Learning-Enhanced Riemannian Gradient Descent Method for Transmit-Receive Joint Design Towards ISRJ Suppression

Xiangfeng Qiu
Weidong Jiang
Yongxiang Liu
National University of Defense Technology, Changsha, China

Symeon Chatzinotas, Fellow, IEEE
National Centre for Scientific Research "Demokritos", Athens, Greece

Fulvio Gini, Fellow, IEEE
Maria Sabrina Greco, Fellow, IEEE
University of Pisa, 56122 Pisa, Italy

***Abstract*— The interrupted sampling repeater jamming (ISRJ) can create false targets that obscure real targets, leading to radar target detection failures. This study investigates the ISRJ countermeasure in multiple-input multiple-output (MIMO) radar through transmit-receive joint design. We initially model the transmit-receive design problem as a jointly constrained optimization problem, aiming to minimize the waveform sidelobes, ISRJ energy, and mutual interference among various waveform-filter pairs. To address the difficulties posed by non-convex constraints, we transform the original constrained problem ~~in~~ Euclidean space into an unconstrained one in Riemannian manifold space. To simultaneously and adaptively update the transmit waveforms and receive filters, we propose a learning-enhanced Riemannian gradient descent (LE-RGD) method, which unfolds the classical Riemannian gradient descent (RGD) method into layers of a neural network. The LE-RGD algorithm directly optimizes transmit waveforms and receive filters through implicit gradient descent iterations, where the optimization strategy is dynamically and adaptively determined by a parameterized network at each iteration. Furthermore, the LE-RGD network is randomly initialized at each problem instance and updated iteratively, facilitating its application in diverse jamming environments without the need for labeled training data. Numerical experiments conclusively show that the LE-RGD method can effectively design transmit waveforms and receive filters with high performance in terms of pulse compression and ISRJ suppression.**

***Index Terms*— Implicit gradient descent, jamming countermeasure, receiving filter, unsupervised learning, waveform design**

Manuscript received XXXXX 00, 0000; revised XXXXX 00, 0000; accepted XXXXX 00, 0000.
The work of Xiangfeng Qiu has been supported by the China Scholarship Council and the National Natural Science Foundation of China under Grants 62022091 and 61921001. The work of F. Gini and M.S. Greco has been partially supported by the Italian Ministry of Education and Research (MUR) in the framework of the FoReLab project (Departments of Excellence).

Authors' address: Xiangfeng Qiu, Weidong Jiang, and Yongxiang Liu are with the College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China, E-mail: (qxf1981993100@163.com; jwd2232@vip.163.com; lyx_bible@sina.com.); Symeon Chatzinotas is with the Institute of Informatics and Telecommunications, National Centre for Scientific Research "Demokritos", Ag. Paraskevi 15310, Athens, Greece, E-mail: (schatzin@ieee.org). Fulvio Gini and Maria Sabrina Greco are with the Department of Information Engineering, University of Pisa, 56122, Pisa, Italy, E-mail: (fulvio.gini@unipi.it; maria.greco@unipi.it.). *(Corresponding author: Xiangfeng Qiu)*

0018-9251 © 2022 IEEE

I. Introduction

A. Works about ISRJ countermeasure

Interrupted sampling repeater jamming (ISRJ), a prevalent coherent jamming technique in pulse-Doppler radar systems, operates by intercepting the radar signal and transmitting it back towards the intended victim radar [1]–[5]. The key characteristic of ISRJ lies in its high coherence with the original transmit waveform, which subsequently leads to the formation of deceptive targets after pulse compression (PC) at the receiver. Thus, this poses a significant challenge to real target detection under interference conditions. Consequently, the investigation and development of ISRJ suppression techniques are of paramount importance in enhancing the radar's functionality and resilience in the face of such interference.

The efficient ISRJ mitigation approaches rely on the ability to discriminate between target and jamming signals, either within a specific domain or across multiple domains. Modern radar systems, empowered by advanced waveform generators, have harnessed sophisticated waveforms to enhance detection capabilities [6]–[9]. This technological improvement facilitates early-stage differentiation at the transmitter, outperforming conventional echo processing methods in countering ISRJ. As a result, extensive research has been dedicated to waveform design strategies, highlighting their significance in anti-jamming efforts [10]–[12].

While waveform design holds promise, a joint approach involving waveform-filter design offers additional avenues for improving interference suppression [13]–[16]. These methodologies typically involve minimizing the ISRJ signal's total energy and the sidelobe energy of the waveform processed by a mismatched filter (MMF), which aims to find a waveform-filter pair that ensures orthogonality between the ISRJ component and the transmit waveform after PC processing. However, the existing literature predominantly addresses single-channel radar systems. The extension to MIMO systems introduces a new challenge, as orthogonality requirements extend not only to the ISRJ component and waveforms but also to the waveforms of distinct channels. As a result, the development of an anti-ISRJ transmit-receive design for MIMO systems remains an open research topic.

B. Applications of deep learning methods

Recent advancements in neural network research, particularly the refinement of network architectures and optimization techniques, have propelled deep learning (DL) to unparalleled success across a myriad of applications [17]. Given the inherent complexity of non-convex optimization challenges in radar waveform design, neural networks, as versatile nonlinear systems, offer a promising approach to designing the desired waveform. A plethora of DL-driven radar waveform strategies have emerged, leveraging various architectures such as fully connected layers [18]–[20] and residual neural networks [21]–[23]. These methods typically employ a standardized neural network framework to establish a non-linear transformation from the initial waveform to an optimized one. They then iteratively adjust layer parameters through a network optimizer, minimizing a predefined loss function until convergence is achieved. Despite their demonstrated progress, these DL-based methodologies encounter a common challenge: their non-transparent, or "black-box" character [24] generally renders them less interpretable, presenting a challenge that continues to be a research hotspot.

Algorithm unrolling is a technique that enhances the explainability of DL models by transforming iterative algorithms into a sequence of custom network layers. This approach allows the trained model to be more transparent, as it mirrors the structure of known algorithms. In the context of radar waveform design, Metwaly et al. [25] employ unrolling for beampattern design using the Projection-descent-reduction (PDR) algorithm as a basis. However, this method faces a challenge in the need for a large number of training samples generated by the PDR process, which can limit the network's performance if the training data is insufficient. Zhang et al. [26], [27] introduce the alternative solution, by proposing the model-based iterative deep unfolding network. This approach unrolls the gradient descent algorithm on the manifold space, enabling unsupervised learning for parameter updates in each layer. While it improves explainability, the proposed network has a limitation in that only the step size of the descent is adjustable among the layer parameters, which restricts the overall network performance. To summarize, algorithm unrolling is a powerful technique that enhances model interpretability by mimicking iterative processes, but it can be hindered by the reliance on extensive training data or fixed parameter settings. Moreover, the aforementioned works do not involve the joint optimization of the transmit-receive, which is an important way to improve radar capabilities.

C. Proposed approach and major contributions

In this article, the ISRJ suppression approach based on the transmit-receive simultaneous design is investigated to improve the application robustness of the MIMO radar in an interference environment. Specifically, we first establish a transmit-receive design model by minimizing the ISRJ signal's total energy and the sidelobe energy of the waveforms processed by the MMF. In addition, the orthogonality of waveforms between different channels is also considered to reduce the mutual interference between channels. To solve the optimization challenges caused by the non-convex constraints of the waveforms and filters, a learning-enhanced Riemann gradient descent (LE-RGD) algorithm is further proposed to ~~search for~~ the optimal waveforms and filters directly and simultaneously. The main contributions of our work are summarized as follows:

1. The transmit-receive joint design model for ISRJ suppression in MIMO radar. In contrast to conventional strategies designed for single-channel radar against ISRJ, this study innovatively develops a new technique in the realm of MIMO radar systems. The central focus of our proposed methodology is to minimize the sidelobes realized by the waveform and filter from the same channel, while concurrently minimizing the cross-correlation levels between the ISRJ component and the filter. To address the challenges of channel interdependence and mutual interference minimization, our approach integrates the concept of response minimization for waveform-filter pairs associated with distinct channels, thereby enhancing the overall system resilience in the jamming environment with the presence of ISRJ.

2. The LE-RGD algorithm for the transmit-receive adaptive optimization. To address the complexities derived from non-convex constraints imposed on waveforms and filters, our approach initiates by reformulating the initial joint design model into an unconstrained one within the realm of Riemannian geometry. This transformation allows for a more efficient search for the optimal solution in manifold space. To realize the learnable optimization of the waveforms and filters, we further propose a novel learning-enhanced implicit gradient descent strategy, which distinguishes itself from conventional alternating minimization (AM) techniques by enabling concurrent transmit-receive design. In addition, by integrating the model-based gradient descent algorithm into a tailored neural network architecture, the proposed method retains the mathematical foundation of the gradient-based method while enhancing its optimization capabilities and adaptabilities.

D. Outlines and notations

The subsequent sections of this study are organized as follows: Section II presents a thorough introduction to the mathematical foundation for the transmit-receive design problem. Section III is dedicated to the in-depth exposition of our proposed LE-RGD algorithm, detailing its application process and analytical underpinnings. Section IV serves as a numerical demonstration, where we showcase the ISRJ suppression capabilities of our proposed algorithm through a series of simulated experiments, employing a variety of experimental parameters to illustrate its effectiveness. Lastly, Section V concludes the paper by summarizing our findings and discussing potential avenues for future research.

In this paper, matrices are denoted by boldface uppercase letters, while vectors are denoted by boldface lowercase letters. We use $\mathrm{Tr}\left(\cdot\right)$, $(\cdot)^H$, $(\cdot)^*$, $()^T$ to denote the trace, conjugate transpose, complex conjugate, and transpose, respectively. $\mathrm{diag}(\cdot)$ denotes a diagonal matrix with diagonal entries given by a vector. The Hadamard product and the modulus are denoted by $\odot$ and $|\cdot|$. $\varnothing$ means element-wise division. $\mathrm{Re}\left(\cdot\right)$ extracts the real part, and $\left\|\cdot\right\|_F$ denotes the F-norm of a matrix. $\mathrm{Sum}\left(\cdot\right)$ denotes summing the matrix by columns, and $\mathrm{Rep}\left(\cdot, N\right)$ means repeating the column vectors by rows N times to obtain a two-dimensional matrix.

