benchmarks with 100 random initializations and the reported results are averaged over the 100 instances. In Fig. 2, we present the converged objectives, complete computational time, and per-iteration computational time achieved by different algorithms. In terms of average converged objective (Fig. 2(a)), the proposed DDCAS-B and DDCAS-P reach the lowest levels of all the compared optimization algorithms, about one-third of the best benchmark algorithms (Amsgrad and Amsgrad-LS). In detail, DDCAS-P (0.05793) slightly beats DDCAS-B (0.06136) by 5.6 percent. This is because the probabilistic combination strategy takes full advantage of the two computed directions and dynamically adjusts the combination weight. Empirically, DDCAS-P obtains an even lower objective value but its convergence trajectory experiences oscillation, while DDCAS-B reduces the objective step by step and makes progress in every iteration. As a sacrifice, DDCAS-B and DDCAS-P take longer computational time than benchmarks. In terms of complete and per-iteration computational time (Figs. 2(b) and 2(c)), DDCAS-B and DDCAS-P are among the longest of all. The complete computational time is approximately 130 s and 220 s, and the per-iteration computational time is about 0.050 s and 0.052 s. This is because DDCAS-B and DDCAS-P apply the line search mechanism on



(a) Converged objective in average. (b) Complete computational time in average. (c) Per-iteration computational time in average.

Fig. 2. Converged objective, complete computational time, and per-iteration computation time of different optimization algorithms, averaged over 100 instances.

two descent directions while other benchmarks either skip the line search step or involve only one descent direction.

To conclude, the above simulations prove the necessity of the shrinkage-expansion line search method and descent direction combination. DDCAS-P is recommended when the achieved objective value is prioritized; DDCAS-B is preferred when progressive objective improvement is required.

5.2. Waveform design

We will show the multiobjective waveform design with the following experiments. Given waveform $\mathbf{X}$, beampattern at location θ_i is calculated as

$$\text{Beampattern}\left(\theta_i\right) = \left\|\mathbf{a}^H\left(\theta_i\right)\mathbf{X}\right\|_2^2. \tag{20}$$

The proposed algorithms are DDCAS-B and DDCAS-P. For comparison, we consider three benchmarks:

- The first benchmark is the classical two-stage approach, which solves the modified waveform covariance matrix problem (21) in the first stage and synthesize the waveform using alternating minimization [53] afterwards for a constant modulus waveform.

$$\begin{aligned} \underset{q,\mathbf{R}}{\text{minimize}} \quad & \frac{1}{I}\sum_{i=1}^{I}\left(qP_d\left(\theta_i\right) - \mathbf{a}^H\left(\theta_i\right)\mathbf{R}\mathbf{a}\left(\theta_i\right)\right)^2 \\ \text{subject to} \quad & \mathbf{R}_{mm} = \frac{1}{M},\ \forall m, \qquad \mathbf{R} \succeq \mathbf{0} \\ & \mathbf{a}^H\left(\theta_l\right)\mathbf{R}\mathbf{a}\left(\theta_l\right) = \Delta_l,\ \forall l. \end{aligned} \tag{21}$$

  We model the communication power level fitting concern as a few linear constraints. This benchmark is shortly written as "SDP + AltMin".
- The second benchmark follows the unconstrained reformulation in [29], which treats phase angles $\boldsymbol{\Theta}$ as optimization variables and solves:

$$\underset{q,\boldsymbol{\Theta}}{\text{minimize}} \quad f_{\text{BM}}\left(q, \mathbf{X}\left(\boldsymbol{\Theta}\right)\right) + \tau f_{\text{CP}}\left(\mathbf{X}\left(\boldsymbol{\Theta}\right)\right) \tag{22}$$

  with $\mathbf{X}(\boldsymbol{\Theta}) = \frac{1}{\sqrt{MN}}\exp\left(j\boldsymbol{\Theta}\right)$. This problem is solved with an LBFGS-based algorithm in [29], shortly written as "$\boldsymbol{\Theta}$-LBFGS".
- The third benchmark (state-of-the-art) is inherited from [15–17]. This benchmark must be adapted so as to cater for the waveform matrix design and constant modulus constraint. After a reasonable modification, we solve the following minimax optimization problem:

