Article

Designing Waveform Sets with Good Correlation and Stopband Properties for MIMO Radar via the Gradient-Based Method

Liang Tang *, Yongfeng Zhu and Qiang Fu

College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China; zoyofo@163.com (Y.Z.); fuqiang1962@vip.sina.com (Q.F.)

* Correspondence: tangliang12@nudt.edu.cn; Tel.: +86-731-8457-6401

Abstract: Waveform sets with good correlation and/or stopband properties have received extensive attention and been widely used in multiple-input multiple-output (MIMO) radar. In this paper, we aim at designing unimodular waveform sets with good correlation and stopband properties. To formulate the problem, we construct two criteria to measure the correlation and stopband properties and then establish an unconstrained problem in the frequency domain. After deducing the phase gradient and the step size, an efficient gradient-based algorithm with monotonicity is proposed to minimize the objective function directly. For the design problem without considering the correlation weights, we develop a simplified algorithm, which only requires a few fast Fourier transform (FFT) operations and is more efficient. Because both of the algorithms can be implemented via the FFT operations and the Hadamard product, they are computationally efficient and can be used to design waveform sets with a large waveform number and waveform length. Numerical experiments show that the proposed algorithms can provide better performance than the state-of-the-art algorithms in terms of the computational complexity.

1. Introduction

Waveform design has received considerable attention in recent years [1] and been employed in many applications, including polarimetric radar [2], multiple-input multiple-output (MIMO) radar [3,4], stealth communications [5] and the code-division multiple-access (CDMA) system. As a category of the general waveform design research [1], the design of waveform sets (i.e., multidimensional waveforms) is an important research content of MIMO radar. Generally, waveform sets are desired to have a good correlation property, which can effectively improve radar resolution, detection performance, imaging quality, the ability to obtain information and the accuracy of MIMO channel estimation [6–8]. In recent years, a large number of scholars has been devoted to designing waveform sets with a good correlation property. The main research covers two aspects: one is the waveform sets with good auto- and cross-correlation properties [9–16], and the other is the complementary sets of sequences (CSS) [17–25].

Waveform sets with good auto- and cross-correlation properties, also known as the orthogonal waveform set (OWS), have low autocorrelation sidelobes and low cross-correlation levels. In the early stage, the simulated annealing- [9] and cross entropy-based [10] methods were proposed for OWS design. However, due to the high computational complexity, these methods are not suited to design long waveforms. To improve the computational efficiency, the CAN (cyclic algorithm-new) algorithm

[11] based on fast Fourier transform (FFT) is proposed to minimize the autocorrelation sidelobes and the cross-correlation. This algorithm is computationally efficient and can be used for the design of long waveforms. As it is impossible to design completely orthogonal (i.e., both the autocorrelation sidelobes and the cross-correlation are zeroes) waveform sets [11,16], [11] proposes to design the waveform sets that are orthogonal only at the specified intervals and extends the classical WeCAN (weighted cyclic algorithm-new) [26] algorithm to MIMO radar. To solve the same problem, [12] develops the LBFGS (limited-memory Broyden, Fletcher, Goldfarb and Shanno) iterative algorithm, which is more efficient than the WeCAN algorithm. However, because of the complicated linear search rule for determining the step size, the LBFGS iterative algorithm is still time consuming. Recently, the majorization-minimization (MM)-based algorithms (i.e., MM-Corr (MM-correlation) and MM-WeCorr (MM-weighted correlation)) are proposed in [14]. These two algorithms are also based on FFT operations and much faster than the CAN and WeCAN algorithms [11].

Another waveform set with a good correlation property is the complementary sets of sequences (CSS). A waveform set is called CSS if and only if the autocorrelation sum of the waveforms is a delta function [17]. CSS design is proposed to overcome the difficulty of generating a single unimodular waveform with ideal (impulse-like) autocorrelation. A common application of the CSS is the pulse compression [18–20]. In pulse compression radar, the complementary sequences are used to modulate consecutive pulses in a coherent pulse train. Then, the autocorrelation sidelobes can be reduced via the coherent [19] or noncoherent [18] accumulation, which can be regarded as the process of obtaining the autocorrelation sum of the complementary sequences. Moreover, due to the good correlation property, CSS has been widely applied to the CDMA system [21], ISI (intersymbol interference) channel estimation [22], orthogonal frequency division multiplexing (OFDM) [23], and many other areas. The main methods of designing CSS are the analytical construction methods, which have great limitation in generating long waveforms. To overcome this problem, [24] introduces a computational framework based on an iterative twisted approximation (ITROX) for periodically complementary sets of sequences design. Subsequently, [25] extends the CAN algorithm [26] and develops a fast algorithm named CANARY (CAN complementary). Additionally, [14] applies the MM method to the design of CSS.