II. Signal model

Consider a MIMO radar system with M independent channels, each transmits a complex-valued baseband waveform of length N. We focus on the scenario where the MMF shares a similar signal configuration with the waveform. Consequently, the matrix representation of the transmit waveform and the receive MMF can be mathematically formalized as follows:

$$\begin{aligned} \boldsymbol{S} &= \left[\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M\right]_{N\times M} \\ \boldsymbol{W} &= \left[\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_M\right]_{N\times M}, \end{aligned} \tag{1}$$

with each column of $\boldsymbol{S}$ representing a distinct waveform and rows corresponding to individual channels. The matrix $\boldsymbol{W}$, on the other hand, encapsulates the filter coefficients applied to the received signals, serving as a crucial component in optimizing the overall system performance. Defining a shift matrix

$$\boldsymbol{J}_r(p,q) = \begin{cases} 1, & \text{if} \quad p-q=r \\ 0, & \text{if} \quad p-q\neq r \end{cases} \quad p,q\in\{1,...,N\}, \tag{2}$$

then the PC response values employing the MMF at the specific time-delay r can be expressed as

$$\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S} = \begin{bmatrix} \boldsymbol{w}_1^H\boldsymbol{J}_r\boldsymbol{s}_1 & \cdots & \boldsymbol{w}_1^H\boldsymbol{J}_r\boldsymbol{s}_M \\ \vdots & \ddots & \vdots \\ \boldsymbol{w}_M^H\boldsymbol{J}_r\boldsymbol{s}_1 & \cdots & \boldsymbol{w}_M^H\boldsymbol{J}_r\boldsymbol{s}_M \end{bmatrix}. \tag{3}$$

Similarly, the PC response value associated with the ISRJ signal and the MMF can be given as

$$\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}_J = \begin{bmatrix} \boldsymbol{w}_1^H\boldsymbol{J}_r\boldsymbol{s}_1^J & \cdots & \boldsymbol{w}_1^H\boldsymbol{J}_r\boldsymbol{s}_M^J \\ \vdots & \ddots & \vdots \\ \boldsymbol{w}_M^H\boldsymbol{J}_r\boldsymbol{s}_1^J & \cdots & \boldsymbol{w}_M^H\boldsymbol{J}_r\boldsymbol{s}_M^J \end{bmatrix}, \tag{4}$$

where $\boldsymbol{s}_m^J = \boldsymbol{p}\odot\boldsymbol{s}_m$, $\boldsymbol{S}_J = \mathrm{diag}\left(\boldsymbol{p}\right)\boldsymbol{S}$. The vector $\boldsymbol{p}$ denotes the discrete form of the interrupted sampling function [15], which contains all of the sampling and repeater information about the ISRJ jammer. Here we assume that the sampling function can be acquired through cognitive approaches [6], [7].

With equations (3) and (4), our established objective function used for the transmit-receive design includes the following four parts:

- To suppress the mutual interference among different channels and reduce the PC sidelobes in each channel, the following metric should be minimized

$$f_1\left(\boldsymbol{S},\boldsymbol{W}\right) = \sum_{r\in\mathcal{R}\backslash 0}\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right\|_F^2 + \left\|\left(\boldsymbol{W}^H\boldsymbol{S}\right)\odot\boldsymbol{A}\right\|_F^2, \tag{5}$$

where $\mathcal{R}\in\{1-N,\cdots,N-1\}$ is the range bins of interest, $\boldsymbol{A}=\mathbf{1}-\boldsymbol{I}$, with $\mathbf{1}$ is a matrix of dimension $M\times M$ with all elements set to 1 and $\boldsymbol{I}$ is a unit diagonal matrix with the same size of $\mathbf{1}$.

- To control the signal-to-noise-ratio (SNR) loss caused by the MMF processing, we minimize the following component to force the PC peak values to approach the preset values

$$f_2\left(\boldsymbol{S},\boldsymbol{W}\right) = \left\|\left(\boldsymbol{W}^H\boldsymbol{S}\right)\odot\boldsymbol{I}-\mu_1\boldsymbol{I}\right\|_F^2, \tag{6}$$

where μ_1 is the desired peak value after the PC process.

- To suppress the ISRJ signal's response levels, we minimize the overall energy of the ISRJ component across all the range bins after MMF processing. Thus, we have

$$f_3\left(\boldsymbol{S},\boldsymbol{W}\right) = \sum_{r\in\mathcal{R}}\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}_J\right\|_F^2. \tag{7}$$

- To control the peak values associated with the ISRJ after the PC processing, we minimize the following function component to force the peaks to approach the predefined small values

$$f_4\left(\boldsymbol{S},\boldsymbol{W}\right) = \left\|\left(\boldsymbol{W}^H\boldsymbol{S}_J\right)\odot\boldsymbol{I}-\mu_2\boldsymbol{I}\right\|_F^2, \tag{8}$$

where μ_2 denotes the desired peak value of the MMF result of ISRJ.

In practical applications, the transmit waveforms are required to avoid amplitude fluctuation, which ~~resulting~~ the constant modulus (CM) constraint of the waveforms [28]. Additionally, the receive filters need to satisfy the constant energy (CE) constraint to precisely control the SNR loss [29]. Consequently, our established transmit-receive design model for ISRJ suppression can be mathematically represented as follows

$$\begin{aligned} &\min_{\boldsymbol{S},\boldsymbol{W}} f\left(\boldsymbol{S},\boldsymbol{W}\right) = \sum_{k=1}^{4} f_k\left(\boldsymbol{S},\boldsymbol{W}\right) \\ &\;\mathrm{s.t.}\;\; \left|\boldsymbol{s}_m\left(n\right)\right| = 1,\;\; \left\|\boldsymbol{w}_m\right\|_2^2 = 1 \end{aligned} \tag{9}$$

where $m\in\{1,\cdots,M\}$, $n\in\{1,\cdots,N\}$, $\boldsymbol{s}_m\left(n\right)$ denotes the n-th sample point of the transmit waveform in the m-th channel, $\boldsymbol{w}_m$ denotes the receive filter associated with the m-the channel. In practical deployments, it is also necessary to assign weights to individual sub-objective functions, reflecting the specific priorities and requirements of the application. To balance the weights of each sub-function, no weighting is performed on each function component.

Remark: Before formally presenting our solution to the problem (9), we first provide some comments on the formulated problem. The given problem, characterized by a bivariate objective function and non-convex constraints, belongs to the class of NP-hard problems. Conventional approaches adopt an AM strategy, which decomposes the original two-variable problem into iterative sequences of subproblems, each focusing on optimizing one variable while keeping the other constant. However, the non-convex nature of the problem poses a significant challenge, as even if each subproblem is optimally solved at each iteration, the overall solution may not converge to a global optimum. The fundamental limitation lies in its reliance on a greedy, non-adaptive optimization strategy, which can lead to local optima trapping during the subproblem-solving process. To address this limitation, we propose a novel approach that utilizes the advantages of interpretable~~, learning based~~ neural networks. Our approach leverages unsupervised learning to imbue the network with an adaptive and training-free optimization behavior, thereby enhancing its ability to navigate the complex landscape of non-convex optimization.

III. Proposed method

In this section, we first present the AM-based RGD algorithm to solve the established optimization problem (9). Then the proposed LE-RGD method is formally introduced, including its network structure and parameters update approaches.

A. AM-based RGD algorithm

Riemannian manifold optimization has emerged as a prevalent framework in waveform optimization, particularly in applications like [28], [30]. Its essence lies in leveraging the geometric properties of constrained spaces to transform a traditional Euclidean optimization problem into an unconstrained one on a Riemannian manifold space. Under this framework, the optimization problem in (9) can be reformulated equivalently as

$$\min_{\boldsymbol{S}\in\mathcal{S},\boldsymbol{W}\in\mathcal{W}} f(\boldsymbol{S},\boldsymbol{W})=\sum_{k=1}^{4} f_k(\boldsymbol{S},\boldsymbol{W}), \tag{10}$$

where $\mathcal{S}$ and $\mathcal{W}$ denote the complex circle manifold (CCM) and complex oblique manifold (COM), respectively, defined by the constraints on the waveforms and filters. Mathematically, the CCM and COM can be expressed as [31]

$$\begin{aligned}\mathcal{S}&=\{\boldsymbol{S}\in\mathbb{C}^{N\times M}|\boldsymbol{s}_m(n)=1\}\\ \mathcal{W}&=\{\boldsymbol{W}\in\mathbb{C}^{N\times M}|\boldsymbol{w}_m^H\boldsymbol{w}_m=1\}.\end{aligned} \tag{11}$$

To tackle this transformed problem, the AM-based approach is the most popular approach. It sequentially optimizes individual subproblems, treating each variable independently while holding the others constant. Starting from an initial point $\boldsymbol{W}_0\in\mathcal{W}$, the AM-based algorithm at the t-th iteration updates the variables as follows:

$$\begin{aligned}\boldsymbol{S}_t&=\arg\min_{\boldsymbol{S}\in\mathcal{S}} f(\boldsymbol{S},\boldsymbol{W}_{t-1})\\ \boldsymbol{W}_t&=\arg\min_{\boldsymbol{W}\in\mathcal{W}} f(\boldsymbol{S}_t,\boldsymbol{W}).\end{aligned} \tag{12}$$

At the t-th iteration, the algorithm operates by optimizing two separate functions, $f(\boldsymbol{S},\boldsymbol{W}_{t-1})$ and $f(\boldsymbol{S}_t,\boldsymbol{W})$, which respectively correspond to the variables $\boldsymbol{S}$ and $\boldsymbol{W}$. This two-step process facilitates the iterative update of the variables within their respective feasible regions. Each subproblem given in equation (12) can be resolved through the RGD algorithm as [32]

$$\begin{aligned}&\boldsymbol{S}_i=\mathrm{Ret}_{\boldsymbol{S}}\left\{\boldsymbol{S}_{i-1}-\alpha_i\cdot\mathcal{P}_{\boldsymbol{S}}\left[\mathrm{Grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S}_{i-1},\boldsymbol{W}_{t-1}\right)\right]\right\}\\ &\boldsymbol{W}_i=\mathrm{Ret}_{\boldsymbol{W}}\left\{\boldsymbol{W}_{i-1}-\gamma_i\cdot\mathcal{P}_{\boldsymbol{W}}\left[\mathrm{Grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S}_t,\boldsymbol{W}_{i-1}\right)\right]\right\},\end{aligned} \tag{13}$$

where

- $\mathrm{Grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)$ and $\mathrm{Grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)$ denote the gradients in Euclidean space of the waveform and filter, which are given in (14) in the bottom of the next page.

Proof: The derivations of the gradients can be found in the Appendix part.

- $\mathcal{P}_{\boldsymbol{S}}\left(\cdot\right)$ and $\mathcal{P}_{\boldsymbol{W}}\left(\cdot\right)$ are employed to project the Euclidean gradients into their Riemannian counterparts, as depicted in the following equations [33]

$$\begin{aligned}&\mathrm{grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)\\ &=\mathcal{P}_{\boldsymbol{S}}\left(\mathrm{Grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)\right)=\mathrm{Grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)\\ &-\mathrm{Re}\left(\left(\mathrm{Grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)\right)^*\odot\boldsymbol{S}\right)\odot\boldsymbol{S},\end{aligned} \tag{15}$$

and

$$\begin{aligned}&\mathrm{grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)\\ &=\mathcal{P}_{\boldsymbol{W}}\left(\mathrm{Grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)\right)=\mathrm{Grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)\\ &-\mathrm{Rep}\left[\mathrm{Re}\left(\mathrm{Sum}\left(\boldsymbol{W}^*\odot\mathrm{Grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)\right)\right)\right]\odot\boldsymbol{W}.\end{aligned} \tag{16}$$

where $\mathrm{grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S},\boldsymbol{W}\right)$ and $\mathrm{grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S},\boldsymbol{W}\right)$ signify the Riemannian gradients of the objective function concerning the waveform and filter. These gradients, combined with the step sizes α_i and γ_i, serve as the basis for calculating descent vectors and generating new points within the current iteration.