$$\begin{aligned} \underset{q,\mathbf{R}}{\text{minimize}} \quad & \max_{\{\theta_i : P_d(\theta_i)\neq 0\}} \left|q - \mathbf{a}^H\left(\theta_i\right)\mathbf{R}\mathbf{a}\left(\theta_i\right)\right| \\ \text{subject to} \quad & \mathbf{a}^H\left(\theta_i\right)\mathbf{R}\mathbf{a}\left(\theta_i\right) \le \varepsilon_0,\ \forall \theta_i : P_d\left(\theta_i\right) = 0 \\ & \mathbf{R}_{mm} = \frac{1}{M},\ \forall m, \qquad \mathbf{R} \succeq \mathbf{0} \\ & \mathbf{a}^H\left(\theta_l\right)\mathbf{R}\mathbf{a}\left(\theta_l\right) = \Delta_l,\ \forall l. \end{aligned} \tag{23}$$

  We set $\varepsilon_0 = 0.1$. The waveform matrix is generated using alternating minimization [53] afterwards for a constant modulus waveform. This benchmark is shortly written as "Minimax + AltMin".

Waveform design performance will be investigated in the following four scenarios:

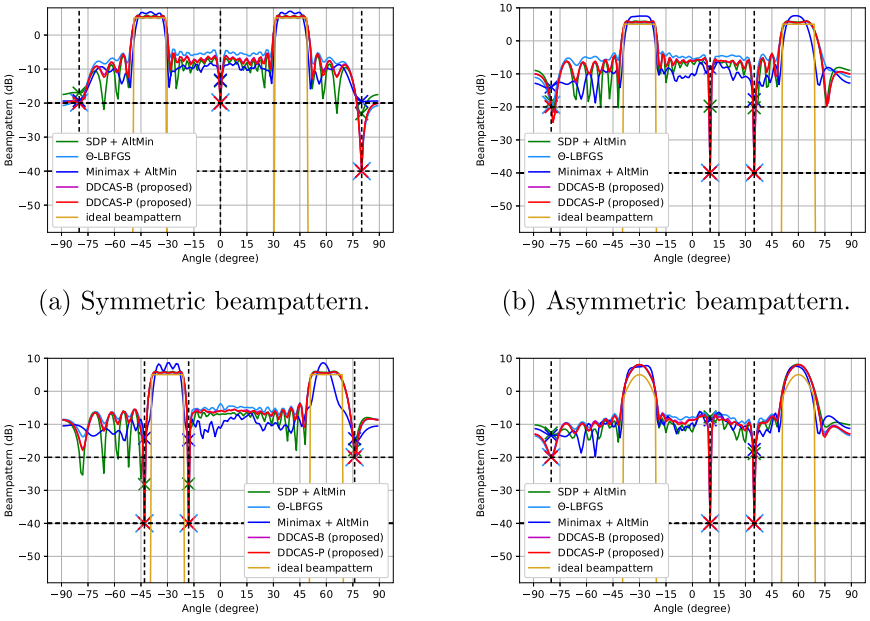
- **Scenario I** (Symmetric beampattern): the target location set is $\Theta_J = \{-40°, 40°\}$, the communication direction set is $\Theta_L = \{-80°, 0°, 80°\}$, and the corresponding power level is $\{-20\text{ dB}, -20\text{ dB}, -40\text{ dB}\}$.
- **Scenario II** (Asymmetric beampattern): the target location set is $\Theta_J = \{-30°, 60°\}$, the communication direction set is $\Theta_L = \{-80°, 10°, 35°\}$, and the corresponding power level is $\{-20\text{ dB}, -40\text{ dB}, -40\text{ dB}\}$.
- **Scenario III** (Asymmetric beampattern with close-to-target communication requirements): the target location set is $\Theta_J = \{-30°, 60°\}$, the communication direction set is $\Theta_L = \{-43°, 18°, 76°\}$, and the corresponding power level is $\{-40\text{ dB}, -40\text{ dB}, -20\text{ dB}\}$.
- **Scenario IV** (Asymmetric beampattern with Gaussian windows): $P_d(\theta)$ takes a Gaussian shape. Other detailed settings follow Scenario II.