In addition to the good correlation property, waveform sets are expected to have a good stopband property when the radar systems work in a crowded electromagnetic environment. Waveforms with the stopband property, also known as the sparse frequency waveforms (SFW) in many literature works, are a kind of waveforms with several frequency stopbands. The applications of SFW include ultra-wide bandwidth (UWB) systems [27], high frequency surface wave radar (HFSWR) [28,29] and cognitive radar [30]. By designing waveforms with the stopband property, it can effectively overcome the narrowband interference in the congested frequency bands. At present, there are many research works on the design of a single waveform with the stopband property [31–36], but they cannot be used in MIMO systems. Therefore, [36] proposes an iterative algorithm combined with the steepest descent (SD) method for MIMO waveform design. By searching along the gradient direction, the convergence speed of this algorithm is improved. However, the computation of the step size along the gradient direction is complicated, which makes the algorithm still costly. In order to improve the computational efficiency, [37] proposes an algorithm named MDISAA-SFW (multi-dimensional iterative spectral approximation algorithm-SFW) based on alternating projection and phase retrieval.

In this paper, we consider the problem of designing unimodular waveform sets with good correlation and stopband properties and propose a gradient-based algorithm, i.e., Gra-WeCorr-SFW (gradient-weighted correlation-SFW). By using the relationship between the correlation function and the power spectrum density (PSD), the design problem is formulated as an unconstrained minimization problem in the frequency domain. Then, the phase gradient is deduced, and its matrix form is given. In order to avoid searching the step size, we use the Taylor series expansion to derive the step size, which is more efficient than the traditional searching methods. Since both the gradient and the step size can be implemented via the FFT operations and the Hadamard product, the proposed algorithm has high computational efficiency. We also deduce the simplified algorithm named Gra-Corr-SFW

(gradient-correlation-SFW), which is faster than the Gra-WeCorr-SFW, for the design problem without considering the correlation weights.

The rest of the paper is organized as follows. In Section 2, the design problem is formulated. In Section 3, we develop a gradient-based algorithm by deducing the phase gradient and the step size and then summarize the algorithm. In Section 4, the simplified algorithm for the design problem without considering the correlation weights is derived. Section 5 provides several numerical experiments to verify the effectiveness of the proposed algorithms. Finally, Section 6 gives the conclusions.

Notation: Boldface upper case letters denote matrices, while boldface lower case letters denote column vectors. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the complex conjugate, transpose and conjugate transpose, respectively. $\|\cdot\|$ and $\|\cdot\|_F$ denote the Euclidean norm and the Frobenius norm. $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary part, respectively. $\mathrm{Diag}(\mathbf{x})$ denotes a diagonal matrix formed with the column vector $\mathbf{x}$. $\circ$ denotes the Hadamard product. $x(m)$ denotes the m-th element of the vector $\mathbf{x}$. $\mathbf{x}^l$ is the l-th iteration of $\mathbf{x}$. $\mathbf{1}_N$ is the all-one vectors of length N. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $\mathcal{F}(\mathbf{x})$ and $\mathcal{F}^{-1}(\mathbf{x})$ denote the 2N-point FFT and IFFT (inverse FFT) operations of $\mathbf{x}$, respectively. $\mathcal{F}_c(\mathbf{X})$ and $\mathcal{F}_c^{-1}(\mathbf{X})$ represent the FFT and IFFT of each column of the matrix $\mathbf{X}$, respectively. In the (I)FFT operations, if the length of $\mathbf{x}$ is less than 2N, $\mathbf{x}$ is padded with trailing zeros to length 2N. $e^{(\cdot)}$ is the exponent arithmetic applied to the scalar, vector or matrix.

2. Problem Formulation

As mentioned in the Introduction, this paper focuses on the problem of designing waveform sets with good correlation and stopband properties. Therefore, we first establish two criteria in the frequency domain to measure the correlation and stopband properties and then formulate the waveform set design problem.

2.1. The Criterion for Good Correlation Property

Let $\{\mathbf{x}_m\}_{m=1}^M$ be the complex waveform set to be designed. The vector form of each waveform can be expressed as:

$$\mathbf{x}_m = [x_m(1), x_m(2), ..., x_m(N)]^T, \; m = 1, ..., M, \tag{1}$$

where N is the waveform length and M denotes the number of the waveforms. Then, the correlation function of $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as:

$$r_{ij}(k) = \sum_{n=k+1}^{N} x_i(n)x_j^*(n-k) = r_{ji}^*(-k), \; i,j = 1, ..., M, \; k = 1 - N, ..., N - 1. \tag{2}$$

Here, we consider the design problems of complementary sets of sequences and waveform sets with both good auto- and cross-correlation properties. For the complementary set of sequences, the common criterion is the complementary integrated sidelobe level (CISL) metric [14,25], which is defined as:

$$\mathrm{CISL} = \sum_{\substack{k=1-N \\ k \neq 0}}^{N-1} \left| \sum_{m=1}^{M} r_{mm}(k) \right|^2. \tag{3}$$

for the waveform sets with both good auto- and cross-correlation properties, we consider the following more general weighted measure [14]:

$$\psi = \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k=1-N}^{N-1} w_t(k) \left| r_{ij}(k) \right|^2 - w_t(0)MN^2, \tag{4}$$

where $w_t(k) = w_t(-k) \geq 0, k = 0, ..., N - 1$ denote the weights assigned to different time lags. In those specified intervals that are expected to have as low as possible sidelobes, $w_t(k) = 1$, else $w_t(k) = 0$.