- $\mathrm{Ret}_{\boldsymbol{S}}\left(\cdot\right)$ and $\mathrm{Ret}_{\boldsymbol{W}}\left(\cdot\right)$ are employed to retract the newly obtained points back into the manifold space, as expressed

$$\begin{aligned}&\mathrm{Ret}_{\boldsymbol{S}}\left(\boldsymbol{S}\right)=\boldsymbol{S}\varnothing\left|\boldsymbol{S}\right|\\ &\mathrm{Ret}_{\boldsymbol{W}}\left(\boldsymbol{W}\right)=\boldsymbol{W}\odot\mathrm{Rep}\left[\mathrm{Sum}^{\frac{1}{2}}\left(\left|\boldsymbol{W}\right|^2\right),N\right].\end{aligned} \tag{17}$$

In the iterative procedure delineated by equation (13), the primary aim is to mitigate the overarching challenge by sequentially optimizing its constituent subproblems. However, it's imperative to recognize that the algorithm may not converge to a global optimum or a satisfactory solution. This limitation arises from the reliance on first-order information, often leading to locally greedy optimization, potentially overlooking the optimal outcome. The fundamental issue of overcoming these limitations lies in the static nature of model-based strategies, as encapsulated by the update functions in equation (13). These update functions, initially designed to address individual subproblems, lack the inherent capacity to produce globally optimal solutions for the entire problem. This inherent constraint confines the algorithm's exploration of the solution space and impedes its ability to achieve optimality.

B. LE-RGD algorithm

In this section, we present a learning-driven optimization methodology designed to tackle the simultaneous transmit-receive design, departing from the conventional approach of minimizing each subproblem individually. By integrating DL principles in the update progress, our approach incorporates a more flexible variables update way and seeks to enhance the overall design efficiency and effectiveness.

1. Overall structure

In Fig.1, the implemented network structure of the proposed LE-RGD algorithm is illustrated. As we can observe from this figure, the blue and green sub-network structures defined in parallel are used to update waveforms and filters respectively. In addition, each sub-network is composed of T layers, containing multiple operators that mimic the mathematical operations of the RGD algorithm. To facilitate more flexible update functions, the LE-RGD approach modifies the iterative framework in (12) and replaces the analytical minimization algorithm with neural networks in the form of

$$\begin{aligned} \boldsymbol{S}_t &= \mathrm{Net}^{\boldsymbol{S}}_{\theta_t}\left(\boldsymbol{S}_{t-1}, \boldsymbol{W}_{t-1}\right) \\ \boldsymbol{W}_t &= \mathrm{Net}^{\boldsymbol{W}}_{\rho_t}\left(\boldsymbol{S}_{t-1}, \boldsymbol{W}_{t-1}\right). \end{aligned} \tag{18}$$

After passing through all T layers, one training round concludes, and the initial transmit-receive variables, $\boldsymbol{S}_0$ and $\boldsymbol{W}_0$, are replaced with the updated values, $\boldsymbol{S}_T$ and $\boldsymbol{W}_T$. Then, the loss function is computed using these updated variables, and a backward pass is performed to optimize the trainable network parameters, $\boldsymbol{\theta}_t$ and $\boldsymbol{\rho}_t$. The optimization continues for a maximum number of rounds, K, after which the final results for the transmit-receive variables are obtained from the network parameters and their respective transformations. Since the network structures used to update waveforms and filters are defined separately and can realize forward propagation at the same network layer, the LE-RGD algorithm can realize the simultaneous design of the transmit-receive, avoiding the problem of poor overall performance caused by greedy optimization of a single variable in the AM strategy.

2. Forward Propagation

In Fig. 1, the t-th layer's sub-network structures for waveform and filter updates are shown in the blue-dotted and green-dotted sections, respectively. As we can observe, the algorithm parameters and variables to be updated are input into each layer of the sub-network, and the Euclidean gradient and Riemannian gradient are obtained successively. Then, the Riemannian gradient is input into the non-linear network, and the network output is processed by the retraction operator to finally obtain the updated variables of the layer. In this way, these layers not only maintain the core principles of the model-based RGD algorithm but also integrate network-based components to learn the optimal hyperparameters for descent directions and step sizes.

By introducing the non-linear network structure, the transmit-receive variables can be updated using a parameterized non-linear function, and the parameters of this function are themselves optimized in each round of training. This allows the algorithm to continuously learn and refine the update strategy, ultimately leading to better optimization results compared to a fixed approach. The formulation of the learning-based update strategy can be expressed as

$$\begin{aligned} \boldsymbol{S}_t &= \boldsymbol{S}_{t-1} + \mathrm{FC}^{\boldsymbol{S}}_{\boldsymbol{\theta}_{t-1}}\left(\operatorname{grad} f_{\boldsymbol{W}_{t-1}}\left(\boldsymbol{S}_{t-1}\right)\right) \\ \boldsymbol{W}_t &= \boldsymbol{W}_{t-1} + \mathrm{FC}^{\boldsymbol{W}}_{\boldsymbol{\rho}_{t-1}}\left(\operatorname{grad} f_{\boldsymbol{S}_t}\left(\boldsymbol{W}_{t-1}\right)\right) \end{aligned} \tag{19}$$

where $\mathrm{FC}^{\boldsymbol{S}}_{\boldsymbol{\theta}_t}(\cdot)$ and $\mathrm{FC}^{\boldsymbol{W}}_{\boldsymbol{\rho}_t}(\cdot)$ denote the fully-connected (FC) networks in the t-th layer. It's worth noting that PyTorch does not support the creation of complex-valued neural networks. When dealing with complex gradients, such as $\operatorname{grad} f_{\boldsymbol{S}}(\boldsymbol{S}, \boldsymbol{W})$ and $\operatorname{grad} f_{\boldsymbol{W}}(\boldsymbol{S}, \boldsymbol{W})$, we first extract their real and imaginary parts. These parts are then

$$\begin{aligned} \operatorname{Grad} f_{\boldsymbol{S}}(\boldsymbol{S}, \boldsymbol{W}) = 2 \times \Bigg\{ &\sum_{r \neq 0} \boldsymbol{J}_r^H \boldsymbol{W}\boldsymbol{W}^H \boldsymbol{J}_r \boldsymbol{S} + \boldsymbol{W}\left[\left(\boldsymbol{W}^H\boldsymbol{S}\right) \odot \boldsymbol{A} + \left(\left(\boldsymbol{W}^H\boldsymbol{S}\right) \odot \boldsymbol{I} - \mu_1 \boldsymbol{I}\right) \odot \boldsymbol{I}\right] \\ &+ \sum_{r} \tilde{\boldsymbol{J}}_r^H \boldsymbol{W}\boldsymbol{W}^H \tilde{\boldsymbol{J}}_r \boldsymbol{S} + \boldsymbol{P}^H\boldsymbol{W}\left(\left[\left(\boldsymbol{W}^H\boldsymbol{P}\boldsymbol{S}\right) \odot \boldsymbol{I} - \mu_2 \boldsymbol{I}\right] \odot \boldsymbol{I}\right) \Bigg\} \\ \operatorname{Grad} f_{\boldsymbol{W}}(\boldsymbol{S}, \boldsymbol{W}) = 2 \times \Bigg\{ &\sum_{r \neq 0} \boldsymbol{J}_r \boldsymbol{S}\boldsymbol{S}^H \boldsymbol{J}_r^H \boldsymbol{W} + \boldsymbol{S}\left[\left(\boldsymbol{S}^H\boldsymbol{W}\right) \odot \boldsymbol{A}^H + \left(\left(\boldsymbol{S}^H\boldsymbol{W}\right) \odot \boldsymbol{I} - \mu_1 \boldsymbol{I}\right) \odot \boldsymbol{I}\right] \\ &+ \sum_{r} \tilde{\boldsymbol{J}}_r \boldsymbol{S}\boldsymbol{S}^H \tilde{\boldsymbol{J}}_r^H \boldsymbol{W} + \boldsymbol{P}\boldsymbol{S}\left(\left[\left(\boldsymbol{S}^H\boldsymbol{P}^H\boldsymbol{W}\right) \odot \boldsymbol{I} - \mu_2 \boldsymbol{I}\right] \odot \boldsymbol{I}\right) \Bigg\} \end{aligned} \tag{14}$$