5.2.1. Single realization performance

We present the beampattern comparison results in Fig. 3. Vertical dashed lines show the communication directions and horizontal dashed lines indicate the desired communication power levels. A "cross" marker indicates the power level in the particular direction derived from a certain algorithm. The beampatterns produced by "SDP + AltMin" and/or "Minimax + AltMin" have the lowest sidelobe levels of all, but fail to meet the communication power requirements, especially where the required power level is −40 dB. The mainlobes produced by "Minimax + AltMin" are likely to deteriorate when the ideal beampattern is asymmetric. This is because the waveform synthesis stage "AltMin" wipes out the effects of the communication power constraints and possibly the beampattern matching objective. The beampatterns produced by $\boldsymbol{\Theta}$-LBFGS satisfy the communication power requirements in all the four scenarios, but the sidelobe levels are relatively higher on most occasions. The reason could be that the LBFGS-based algorithm terminates in a plateau region or converges to a low-quality local minimum. Numerous local minima spread across the domain due to the periodic property of $\exp(j\boldsymbol{\Theta})$. In contrast, the proposed algorithms DDCAS-B and DDCAS-P satisfy the beampattern matching and communication power requirements simultaneously. Note that the beampatterns produced by DDCAS-B and DDCAS-P mostly overlap each other with negligible visual differences. Furthermore, it can be shown that both DDCAS-B and DDCAS-P achieve at least a 12 dB sidelobe decay in Scenarios I–IV.

5.2.2. Multiple realization performance

For performance generalization, we have tested all algorithms with 100 random initializations and the reported results are averaged over the 100 instances. The numerical results of the converged objective and individual losses are displayed in Table 2. The best results are



(a) Symmetric beampattern.

(b) Asymmetric beampattern.

(c) Asymmetric beampattern with close-to-target communication requirements.

(d) Asymmetric beampattern with Gaussian windows.

Fig. 3. Beampattern comparison of different algorithms under various scenarios.

Table 2
Converged objective and individual losses of different algorithms, averaged over 100 instances.

(a) Symmetric beampattern.

Algorithm	Objective	Beampattern loss	Communication loss
SDP + AltMin	0.1900	0.07203	1.180
$\boldsymbol{\Theta}$-LBFGS	0.09165	0.09163	2.540×10^{-4}
Minimax + AltMin	0.5353	0.3846	1.507
DDCAS-B (proposed)	0.07610	0.07610	2.955×10^{-5}
DDCAS-P (proposed)	0.07544	0.07544	3.573×10^{-5}

(b) Asymmetric beampattern.

Algorithm	Objective	Beampattern loss	Communication loss
SDP + AltMin	0.3835	0.1168	2.667
$\boldsymbol{\Theta}$-LBFGS	0.1261	0.1261	1.792×10^{-4}
Minimax + AltMin	1.270	0.7544	5.151
DDCAS-B (proposed)	0.1207	0.1207	3.630×10^{-5}
DDCAS-P (proposed)	0.1206	0.1206	4.813×10^{-5}

(c) Asymmetric beampattern with close-to-target communication requirements.

Algorithm	Objective	Beampattern loss	Communication loss
SDP + AltMin	0.2285	0.1196	1.089
$\boldsymbol{\Theta}$-LBFGS	0.1387	0.1387	1.899×10^{-4}
Minimax + AltMin	1.793	1.349	4.438
DDCAS-B (proposed)	0.1244	0.1244	4.078×10^{-6}
DDCAS-P (proposed)	0.1240	0.1240	3.880×10^{-6}

(d) Asymmetric beampattern with Gaussian windows.