Let:

$$\mathbf{r}_{ij} = \left[r_{ij}(0), r_{ij}(1), ..., r_{ij}(N-1), 0, r_{ij}(1-N), ..., r_{ij}(-1) \right]^T,$$
$$\mathbf{w}_t = \left[w_t(0), w_t(1), ..., w_t(N-1), 1, w_t(1-N), ..., w_t(-1) \right]^T \tag{5}$$

be the correlation vector and the corresponding correlation weight vector, respectively. The Fourier transform of the correlation function $\mathbf{r}_{ij}$ (i.e., power spectrum density (PSD)) can be written as:

$$\mathbf{p}_{ij} = \mathcal{F}\left(\mathbf{r}_{ij}\right) = \mathbf{F}\mathbf{r}_{ij},$$
$$\mathbf{r}_{ij} = \mathcal{F}^{-1}\left(\mathbf{p}_{ij}\right) = \frac{1}{2N}\mathbf{F}^H\mathbf{p}_{ij}, \tag{6}$$

where $\mathbf{F}$ is the $2N \times 2N$ discrete Fourier transform (DFT) matrix with the following expression:

$$\mathbf{F}(m,n) = e^{-j\frac{2mn\pi}{2N}}, 1 \le m, n \le 2N. \tag{7}$$

According to (5) and (6), (4) can be expressed in the frequency domain as:

$$\psi = \sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{r}_{ij}^H\mathbf{D}_t\mathbf{r}_{ij} - w_t(0)MN^2 = \frac{1}{4N^2}\sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{p}_{ij}^H\mathbf{F}\mathbf{D}_t\mathbf{F}^H\mathbf{p}_{ij} - w_t(0)MN^2, \tag{8}$$

where $\mathbf{D}_t = \mathrm{Diag}\left(\mathbf{w}_t\right)$. It is worth noting that the value of the $N+1$-th element of $\mathbf{w}_t$ has no effect on the objective function. In order to facilitate the derivation below, the $N+1$-th element of $\mathbf{w}_t$ is set to be one.

Actually, when $w_t(k) = 1, k = 1-N, ..., N-1$ (i.e., the correlation weights are not taken into account), Criterion (8) is equivalent to Criterion (3) (see Appendix A), which means that Criterion (3) is a special case of Criterion (8). Thus, here, we just consider Criterion (8). By ignoring the constant term in (8), the criterion related to the correlation property is given by:

$$J_{CF} = \frac{1}{4N^2}\sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{p}_{ij}^H\mathbf{F}\mathbf{D}_t\mathbf{F}^H\mathbf{p}_{ij}. \tag{9}$$

2.2. The Criterion for the Good Stopband Property

The waveform set with a good stopband property means that the PSD of each waveform has several frequency stopbands. Without loss of generality, we consider that the frequency is normalized. Define the set of frequency stopbands as:

$$\Omega_f = \bigcup_{k=1}^{n_s}(f_{k1}, f_{k2}) \subset [0,1], \tag{10}$$

where (f_{k1}, f_{k2}) denotes one stopband and n_s denotes the number of the stopbands. Considering the $2N$-point FFT operations, the corresponding point set of the stopband set Ω_f can be expressed as:

$$\overline{\Omega}_f = \bigcup_{k=1}^{n_s}(2Nf_{k1}, 2Nf_{k2}) \subset [0, 2N]. \tag{11}$$

Define the frequency weight vector as:

$$\mathbf{w}_f = \left[w_f(1), w_f(2), ..., w_f(2N) \right]^T,$$
$$w_f(p) = \begin{cases} 1, & p \in \overline{\Omega}_f \\ 0, & \text{otherwise} \end{cases}. \tag{12}$$

As the PSD of each waveform is nonnegative, the criterion related to the stopband property (or PSD) can be formulated as:

$$J_{PSD} = \sum_{i=1}^{M} \left\| \mathbf{w}_f \circ \mathbf{p}_{ii} \right\|^2 = \sum_{i=1}^{M} \mathbf{p}_{ii}^{H} \mathbf{D}_f \mathbf{p}_{ii}, \tag{13}$$

where $\mathbf{D}_f = \mathrm{Diag}\left(\mathbf{w}_f\right)$.