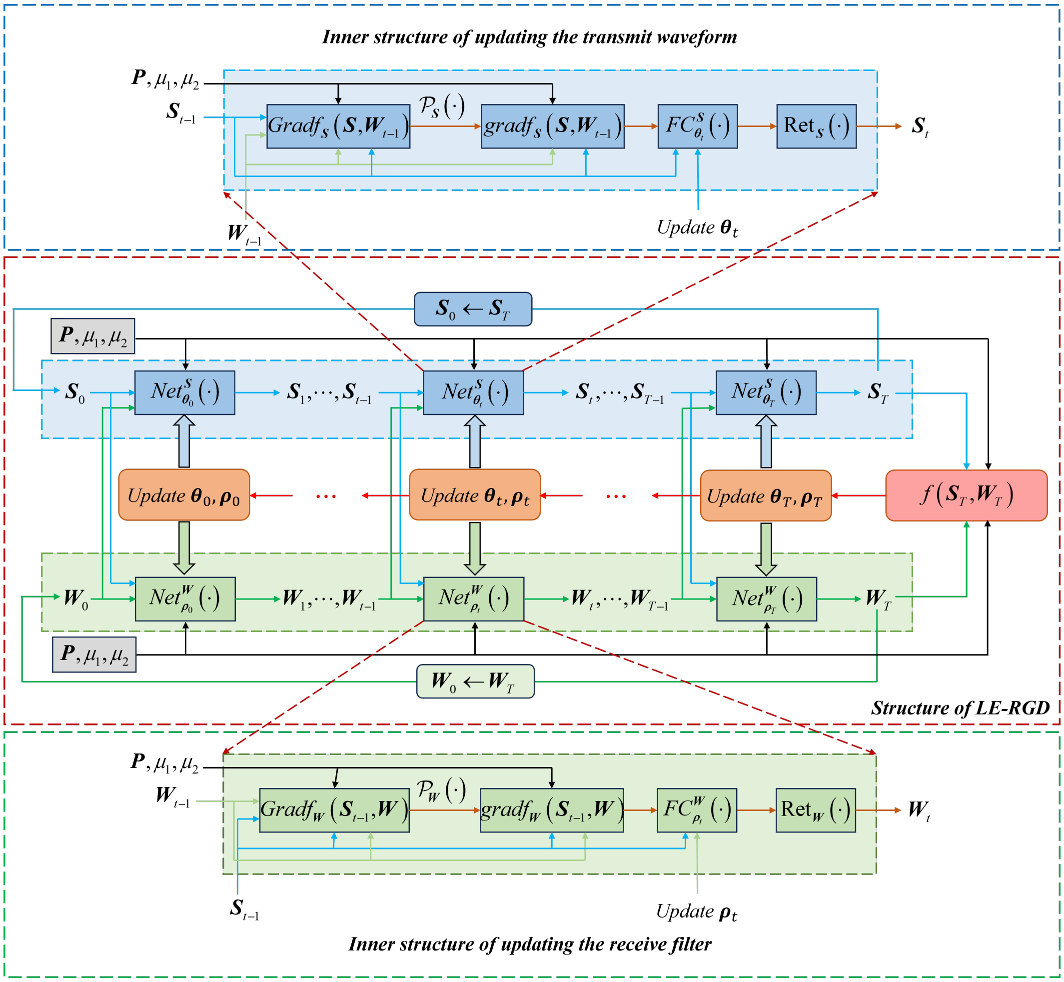


Fig. 1: The overall structure of the proposed LE-RGD algorithm. The red-dotted part is the main skeleton of LE-RGD, the transmit waveform's updating mechanism is symbolically represented by the blue-dotted framework, and the green-dotted line denotes the corresponding network layout utilized for the update of the receive filter. The parameters of the networks for updating the waveforms and filters in the t-th layer are denoted by $\boldsymbol{\theta}_t$ and $\boldsymbol{\rho}_t$. The orange boxes in the middle represent the learnable parameters in the network.

concatenated together and fed into a fully connected (FC) network. At the output end, we reverse the process by combining the real and imaginary outputs in a way that reconstructs the original complex value. This ensures that the output has the same dimensions as the input gradients, allowing us to maintain the complex structure throughout the network.

3. Backward Propagation

In the LE-RGD algorithm, the strategy for variable optimization is controlled by the network parameters, $\boldsymbol{\theta}_t$ and $\boldsymbol{\rho}_t$. These parameters are updated during the backward propagation process, with the loss value serving as the driving force for backward propagation. Unlike traditional machine learning models that require pre-training on a dataset, the network parameters in LE-RGD are updated as the objective function in (9) is minimized. During the backward propagation, the parameters of the FC networks can be updated by the Adam optimizer [34]

$$\begin{aligned}\boldsymbol{\theta}_t &\leftarrow \boldsymbol{\theta}_t + \mu \cdot \mathrm{Adam}\left(\nabla_{\boldsymbol{\theta}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)\right)\\ \boldsymbol{\rho}_t &\leftarrow \boldsymbol{\rho}_t + \mu \cdot \mathrm{Adam}\left(\nabla_{\boldsymbol{\rho}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)\right)\end{aligned} \tag{20}$$

where $\nabla_{\boldsymbol{\theta}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)$ and $\nabla_{\boldsymbol{\rho}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)$ denote the gradients of the loss function concerning the parameters and they can be automatically computed using the built-in automatic differentiation feature. In this way, each update of $\boldsymbol{\theta}_t$ and $\boldsymbol{\rho}_t$ can be seen as a form of *"training while solving"*, where each iteration of the optimization problem acts as a single training sample. This approach allows for a *plug-and-play* implementation, as the algorithm does not necessitate a pre-existing training dataset. The network parameters dynamically adapt to minimize the loss function, resulting in a more efficient and adaptive optimization process for the transmit-receive variables.

4. Algorithm steps and performance analysis

The LE-RGD algorithm, detailed in Algorithm 1, introduces trainable variable update functions, $\mathrm{FC}^{\boldsymbol{S}}_{\boldsymbol{\theta}_t}(\cdot)$

Algorithm 1: The LE-RGD for the joint design of transmit-receive toward ISRJ suppression

Input: $f\left(\boldsymbol{S}, \boldsymbol{W}\right)$, M, N, $\boldsymbol{P}$, $\boldsymbol{I}$, $\boldsymbol{A}$, μ
Initialize: K, T, $\boldsymbol{S}_0 \in \mathcal{S}$, $\boldsymbol{W}_0 \in \mathcal{W}$, $\boldsymbol{\theta}_t$, $\boldsymbol{\rho}_t$
for $k \leftarrow 0, 1, \cdots, K$ **do**
  $t = 0$;
  `/* forward propagation */`
  while $t \leq T$ **do**
    $\Delta\boldsymbol{S} = \mathrm{FC}^{\boldsymbol{S}}_{\boldsymbol{\theta}_t}\left(\mathrm{grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S}_t, \boldsymbol{W}_t\right)\right)$;
    $\Delta\boldsymbol{W} = \mathrm{FC}^{\boldsymbol{W}}_{\boldsymbol{\rho}_t}\left(\mathrm{grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S}_t, \boldsymbol{W}_t\right)\right)$;
    $\boldsymbol{S}_{t+1} = \mathrm{Ret}_{\boldsymbol{S}}\left(\boldsymbol{S}_t + \Delta\boldsymbol{S}\right)$;
    $\boldsymbol{W}_{t+1} = \mathrm{Ret}_{\boldsymbol{W}}\left(\boldsymbol{W}_t + \Delta\boldsymbol{W}\right)$;
    $t \leftarrow t+1$;
  end
  $t = 1$;
  $f(\boldsymbol{S}_T, \boldsymbol{W}_T) = \sum_{i=1}^{4} f_i(\boldsymbol{S}_T, \boldsymbol{W}_T)$;
  `/* backward propagation */`
  while $t \leq T$ **do**
    $\boldsymbol{\theta}_t \leftarrow \boldsymbol{\theta}_t + \mu \cdot \mathrm{Adam}\left(\nabla_{\boldsymbol{\theta}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)\right)$;
    $\boldsymbol{\rho}_t \leftarrow \boldsymbol{\rho}_t + \mu \cdot \mathrm{Adam}\left(\nabla_{\boldsymbol{\rho}_t} f\left(\boldsymbol{S}_T, \boldsymbol{W}_T\right)\right)$;
    $t \leftarrow t+1$;
  end
end

for waveforms and $\mathrm{FC}^{\boldsymbol{W}}_{\boldsymbol{\rho}_t}(\cdot)$ for filters, to replace the fixed update formulas from (13). At each layer t, these neural network-based update rules leverage the gradients, $\mathrm{grad}\, f_{\boldsymbol{S}}\left(\boldsymbol{S}_{t-1}, \boldsymbol{W}_{t-1}\right)$ and $\mathrm{grad}\, f_{\boldsymbol{W}}\left(\boldsymbol{S}_{t-1}, \boldsymbol{W}_{t-1}\right)$, to predict the next step in the optimization of transmit-receive variables. This continuous adaptation allows the algorithm to learn a more effective strategy for solving the joint transmit-receive design problem, inheriting the core RGD-based iterative approach while enhancing its adaptability and learning capabilities.

The computational burden of the LE-RGD algorithm mainly results from the calculations of gradients from (14) and the cost function defined in (9), both of which have a complexity of $\mathcal{O}\left(MN^2\right)$. Note that the operations in the gradients and cost function can be implemented using discrete Fourier transform (DFT). Thus, we can estimate the per-iteration complexity to be around $\mathcal{O}\left(MN\log\left(N\right)\right)$. As depicted in Fig.1, each layer's update generates a point in the manifold space with a decreased cost, assuming appropriate adjustments to the step size and descent direction. To achieve this, the network employs backward propagation to determine the gradients of the loss function concerning all parameters using the chain rule. Subsequently, the network parameters are refined using the Adam optimizer to guarantee a non-increasing loss, thus enabling convergence of the method.

IV. Experiments and analysis

In this section, we evaluate the effectiveness of the proposed LE-RGD algorithm through a series of simulation experiments. The implementation of the LE-RGD algorithm is accomplished using Python version 3.10.11, specifically with the Pytorch library version 2.2.2.

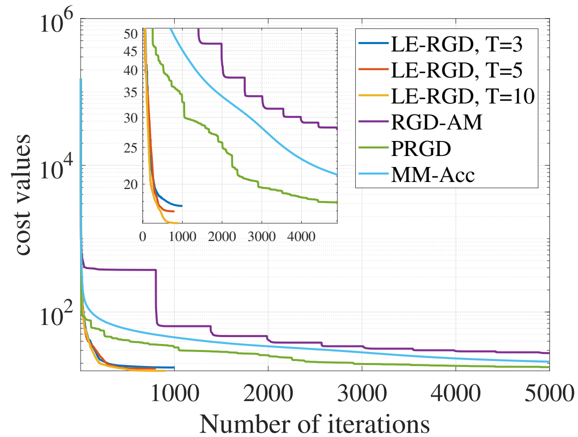
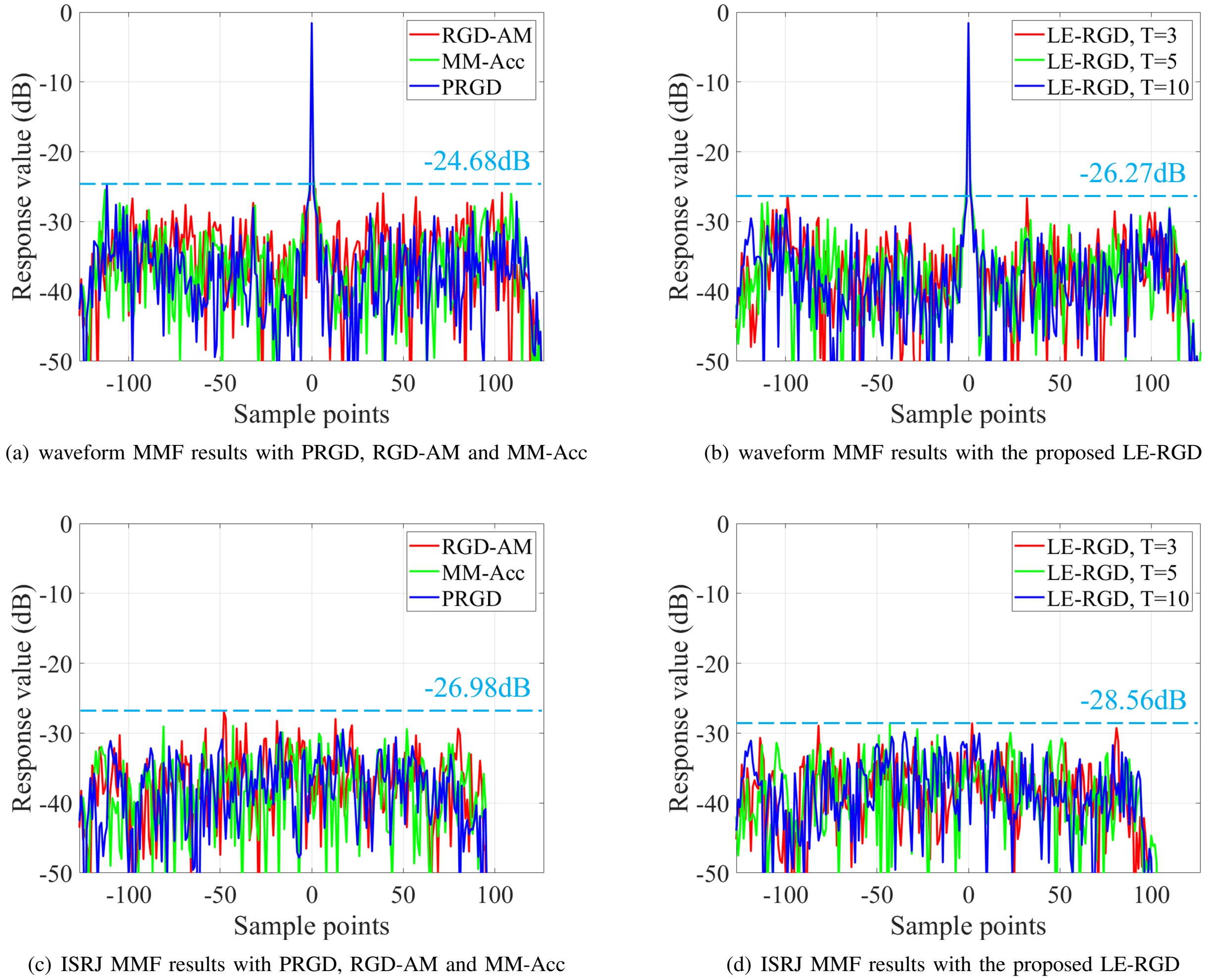