Algorithm	Objective	Beampattern loss	Communication loss
SDP + AltMin	0.5402	0.02440	5.158
$\boldsymbol{\Theta}$-LBFGS	0.02426	0.02426	2.992×10^{-5}
Minimax + AltMin	0.7032	0.1856	5.176
DDCAS-B (proposed)	0.01895	0.01895	1.714×10^{-5}
DDCAS-P (proposed)	0.01884	0.01884	2.079×10^{-5}

highlighted in red. The objective computation follows problem (1). In terms of objective value, DDCAS-P achieves the lowest level in all the studied scenarios, narrowly beating DDCAS-B. $\boldsymbol{\Theta}$-LBFGS ranks third because it shares almost the same objective as the proposed algorithms but terminates early in a far-from-optimal region. "SDP + AltMin" and "Minimax + AltMin" rank last due to the two-stage optimization scheme; the "AltMin" stage kills communication properties and even causes beampattern mismatch. Next we investigate into individual losses. When the beampattern loss is concerned, "SDP + AltMin" does the best job in Scenarios I, II, III, and V while DDCAS-P ranks top in Scenario IV. The beampattern matching performance of "Minimax + AltMin" is not so satisfactory. As for the communication losses, DDCAS-B and DDCAS-P outperform the other algorithms by at least one order of magnitude on most occasions in Scenarios I - III and their superiority is still maintained in Scenario IV; "SDP + AltMin" and "Minimax + AltMin" fail to meet communication requirements.

To illustrate the dual functional capability on the communication side, we adopt a practical evaluation metric named Mean Relative Symbol Deviation (MRSD), which is calculated as MRSD $= \frac{1}{L}\sum_{l=1}^{L}\frac{\left|\Delta_{\text{design},l}-\Delta_{\text{standard},l}\right|}{\Delta_{\text{standard},l}}$. This metric is driven by the low-power communication requirement so that the communication signal can be perfectly hidden in the background. For transmission security, only the symbols whose power levels are close to standard can be successfully decoded. Thus, the MRSD metric should be as small as possible. The MRSD comparison results are presented in Table 3. Under the four studied scenarios, the proposed algorithms (either DDCAS-B or DDCAS-P) achieve the lowest MRSD. The waveforms generated by benchmarks "SDP + AltMin" and "Minimax + AltMin" cannot realize the communication purpose due to extremely large MRSD values. Benchmark $\boldsymbol{\Theta}$-LBFGS is slightly inferior to the proposed algorithms on MRSD because of early termination in a suboptimal region. Quantitatively, DDCAS-B and DDCAS-P reduce the MRSD value of $\boldsymbol{\Theta}$-LBFGS by 51% and 46%,

Table 3
Communication-oriented metric MRSD comparison under the four studied scenarios, averaged over 100 instances.

	I	II	III	IV
SDP + AltMin	19.94	65.81	12.06	624.4
$\boldsymbol{\Theta}$-LBFGS	0.02950	0.01533	0.01497	8.451×10^{-3}
Minimax + AltMin	35.51	526.8	237.26	527.6
DDCAS-B (proposed)	0.01023	9.128×10^{-3}	3.777×10^{-3}	6.496×10^{-3}
DDCAS-P (proposed)	0.01137	0.01086	3.374×10^{-3}	7.172×10^{-3}

respectively. Hence, dual functionality on the communication side is fulfilled.

To conclude, the above simulations have shown that the proposed algorithms suffice to generate qualified waveforms in different engineering scenarios. If the beampattern shape is the only concern, either DDCAS-B or DDCAS-P works. When the optimality of objective or the beampattern loss is prioritized, DDCAS-P is recommended. When computational efficiency is preferred, one may want to choose DDCAS-B.

6. Conclusion

In this paper, we have studied the transmit waveform optimization problem in radar-centric MIMO DFRC systems. We have considered a penalty-based formulation which addresses the concerns of transmit beampattern matching and communication power level fitting. A constant modulus constraint has been imposed on the waveform for hardware concerns. We have rewritten the original problem equivalently into an unconstrained one via a change of variable. Thereafter, we have proposed a novel and provably convergent algorithm named DDCAS which combines the descent directions provided by gradient and Amsgrad. In addition, we have developed the shrinkage-expansion line search method which allows a customized initial step and enables a larger step size than initialized, as opposed to the traditional Armijo step size rule. Algorithm DDCAS contains two weight combination strategies: binary and probabilistic combination. The binary combination strategy has shown provable stationarity convergence and monotonic property in objective. The probabilistic combination strategy has experimentally achieved an even lower converged objective than its binary counterpart. Furthermore, a convergence rate of $\mathcal{O}\left(\frac{\log T}{\sqrt{T}}\right)$ has been verified under certain assumptions. Numerical simulations have revealed that algorithm DDCAS manages to synthesize qualified waveforms which achieve radar and communication purposes. We have also specified the occasions for which the combination strategies are well suited.

CRediT authorship contribution statement

Licheng Zhao: Writing – original draft, Methodology, Conceptualization. **Rui Zhou:** Writing – review & editing. **Qingjiang Shi:** Supervision.

Data availability

No data was used for the research described in the article.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.sigpro.2024.109635.

References

[1] F. Liu, C. Masouros, A.P. Petropulu, H. Griffiths, L. Hanzo, Joint radar and communication design: Applications, state-of-the-art, and the road ahead, IEEE Trans. Commun. 68 (2020) 3834–3862.

[2] D. Ma, N. Shlezinger, T. Huang, Y. Liu, Y.C. Eldar, Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies, IEEE Signal Process. Mag. 37 (2020) 85–97.

[3] C. Yang, H.-R. Shao, Wifi-based indoor positioning, IEEE Commun. Mag. 53 (2015) 150–157.

[4] P.K. Hughes, J.Y. Choe, Overview of advanced multifunction rf system (amrfs), in: Proceedings 2000 IEEE International Conference on Phased Array Systems and Technology (Cat. No. 00TH8510), IEEE, 2000, pp. 21–24.

[5] Y. Zeng, R. Zhang, T.J. Lim, Wireless communications with unmanned aerial vehicles: Opportunities and challenges, IEEE Commun. Mag. 54 (2016) 36–42.

[6] S.D. Blunt, J.G. Metcalf, C.R. Biggs, E. Perrins, Performance characteristics and metrics for intra-pulse radar-embedded communication, IEEE J. Sel. Areas Commun. 29 (2011) 2057–2066.

[7] A. Hassanien, M.G. Amin, Y.D. Zhang, F. Ahmad, Signaling strategies for dual-function radar communications: An overview, IEEE Aerosp. Electron. Syst. Mag. 31 (2016) 36–45.

[8] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, Y.C. Eldar, Joint transmit beamforming for multiuser MIMO communications and MIMO radar, 68 (2020) 3929–3944.

[9] A. Hassanien, M.G. Amin, E. Aboutanios, B. Himed, Dual-function radar communication systems: A solution to the spectrum congestion problem, IEEE Signal Process. Mag. 36 (2019) 115–126.

[10] P. Kumari, J. Choi, N. González-Prelcic, R.W. Heath, IEEE 802.11 ad-based radar: An approach to joint vehicular communication-radar system, IEEE Trans. Veh. Technol. 67 (2017) 3012–3027.

[11] J. Moghaddasi, K. Wu, Multifunctional transceiver for future radar sensing and radio communicating data-fusion platform, IEEE Access 4 (2016) 818–838.

[12] C. Sahin, J. Jakabosky, P.M. McCormick, J.G. Metcalf, S.D. Blunt, A novel approach for embedding communication symbols into physical radar waveforms, in: 2017 IEEE Radar Conference (RadarConf), IEEE, 2017, pp. 1498–1503.