Fig. 2: The cost values versus the number of iterations corresponding to different methods.

A. The joint design in single-channel system

In this part, we first evaluate the convergence performance of the proposed LE-RGD algorithm by comparing it with the AM-based RGD (AM-RGD) method (13), the product manifold-based RGD (PRGD) method [35], and the Majorization-minimization algorithm with acceleration (MM-Acc) [13]. Since the MM-Acc method can only be used in single-channel radar systems, here we set $M = 1$ and $N = 128$ to make the comparison. For the ISRJ parameters, the duty ratio is set to 0.25, and the interrupted sampling period is set to $4/N$. Then the sampling vector $\boldsymbol{p}$ can be determined in these two parameters [15]. For the algorithm parameters, we set $\mu_1 = \sqrt{N} \times 10^{-1.5/20}$, and $\mu_2 = \sqrt{N} \times 10^{-40/20}$. For all of these methods, the maximum number of iterations K is set to 5×10^3. Moreover, the algorithms stop running when the decreased value of the loss function corresponding to two adjacent iterations is less than 1×10^{-4}. The network layers $\mathrm{FC}^{\boldsymbol{S}}_{\boldsymbol{\theta}_t}(\cdot)$ and $\mathrm{FC}^{\boldsymbol{W}}_{\boldsymbol{\rho}_t}(\cdot)$, contain 3 hidden layers with $3 \times N$ units in each. The Adam optimizer with a learning rate of 1×10^{-3} is employed to update the network parameters $\boldsymbol{\theta}_t$ and $\boldsymbol{\rho}_t$ in each layer.

In Fig.2, the iterative progress of different algorithms is illustrated for comparison. To proceed with the design progress, all of the waveforms are initialized with complex-valued matrices with uniformly distributed phases. Since all of the algorithms are initialized with the same starting point, the cost value starts to decrease from an identical value monotonically until a convergence state is reached. For the AM-based algorithms, RGD-AM and MM-Acc, it is hard for them to find a good optimum within the maximum number of iterations as each subproblem in these two methods is solved greedily. The PRGD algorithm first considered a product manifold space composed of the Cartesian product of the waveform



(a) waveform MMF results with PRGD, RGD-AM and MM-Acc

(b) waveform MMF results with the proposed LE-RGD

(c) ISRJ MMF results with PRGD, RGD-AM and MM-Acc

(d) ISRJ MMF results with the proposed LE-RGD

Fig. 3: The realized MMF results of the waveform and ISRJ associated with different methods

and the filter, then converted the two-variable optimization problem into a single-variable optimization problem, and further performed gradient descent on the product manifold space. Since the transmit waveform and receive filter can be updated at the same time, its performance exceeds that of the algorithms based on the AM strategy. However, due to the employment of the fixed update strategy, its convergence performance has the potential to be further improved. In addition, the proposed LE-RGD method reaches a convergence state and smaller cost values within 1000 iterations with the parameters setting of $T = 3, 5, 10$. Moreover, it tends to find a better optimum when the number of network layers is improved.

In Fig.3, the PC results of the waveform and ISRJ corresponding to different design algorithms are illustrated to demonstrate their ISRJ suppression performance. In Fig.3 (a) and (b), the PC results of the optimized waveform-filter pairs are presented. Since they are processed under the MMF theme, all of the algorithms produce different SNR losses around $-1.53\,\mathrm{dB}$. In addition, our proposed LE-RGD generates peak sidelobes (PSL) of $-26.27\,\mathrm{dB}$, $-27.21\,\mathrm{dB}$, and $-28.20\,\mathrm{dB}$ with $T = 3, 5, 10$, surpassing those resulting from RGD-AM, MM-Acc, and PRGD methods. In Fig.3 (c) and (d), we present the PC results of the ISRJ components with the designed receive filters. As we can observe, the ISRJ components cannot obtain processing gain during the process, resulting in no peaks in the PC results, proving the effectiveness of the designed transmit-receive in ISRJ suppression. More precisely, the resulting highest response values are $-28.56\,\mathrm{dB}$, $-28.66\,\mathrm{dB}$, and $-29.81\,\mathrm{dB}$ for the proposed LE-RGD method with $T = 3$, $T = 5$, and $T = 10$, respectively. The realized response values are lower than those produced by AM-based algorithms, demonstrating the excellence of the proposed method in ISRJ suppression.

B. The joint design in multiple-channel system

To evaluate the ISRJ suppression performance in MIMO systems, we further consider a waveform set with $M = 3$, and $N = 128$. Since the MM-Acc algorithm can only solve the joint design problem in single-channel systems, we compare the proposed method with RGD-AM

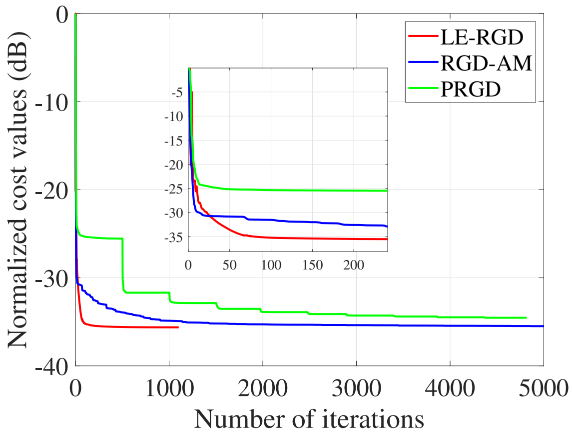


Fig. 4: The cost values versus the number of iterations corresponding to different methods.

and PRGD with our proposed method in this experiment. Moreover, the ISRJ and algorithm parameters are kept consistent with the experiment in Section IV.A, and the number of layers is set to $T = 3$.

In Fig.4, the normalized cost values versus the number of iterations associated with different algorithms are demonstrated. As we can observe, the PRGD algorithm has the fastest convergence speed in the first 25 iterations but is subsequently surpassed by the LE-RGD algorithm. Moreover, the proposed LE-RGD algorithm stops at the 1102-th iteration with a cost value of $-35.62\,\mathrm{dB}$, better than the outputted optima with the PRGD and RGD-AM algorithms, demonstrating the superiority in convergence speed of the proposed LE-RGD algorithm.

In Fig.5, we present the PC results of different waveform-filter pairs and ISRJ-filter pairs. As we can find in Fig.5 (a), for the MMF results of the waveform-filter pairs from the identical channels, high peaks can be formed in the zero-delay positions, accompanied by the PSL around $-19\,\mathrm{dB}$. Additionally, there are no high peaks for the waveform-filter pairs from different channels, showing that the proposed method can effectively mitigate the mutual interference among different channels. Moreover, we also present the MMF results corresponding to different ISRJ-filter pairs in Fig.5 (b). We can also find no high peaks in these results, either for the ISRJ-filter pairs from the identical channels or different channels, demonstrating the effectiveness of the proposed method in ISRJ suppression.

In TableI, we present several performance metrics to evaluate the transmit-receive joint design performance from different algorithms. In this table, the cross-channel peaks (CCP) denote the MMF peak values for the waveform-filter pairs from different channels. As we can observe, the proposed LE-RGD algorithm reaches an SNR loss of $-1.66\,\mathrm{dB}$, closer to the preset values μ_1 than those of the other two methods. In addition, it can also obtain the lowest PSL, CCP, and ISRJ peaks, demonstrating the transmit-receive optimization superiority in ISRJ suppression and target detection.

TABLE I: The performance metrics in dB realized by different algorithms.

	SNR Loss	PSL	CCP	ISRJ peaks
RGD-AM	-1.67	-17.8	-14.4	-23.52
PRGD	-1.67	-18.23	-16.29	-25.88
LE-RGD	-1.66	-18.97	-16.46	-26.87

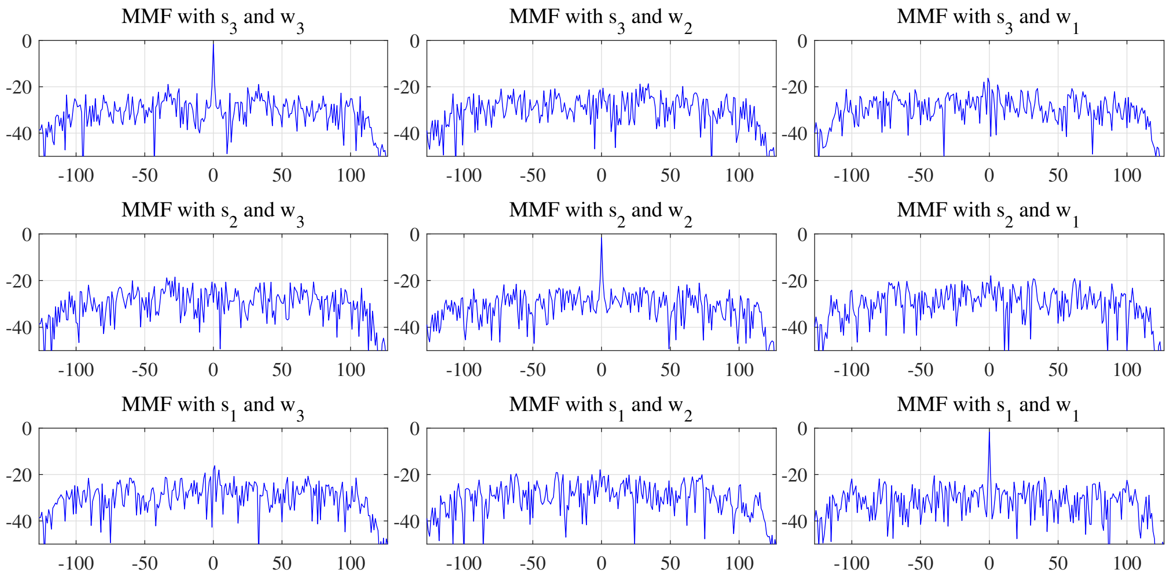
TABLE II: The simulated parameters of the ISRJ and targets.

ISRJ and target parameters	Values
ISRJ sample interval	$5\,\mu$s
ISRJ duty ratio	0.25
ISRJ location	$4.7\,\mathrm{km}$, $4.8\,\mathrm{km}$, $4.9\,\mathrm{km}$
Target location	$4.4\,\mathrm{km}$, $5.0\,\mathrm{km}$
Radar cross section (RCS)	$0\,\mathrm{dB}$, $-5\,\mathrm{dB}$
Signal to Jamming ratio (SJR)	$-15\,\mathrm{dB}$
SNR	$15\,\mathrm{dB}$

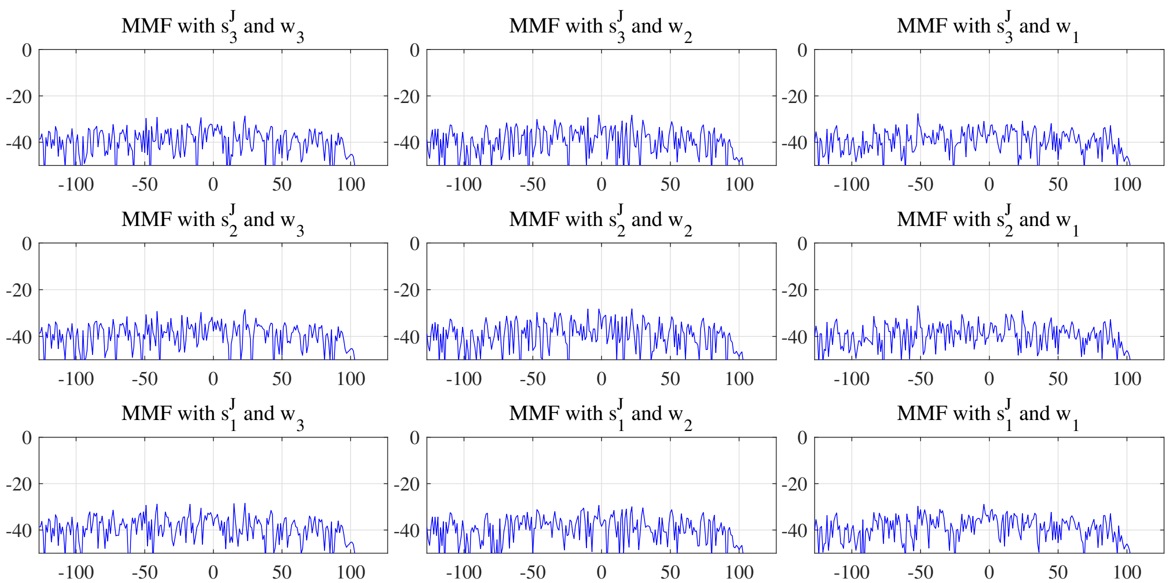
C. Evaluation in ISRJ suppression

To evaluate the ISRJ suppression and target detection performance, we compare the random phased-coded (RPC) waveform and matched filter (MF), the linear frequency modulated (LFM) waveform and MF, and the waveform and MMF designed using LE-RGD in the ISRJ scenario. The pulse width and bandwidth are set to $20\,\mu$s and $25\,\mathrm{MHz}$ for each waveform. Besides, the detailed parameters of the ISRJ and the targets are presented in Table II. We assume that the ISRJ jammer forwards the intercepted radar waveform three times with different times delays after intercepting it, and finally forms three false targets at $4700\,\mathrm{m}$, $4800\,\mathrm{m}$, and $4900\,\mathrm{m}$.

In Fig.6, the PC results corresponding to different waveforms are demonstrated to compare the ISRJ suppression performance. As we can observe in Fig.6 (a) and (b), due to the existence of the ISRJ components in the echoes, three false targets are generated after the MF processing in both channels. Precisely, the amplitudes of the false target are more than $4\,\mathrm{dB}$ higher than the real targets, posing great challenges to real target detection. In Fig.6 (c) and (d), the MF results associated with the LFM waveforms are presented. We assume that the LFM waveforms transmitted by the two channels have different frequency-modulated slopes, namely $1.25\,\mathrm{MHz}\,/\,\mu$s and $-1.25\,\mathrm{MHz}\,/\,\mu$s. Compared with the PC results of RPC waveforms, a group of false targets (whose amplitudes are more than $5\,\mathrm{dB}$ higher than the real targets) is formed in both channels, bringing more difficulties in identifying the real targets. In Fig.6 (e) and (f), the MMF results corresponding to the optimized waveform-filter pairs are



(a) the MMF results of different waveform-filter pairs



(b) the MMF result of different ISRJ-filter pairs

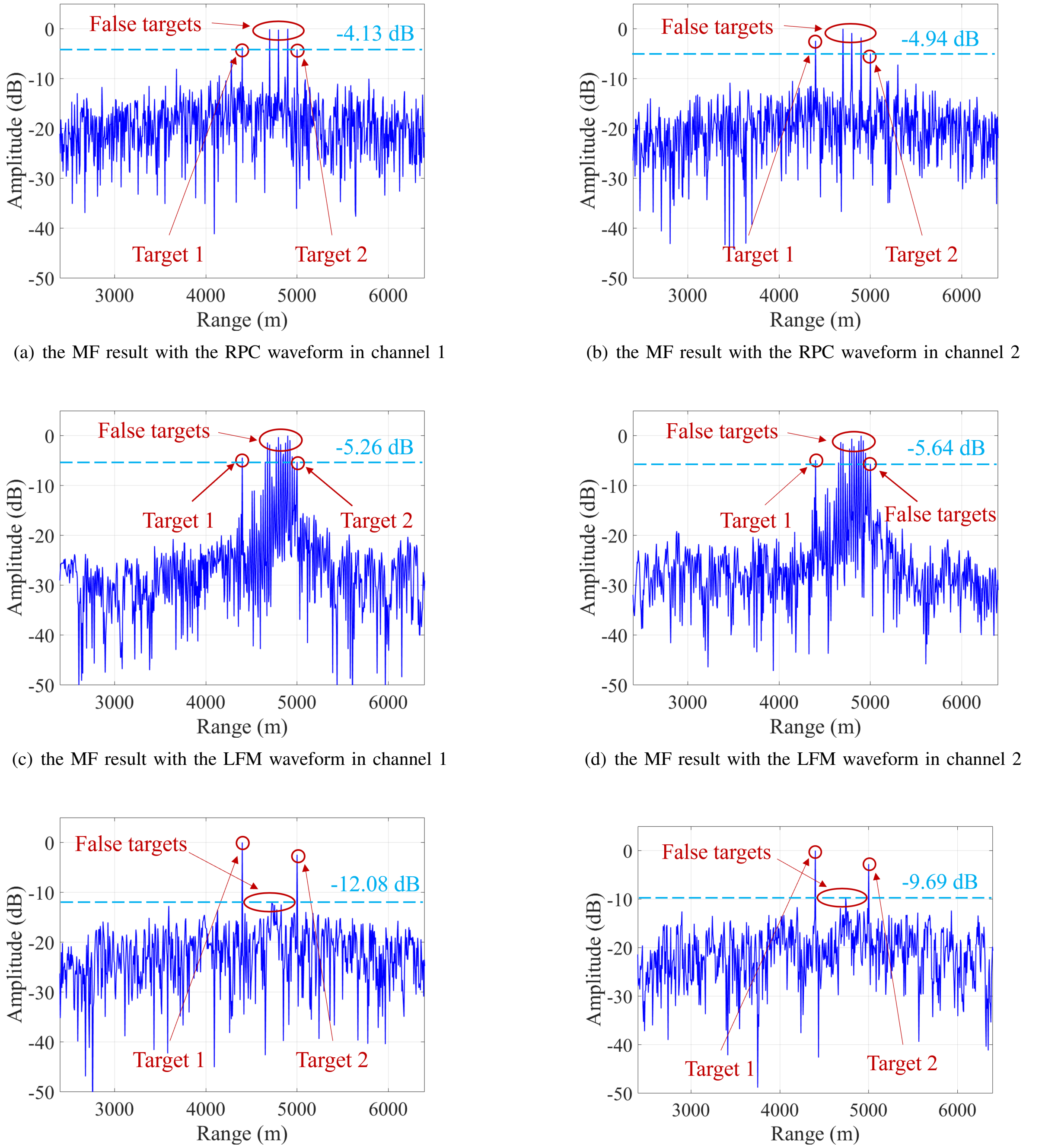
Fig. 5: The MMF results of different waveform-filter pairs and ISRJ-filter pairs. For each subfigure, the abscissa represents the range bins, and the ordinate represents the normalized response value in dB.

given as a comparison. As we can see, there are no high peaks within the interference area, and two real targets can be easily detected as they present $10\,\mathrm{dB}$ higher peaks than the ISRJ components. This further verifies the effectiveness and superiority of the proposed transmit-receive joint design method.

V. Conclusions

This work investigated the transmit-receive joint design problem in the MIMO system for the ISRJ suppression. Initially, we formulated the joint design problem as a joint-constrained non-convex problem by minimizing four different subfunctions. Then it is transformed into an unconstrained one on the Riemannian manifold space. In addition, by unfolding the classical RGD algorithm, we proposed a LE-RGD method to realize a trainable and adaptive update rule for the transmit waveform and receive filter. Finally, we conducted various numerical experiments to verify the effectiveness of our proposed approach.