[13] P.M. McCormick, S.D. Blunt, J.G. Metcalf, Simultaneous radar and communications emissions from a common aperture, part i: Theory, in: 2017 IEEE Radar Conference (RadarConf), IEEE, 2017.

[14] T. Huang, N. Shlezinger, X. Xu, Y. Liu, Y.C. Eldar, Majorcom: A dual-function radar communication system using index modulation, IEEE Trans. Signal Process. 68 (2020) 3423–3438.

[15] A. Hassanien, M.G. Amin, Y.D. Zhang, F. Ahmad, Dual-function radar-communications: Information embedding using sidelobe control and waveform diversity, IEEE Trans. Signal Process. 64 (2015) 2168–2181.

[16] A. Ahmed, Y.D. Zhang, B. Himed, Multi-user dual-function radar-communications exploiting sidelobe control and waveform diversity, in: 2018 IEEE Radar Conference (RadarConf18), IEEE, 2018a, pp. 0698–0702.

[17] A. Ahmed, Y.D. Zhang, Y. Gu, Dual-function radar-communications using QAM-based sidelobe modulation, Digit. Signal Process. 82 (2018b) 166–174.

[18] A. Ahmed, Y. Gu, D. Silage, Y.D. Zhang, Power-efficient multi-user dual-function radar-communications, in: 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications, SPAWC, IEEE, 2018c, pp. 1–5.

[19] X. Wang, A. Hassanien, M.G. Amin, Sparse transmit array design for dual-function radar communications by antenna selection, Digit. Signal Process. 83 (2018) 223–234.

[20] C. Wen, Y. Huang, T.N. Davidson, Efficient transceiver design for MIMO dual-function radar-communication systems, IEEE Trans. Signal Process. 71 (2023a) 1786–1801.

[21] C. Wen, Y. Huang, L. Zheng, W. Liu, T.N. Davidson, Transmit waveform design for dual-function radar-communication systems via hybrid linear-nonlinear precoding, IEEE Trans. Signal Process. 71 (2023b) 2130–2145.

[22] F. Liu, C. Masouros, A. Li, H. Sun, L. Hanzo, MU-MIMO communications with MIMO radar: From co-existence to joint transmission, IEEE Trans. Wireless Commun. 17 (2018a) 2755–2770.

[23] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, A. Petropulu, Toward dual-functional radar-communication systems: Optimal waveform design, IEEE Trans. Signal Process. 66 (2018b) 4264–4279.

[24] R. Liu, M. Li, Q. Liu, A.L. Swindlehurst, Dual-functional radar-communication waveform design: A symbol-level precoding approach, IEEE J. Sel. Topics Signal Process. 15 (2021) 1316–1331.

[25] Z. Liu, H. Zhang, T. Huang, F. Xu, Y.C. Eldar, Hybrid RIS-assisted MIMO dual-function radar-communication system, IEEE Trans. Signal Process. 72 (2024) 1650–1665.

[26] X. Jin, T. Lv, W. Ni, Z. Lin, Q. Zhu, E. Hossain, H.V. Poor, A reconfigurable subarray architecture and hybrid beamforming for millimeter-wave dual-function-radar-communication systems, IEEE Trans. Wireless Commun. (2024).

[27] P. Stoica, J. Li, Y. Xie, On probing signal design for MIMO radar, IEEE Trans. Signal Process. 55 (2007) 4151–4161.