Our established interference countermeasure realized the satisfactory performance in ISRJ suppression. However, since the sampling vector of ISRJ is assumed to



(a) the MF result with the RPC waveform in channel 1

(b) the MF result with the RPC waveform in channel 2

(c) the MF result with the LFM waveform in channel 1

(d) the MF result with the LFM waveform in channel 2

(e) the MMF result with the optimized waveform-filter in channel 1

(f) the MMF result with the optimized waveform-filter in channel 2

Fig. 6: the MF and MMF results associating with different waveforms in the ISRJ scenario.

be precisely known in the modeling process, this poses a huge challenge to the practical application of the proposed method. In this case, studying the estimation method of the sampling vector and the robust anti-ISRJ method via transmit-receive optimization can be the next research direction. In addition, in this work, we only considered the waveform optimization problem within a single pulse repetition interval, ignoring the Doppler modulation effect of the ISRJ jammer on the slow time dimension. Therefore, it is also worthwhile to design the transmit waveform and receive filter in both the fast and slow time dimensions to suppress the ISRJ containing Doppler modulation. Moreover, the nonlinear network structure used to update the transmit waveform and receive filter only includes some simple FC networks. The nonlinear mapping performance of other popular network structures, such as convolutional neural networks (CNN) and long short-term memory (LSTM) networks, is also worth exploring.

VI. APPENDIX

The Euclidean gradient of the function $f(\boldsymbol{S},\boldsymbol{W})$ w.r.t. the waveforms and filters can be derived by employing the following definition [36]

$$\begin{aligned}\mathrm{Grad}\, f_{\boldsymbol{S}}(\boldsymbol{S},\boldsymbol{W}) &= 2\frac{\partial}{\partial \boldsymbol{S}^*} f(\boldsymbol{S},\boldsymbol{W})\\ \mathrm{Grad}\, f_{\boldsymbol{W}}(\boldsymbol{S},\boldsymbol{W}) &= 2\frac{\partial}{\partial \boldsymbol{W}^*} f(\boldsymbol{S},\boldsymbol{W}),\end{aligned} \tag{21}$$

where $\frac{\partial}{\partial \boldsymbol{S}^*} f(\boldsymbol{S},\boldsymbol{W})$ denotes the partial derivative of the cost function w.r.t. $\boldsymbol{S}^*$, which can be calculated as

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{S}^*} f(\boldsymbol{S},\boldsymbol{W}) &= \sum_{k=1}^{4} \frac{\partial}{\partial \boldsymbol{S}^*} f_k(\boldsymbol{S},\boldsymbol{W})\\ \frac{\partial}{\partial \boldsymbol{W}^*} f(\boldsymbol{S},\boldsymbol{W}) &= \sum_{k=1}^{4} \frac{\partial}{\partial \boldsymbol{W}^*} f_k(\boldsymbol{S},\boldsymbol{W}).\end{aligned} \tag{22}$$

Before presenting the computation of the partial derivative, we first introduce a commonly used theorem.

Lemma 1: For a real-valued complex function $g(\boldsymbol{Z},\boldsymbol{Z}^*)$ with $\boldsymbol{Z} \in \mathbb{C}^{N\times M}$, if its differential can be expressed as

$$dg(\boldsymbol{Z},\boldsymbol{Z}^*) = \mathrm{Tr}\left\{\boldsymbol{A}_0^T d\boldsymbol{Z} + \boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}, \tag{23}$$

then the derivatives of $g(\boldsymbol{Z},\boldsymbol{Z}^*)$ can be obtained as [37]

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{Z}} g(\boldsymbol{Z},\boldsymbol{Z}^*) &= \boldsymbol{A}_0,\\ \frac{\partial}{\partial \boldsymbol{Z}^*} g(\boldsymbol{Z},\boldsymbol{Z}^*) &= \boldsymbol{A}_1.\end{aligned} \tag{24}$$

Here the variable $\boldsymbol{Z}$ and its conjugate transpose are treated as two independent variables, and $d(\cdot)$ represents the differential operation.

For the function component $f_1(\boldsymbol{S},\boldsymbol{W})$, taking $\boldsymbol{S}$ as a variable and differentiating the function, we have

$$\begin{aligned} d\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right\|_F^2 &= d\,\mathrm{Tr}\left(\boldsymbol{S}^H\boldsymbol{J}_r^H\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right)\\ &= \mathrm{Tr}\left[(\boldsymbol{\Phi}_1\boldsymbol{S})^T d\boldsymbol{S}^* + \boldsymbol{S}^H\boldsymbol{\Phi}_1 d\boldsymbol{S}\right],\end{aligned} \tag{25}$$

where $\boldsymbol{\Phi}_1 = \boldsymbol{J}_r^H\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{J}_r$. Further, we can obtain

$$\frac{\partial}{\partial \boldsymbol{S}^*}\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right\|_F^2 = \boldsymbol{\Phi}_1\boldsymbol{S} = \boldsymbol{J}_r^H\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}. \tag{26}$$

In the same way, by taking $\boldsymbol{W}$ as the variable, we have

$$\begin{aligned} d\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right\|_F^2 &= d\,\mathrm{Tr}\left(\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{J}_r^H\boldsymbol{W}\right)\\ &= \mathrm{Tr}\left[(\boldsymbol{\Phi}_2\boldsymbol{W})^T (d\boldsymbol{W})^* + \boldsymbol{W}^H\boldsymbol{\Phi}_2 (d\boldsymbol{W})\right]\end{aligned} \tag{27}$$

where $\boldsymbol{\Phi}_2 = \boldsymbol{J}_r\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{J}_r^H$. Further, we have

$$\frac{\partial}{\partial \boldsymbol{W}^*}\left\|\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S}\right\|_F^2 = \boldsymbol{\Phi}_2\boldsymbol{W} = \boldsymbol{J}_r\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{J}_r^H\boldsymbol{W}. \tag{28}$$

Similarly, we can also obtain

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{S}^*}\left\|(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{A}\right\|_F^2 &= \boldsymbol{W}\left[(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{A}\odot\boldsymbol{A}\right]\\ \frac{\partial}{\partial \boldsymbol{W}^*}\left\|(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{A}\right\|_F^2 &= \boldsymbol{S}\left[(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{A}\odot\boldsymbol{A}\right]^H.\end{aligned} \tag{29}$$

Combing equations (26), (28) and (29), the following equation holds

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{S}^*} f_1 &= \sum_{r\neq 0}\boldsymbol{J}_r^H\boldsymbol{W}\boldsymbol{W}^H\boldsymbol{J}_r\boldsymbol{S} + \boldsymbol{W}\left[(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{A}\right]\\ \frac{\partial}{\partial \boldsymbol{W}^*} f_1 &= \sum_{r\neq 0}\boldsymbol{J}_r\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{J}_r^H\boldsymbol{W} + \boldsymbol{S}\left[(\boldsymbol{S}^H\boldsymbol{W})\odot\boldsymbol{A}^H\right].\end{aligned} \tag{30}$$

Repeating ***Lemma 1*** to the other three function components, we can obtain the following derivatives:

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{S}^*} f_2 &= \boldsymbol{W}\left\{\left[(\boldsymbol{W}^H\boldsymbol{S})\odot\boldsymbol{I} - \mu_1\boldsymbol{I}\right]\odot\boldsymbol{I}\right\}\\ \frac{\partial}{\partial \boldsymbol{W}^*} f_2 &= \boldsymbol{S}\left\{\left[(\boldsymbol{S}^H\boldsymbol{W})\odot\boldsymbol{I} - \mu_1\boldsymbol{I}\right]\odot\boldsymbol{I}\right\}\\ \frac{\partial}{\partial \boldsymbol{S}^*} f_3 &= \sum_{r}\tilde{\boldsymbol{J}}_r^H\boldsymbol{W}\boldsymbol{W}^H\tilde{\boldsymbol{J}}_r\boldsymbol{S}\\ \frac{\partial}{\partial \boldsymbol{W}^*} f_3 &= \sum_{r}\tilde{\boldsymbol{J}}_r\boldsymbol{S}\boldsymbol{S}^H\tilde{\boldsymbol{J}}_r^H\boldsymbol{W}\\ \frac{\partial}{\partial \boldsymbol{S}^*} f_4 &= \boldsymbol{P}^H\boldsymbol{W}\left\{\left[(\boldsymbol{W}^H\boldsymbol{P}\boldsymbol{S})\odot\boldsymbol{I} - \mu_2\boldsymbol{I}\right]\odot\boldsymbol{I}\right\}\\ \frac{\partial}{\partial \boldsymbol{W}^*} f_4 &= \boldsymbol{P}\boldsymbol{S}\left\{\left[(\boldsymbol{S}^H\boldsymbol{P}^H\boldsymbol{W})\odot\boldsymbol{I} - \mu_2\boldsymbol{I}\right]\odot\boldsymbol{I}\right\}.\end{aligned} \tag{31}$$

Thus, combining with equations (21), (22), (30) and (31), we can obtain the Euclidean gradients of the cost function $f(\boldsymbol{S},\boldsymbol{W})$, as shown in (14).

REFERENCES

[1] WANG, XueSong, LIU, JianCheng, ZHANG, WenMing, FU, QiXiang, LIU, and Zhong, “Mathematic principles of interrupted-sampling repeater jamming (isrj),” *Science in China*, 2007.

[2] J. Wei, Y. Li, R. Yang, J. Wang, M. Ding, and J. Ding, “A nonuniformly distributed multipulse coded waveform to combat azimuth interrupted sampling repeater jamming in sar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 9054–9066, 2023.

[3] H. Wang, T. Hu, T. Li, D. Wu, Z. Xu, X. Xu, and Z. Tian, “Intelligent interference waveform design for radar detection based on the cross-correlation value function,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 2, pp. 2061–2070, 2024.

[4] Y. Zhang, Y. Wei, and L. Yu, “Isrj rejection with mimo radar using space-time complex modulation in multi-jamming scenario,” *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[5] S. Wei, Y. Fang, Y. Song, S. Wei, and L. Zhang, “Joint jam perception and adaptive waveform optimization for anti-interrupted sampling repeater jamming,” *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[6] M. S. Greco, F. Gini, P. Stinco, and K. Bell, “Cognitive radars: On the road to reality: Progress thus far and possibilities for the future,” *IEEE Signal Processing Magazine*, vol. 35, no. 4, pp. 112–125, 2018.

[7] S. Z. Gurbuz, H. D. Griffiths, A. Charlish, M. Rangaswamy, M. S. Greco, and K. Bell, “An overview of cognitive radar: Past, present, and future,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 12, pp. 6–18, 2019.

[8] S. D. Blunt and E. L. Mokole, “Overview of radar waveform diversity,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 11, pp. 2–42, 2016.

[9] F. Gini, A. De Maio, and L. Patton, *Waveform design and diversity for advanced radar systems*. Institution of engineering and technology London, UK, 2012.

[10] F. Cao, Z.-h. Chen, X.-w. Feng, C. He, and J.-f. Xu, "Optimal design of anti-interrupted sampling repeater jamming waveform for missile-borne radar based on an improved genetic algorithm," *IET Signal Processing*, vol. 15, no. 9, pp. 622–632, 2021.

[11] Z. Liu, Y. Quan, S. Du, Y. Wu, M. Sha, and M. Xing, "A novel eccm scheme against interrupted-sampling repeater jamming using intra-pulse dual-parameter agile waveform," *Digital Signal Processing*, vol. 129, p. 103652, 2022.

[12] J. He, Z. Cheng, and Z. He, "Joint design of detecting and screening waveforms against interrupted sampling repeater jamming with inaccurate jamming knowledge," *Digital Signal Processing*, vol. 130, p. 103679, 2022.

[13] K. Zhou, D. Li, S. Quan, T. Liu, Y. Su, and F. He, "Sar waveform and mismatched filter design for countering interrupted-sampling repeater jamming," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2021.

[14] M. Ge, X. Yu, Z. Yan, G. Cui, and L. Kong, "Joint cognitive optimization of transmit waveform and receive filter against deceptive interference," *Signal Processing*, vol. 185, p. 108084, 2021.

[15] F. Wang, N. Li, C. Pang, C. Li, Y. Li, and X. Wang, "Complementary sequences and receiving filters design for suppressing interrupted sampling repeater jamming," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[16] Y. Gao, H. Fan, L. Ren, Z. Liu, Q. Liu, and E. Mao, "Joint design of waveform and mismatched filter for interrupted sampling repeater jamming suppression," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[17] B. Kang, J. Kweon, M. Rangaswamy, and V. Monga, "Deep learning for radar waveform design: Retrospectives and the road ahead," in *2023 IEEE International Radar Conference (RADAR)*. IEEE, 2023, pp. 1–6.

[18] Y. Pei, J. Hu, K. Zhong, and J. Wu, "Mimo radar waveform optimization by deep learning method," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 811–814.

[19] J. Zhao, W. Li, Y. Gan, H. Wang, and K. Zou, "Waveform design of cognitive radar based on deep learning in the presence of clutter," 2021.

[20] K. Zhong, W. Zhang, Q. Zhang, J. Hu, P. Wang, X. Yu, and Q. Zhou, "Mimo radar waveform design via deep learning," in *2021 IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–5.

[21] W. Zhang, J. Hu, Z. Wei, H. Ma, X. Yu, and H. Li, "Constant modulus waveform design for mimo radar transmit beampattern with residual network," *Signal Processing*, vol. 177, p. 107735, 2020.

[22] J. Hu, Z. Wei, Y. Li, H. Li, and J. Wu, "Designing unimodular waveform (s) for mimo radar by deep learning method," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 2, pp. 1184–1196, 2020.

[23] M. Xia, S. Chen, and X. Yang, "New optimization method based on neural networks for designing radar waveforms with good correlation properties," *IEEE Access*, vol. 9, pp. 91 314–91 323, 2021.

[24] V. Hassija, V. Chamola, A. Mahapatra, A. Singal, D. Goel, K. Huang, S. Scardapane, I. Spinelli, M. Mahmud, and A. Hussain, "Interpreting black-box models: a review on explainable artificial intelligence," *Cognitive Computation*, vol. 16, no. 1, pp. 45–74, 2024.

[25] K. Metwaly, J. Kweon, K. Alhujaili, M. Greco, F. Gini, and V. Monga, "Interpretable, unrolled deep radar beampattern design," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[26] K. Zhong, J. Hu, Y. Wang, Z. Zhao, X. Yu, G. Cui, B. Liao, and H. Li, "Mimo radar unimodular waveform design with learned complex circle manifold network," *IEEE Transactions on Aerospace and Electronic Systems*, 2024.

[27] Z. Zhao, J. Hu, K. Zhong, Y. Zuo, H. Li, and B. Zhang, "Mimo radar waveform design for range-isl optimization via iterative deep unfolding network," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[28] X. Qiu, X. Zhang, K. Huo, and W. Jiang, "Quartic riemannian adaptive regularization with cubics for radar waveform design," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[29] X. Qiu, C. Yang, W. Jiang, and Y. Liu, "Joint design of transmit waveform and receive filter for cross ambiguity function synthesis," *IEEE Wireless Communications Letters*, 2023.

[30] X. Qiu, W. Jiang, X. Zhang, and K. Huo, "Quartic riemannian trust region algorithm for cognitive radar ambiguity function shaping," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[31] C. Liu and N. Boumal, "Simple algorithms for optimization on riemannian manifolds with constraints," *Applied Mathematics & Optimization*, vol. 82, no. 3, pp. 949–981, 2020.

[32] S. T. Smith, "Optimization techniques on riemannian manifolds," *arXiv preprint arXiv:1407.5965*, 2014.

[33] N. Boumal, *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023.

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

[35] J. Li, G. Liao, Y. Huang, Z. Zhang, and A. Nehorai, "Riemannian geometric optimization methods for joint design of transmit sequence and receive filter on mimo radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5602–5616, 2020.

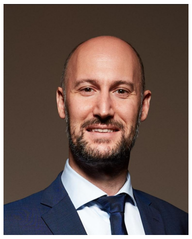
[36] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2002.

[37] A. Hjørungnes, *Complex-valued matrix derivatives: with applications in signal processing and communications*. Cambridge University Press, 2011.

Xiangfeng Qiu was born in Anhui, China, in 1997. He received a B.S. degree from Hainan University, Haikou, China, in 2019, and an M.S. degree from the National University of Defense Technology (NUDT), China, in 2021. From 2016 to 2018, he attended a joint training program at Tianjin University in Tianjin, China. He is currently a Ph.D. student at NUDT and a visiting scholar at the University of Pisa. His current research interests include numerical optimization and radar waveform design.

Weidong Jiang was born in Chongqing, China, in 1968. He received a B.S. degree in communication engineering and a Ph.D. degree in electronic science and technology from the National University of Defense Technology (NUDT), China, in 1991 and 2001, respectively. He is currently a Professor with NUDT. His current research interests include MIMO radar signal processing and radar system technology.

Yongxiang Liu received the B.S. and Ph.D. degrees from the National University of Defense Technology (NUDT), China, in 1999 and 2004, respectively. Since 2004, he has been with NUDT and is currently a Professor with the College of Electrical Science and Engineering, researching radar target recognition, time-frequency analysis, micro motions, and array signal processing. He worked at Imperial College London as an Academic Visitor in 2008.



Symeon Chatzinotas (Fellow, IEEE) received the M.Eng. degree in telecommunications from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2003, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, Guildford, U.K., in 2006 and 2009, respectively.,He is currently a Full Professor or a Chief Scientist I and the Head of the SIGCOM Research Group, Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Esch-sur-Alzette, Luxembourg. In the past, he was a Visiting Professor with the University of Parma, Parma, Italy, and he was involved in numerous research and development projects for the National Center for Scientific Research Demokritos, the Center of Research and Technology Hellas, and the Center of Communication Systems Research, University of Surrey. He has coauthored more than 400 technical papers in refereed international journals, conferences, and scientific books.,Prof. Chatzinotas was the co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award, the CROWNCOM 2015 Best Paper Award, and the 2018 EURASIP JWCN Best Paper Award. He is currently in the editorial board of the IEEE Open Journal of Vehicular Technology and the International Journal of Satellite Communications and Networking.

Fulvio Gini (Fellow, IEEE) received the Doctor Engineer (cum laude) and Research Doctor degrees in electronic engineering from the University of Pisa, Pisa, Italy, in 1990 and 1995, respectively. In 1993, he joined the Department of Ingegneria dell'Informazione, University of Pisa, where he has been a Full Professor since 2006 and the Deputy Head of the department since November 2016. He authored or coauthored 13 book chapters, 154 journal articles, and 183 conference papers, and holds two national patents.,Dr. Gini is a member of the IEEE AES Society Board of Governors (BoG) from 2017 to 2022 and the IEEE SP Society BoG from 2021 to 2023. He was a co-recipient of the 2001 and 2012 IEEE AES Society's Barry Carlton Award for Best Paper published in the IEEE Transactions on Aerospace and Electronic Systems (T-AES) and the 2020 EURASIP JASP Best Paper Award. He was a recipient of the 2022 IEEE AESS Warren White Award for Excellence in Radar Engineering, the 2020 EURASIP Meritorious Service Award, the 2003 IEEE Achievement Award, and the 2003 IEEE AES Society Nathanson Award. He is the IEEE AES Society Awards Chair. He was the EURASIP President from 2013 to 2016. He was the General Co-Chair for ICASSP 2014. He has been an Associate Editor of the IEEE T-AES since 2007 and the Elsevier Signal Processing journal since 2006. He has also been the Specialty Chief Editor of the "Radar Signal Processing" section of the journal Frontiers in Signal Processing since 2021. His H-index is 52, and he has 10467 citations.

Maria Sabrina Greco (Fellow, IEEE) graduated in electronic engineering from the University of Pisa, Pisa, Italy, in 1993, and received the Ph.D. degree in telecommunication engineering from the University of Pisa, in 1998.,From December 1997 to May 1998, she was with Georgia Tech Research Institute, Atlanta, GA, USA, as a Visiting Research Scholar. In 1993, she joined the Department of Information Engineering, University of Pisa, where, since 2017, she has been a Full Professor. From May 2015 to June 2015 and from January 2018 to February 2018, she visited as an Invited Professor at Universitá Paris-Sud, CentraleSupélec, Paris, France. She has coauthored many book chapters and about 200 journal articles and conference papers. Her research interests include statistical signal processing, estimation and detection theory, in particular, clutter models, coherent and incoherent detection in non-Gaussian clutter, CFAR techniques, radar waveform diversity, bistatic or mustistatic active and passive radars, and cognitive radars.,Dr. Greco is a member of the IEEE AESS Board of Governors (BoG) and the IEEE SPS BoG from 2015 to 2017. She was a co-recipient of the 2001 and 2012 IEEE Aerospace and Electronic Systems Society's Barry Carlton Awards for Best Paper and the 2019 EURASIP JASP Best Paper Award and a recipient of the 2008 Fred Nathanson Young Engineer of the Year Award for contributions to signal processing, estimation, and detection theory, and the IEEE AESS Board of Governors Exceptional Service Award for Exemplary Service and Dedication and Professionalism, as the Editor-in-Chief of the IEEE Aerospace and Electronic Systems Magazine (AES). She was the Chair for the IEEE AESS Radar Panel from 2015 to 2016. She was the general chair, the technical program chair, and an organizing committee member for many international conferences for more than ten years. She was the Lead-Guest Editor for a few special issues. She is an Associate Editor of the IET Proceedings-Radar, Sonar and Navigation and the IET Signal Processing, and the Editor-in-Chief of the Springer Journal of Advances in Signal Processing. She was the SPS Distinguished Lecturer from 2014 to 2015, the AESS Distinguished Lecturer from 2015 to 2020, and the AESS VP Publications from 2018 to 2020. She is currently the IEEE SPS Director-at-Large for Region 8.