[28] C. Wen, Y. Huang, J. Peng, J. Wu, G. Zheng, Y. Zhang, Slow-time FDA-MIMO technique with application to STAP radar, IEEE Trans. Aerosp. Electron. Syst. 58 (2021) 74–95.
[29] Y.-C. Wang, X. Wang, H. Liu, Z.-Q. Luo, On the design of constant modulus probing signals for MIMO radar, IEEE Trans. Signal Process. 60 (2012) 4432–4438.
[30] Z. Cheng, Z. He, S. Zhang, J. Li, Constant modulus waveform design for MIMO radar transmit beampattern, IEEE Trans. Signal Process. 65 (2017) 4912–4923.
[31] J. Song, P. Babu, D.P. Palomar, Sequence design to minimize the weighted integrated and peak sidelobe levels, IEEE Trans. Signal Process. 64 (2016) 2051–2064.
[32] J.F. Sturm, Using sedumi 1.02 . a matlab toolbox for optimization over symmetric cones, Optim. Methods Softw. 11 (1999) 625–653.
[33] L. Zhao, J. Song, P. Babu, D.P. Palomar, A unified framework for low autocorrelation sequence design via majorization–minimization, IEEE Trans. Signal Process. 65 (2016) 438–453.
[34] L. Zhao, D.P. Palomar, Maximin joint optimization of transmitting code and receiving filter in radar and communications, IEEE Trans. Signal Process. 65 (2016) 850–863.
[35] M. Hong, Z.-Q. Luo, M. Razaviyayn, Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems, SIAM J. Optim. 26 (2016) 337–364.
[36] Y. Wang, W. Yin, J. Zeng, Global convergence of admm in nonconvex nonsmooth optimization, J. Sci. Comput. 78 (2019) 29–63.
[37] R. Varadhan, C. Roland, Simple and globally convergent methods for accelerating the convergence of any EM algorithm, scandinavian, J. Stat. 35 (2008) 335–353.
[38] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: International Conference on Learning Representations, 2015.
[39] S.J. Reddi, S. Kale, S. Kumar, On the convergence of adam and beyond, in: International Conference on Learning Representations, 2018.
[40] Y.E. Nesterov, A method for solving the convex programming problem with convergence rate O(1/kˆ2), in: Soviet Mathematics Doklady, vol. 269, 1983, pp. 543–547.
[41] C. Jin, R. Ge, P. Netrapalli, S.M. Kakade, M.I. Jordan, How to escape saddle points efficiently, in: International Conference on Machine Learning, PMLR, 2017, pp. 1724–1732.
[42] S.S. Du, C. Jin, J.D. Lee, M.I. Jordan, A. Singh, B. Poczos, Gradient descent can take exponential time to escape saddle points, Adv. Neural Inf. Process. Syst. vol. 30 (2017).
[43] N.S. Keskar, R. Socher, Improving generalization performance by switching from adam to sgd, 2017, arXiv preprint arXiv:1712.07628.
[44] C. Jin, P. Netrapalli, M.I. Jordan, Accelerated gradient descent escapes saddle points faster than gradient descent, in: Conference on Learning Theory, PMLR, 2018, pp. 1042–1085.
[45] M. Staib, S. Reddi, S. Kale, S. Kumar, S. Sra, Escaping saddle points with adaptive gradient methods, in: International Conference on Machine Learning, PMLR, 2019, pp. 5956–5965.
[46] C. Zhang, T. Li, Escape saddle points by a simple gradient-descent based algorithm, Adv. Neural Inf. Process. Syst..
[47] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, 2009.
[48] H. He, J. Li, P. Stoica, Waveform Design for Active Sensing Systems: A Computational Approach, Cambridge University Press, 2012.
[49] J. Duchi, E. Hazan, Y. Singer, Adaptive subgradient methods for online learning and stochastic optimization, J. Mach. Learn. Res. 12 (2011).
[50] M. Fazel, H. Hindi, S.P. Boyd, Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices, in: Proceedings of the 2003 American Control Conference, vol. 3, IEEE, 2003, 2003, pp. 2156–2162.
[51] Y. Nesterov, Smooth minimization of non-smooth functions, Math. Program. 103 (2005) 127–152.
[52] X. Chen, S. Liu, R. Sun, M. Hong, On the convergence of a class of adam-type algorithms for non-convex optimization, in: International Conference on Learning Representations, 2019.
[53] P. Stoica, J. Li, X. Zhu, Waveform synthesis for diversity-based transmit beampattern design, IEEE Trans. Signal Process. 56 (2008) 2593–2598.