2.3. The Minimization Problem

In order to obtain good correlation and stopband properties, both J_{CF} and J_{PSD} should be minimized. Thus, the waveform design problem can be regarded as a multi-objective optimization (also known as Pareto optimization) problem, i.e., finding the Pareto optimal solutions that satisfy the constraints. Here, we apply the traditional weighting method to solve the problem and, thus, formulate a single-objective function as follows:

$$J_T = \lambda J_{PSD} + (1 - \lambda) J_{CF} = \sum_{i=1}^{M} \mathbf{p}_{ii}^{H} \mathbf{Q} \mathbf{p}_{ii} + \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ij}^{H} \mathbf{Q}' \mathbf{p}_{ij}, \tag{14}$$

where $\lambda \in [0, 1]$ is a weight coefficient by which we can balance the relative weight between J_{CF} and J_{PSD}, and:

$$\mathbf{Q} = \lambda \mathbf{D}_f, \quad \mathbf{Q}' = \frac{1}{4N^2} (1 - \lambda) \mathbf{F} \mathbf{D}_t \mathbf{F}^{H}. \tag{15}$$

When $\lambda = 0$, (14) is the criterion for designing waveform sets with a good correlation property. Additionally, when $\lambda \in (0, 1]$, (14) becomes the criterion for SFW design. Generally, for maximizing the transmitter efficiency and reducing the requirement to the hardware, the unimodular constraint [38] is required in the waveform design. Therefore, the design problem can be formulated as the following minimization problem:

$$\min_{\{\mathbf{x}_m\}_{m=1}^{M}} \quad J_T \\ \text{s.t.} \quad |x_m(n)| = 1, \quad n = 1, ..., N, \quad m = 1, ..., M. \tag{16}$$

Let $\boldsymbol{\phi}_m$ denote the phase vector of $\mathbf{x}_m$, i.e., $\mathbf{x}_m = \left[e^{j\phi_m(1)}, e^{j\phi_m(2)}, ..., e^{j\phi_m(N)} \right]^{T}$. Then, the problem (16) can be reformulated as the following unconstrained problem with regard to $\{\boldsymbol{\phi}_m\}_{m=1}^{M}$:

$$\min_{\{\boldsymbol{\phi}_m\}_{m=1}^{M}} \quad J_T = \sum_{i=1}^{M} \mathbf{p}_{ii}^{H} \mathbf{Q} \mathbf{p}_{ii} + \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ij}^{H} \mathbf{Q}' \mathbf{p}_{ij}. \tag{17}$$

3. Problem Optimization via the Gradient Method

In this section, we optimize Problem (17) by using the gradient-based algorithm, which is able to guarantee that the objective function is monotonically decreasing at each iteration. In the following, we first derive the phase gradient and the step size and then briefly summarize the algorithm.

3.1. Phase Gradient

To facilitate the derivation, let $\phi_m(n) = \phi_{mn}$. Before deriving the phase gradient $\nabla_{\boldsymbol{\phi}_m} J_T$, we first deduce the derivative of J_T with respect to the phase ϕ_{mn}:

$$\frac{\partial J_T}{\partial \phi_{mn}} = \partial \left(\sum_{i=1}^{M} \mathbf{p}_{ii}^H \mathbf{Q} \mathbf{p}_{ii} + \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ij}^H \mathbf{Q}' \mathbf{p}_{ij} \right) \Big/ \partial \phi_{mn}$$

$$= \partial \left(\mathbf{p}_{mm}^H \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right) \Big/ \partial \phi_{mn} + \partial \left(\sum_{\substack{j=1 \\ j \neq m}}^{M} \mathbf{p}_{mj}^H \mathbf{Q}' \mathbf{p}_{mj} \sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{im}^H \mathbf{Q}' \mathbf{p}_{im} \right) \Big/ \partial \phi_{mn}. \tag{18}$$

Since $\sum\limits_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{im}^H \mathbf{Q}' \mathbf{p}_{im} = \sum\limits_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{mi}^H \mathbf{Q}'^* \mathbf{p}_{mi}$ (see Appendix B), (18) can be rewritten as:

$$\frac{\partial J_T}{\partial \phi_{mn}} = \partial \left(\mathbf{p}_{mm}^H \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right) \Big/ \partial \phi_{mn} + 2 \sum_{\substack{i=1 \\ i \neq m}}^{M} \partial \left(\mathbf{p}_{mi}^H \mathrm{Re}\left(\mathbf{Q}' \right) \mathbf{p}_{mi} \right) \Big/ \partial \phi_{mn}. \tag{19}$$

The first term in (19) can be simplified as:

$$\frac{\partial \left(\mathbf{p}_{mm}^H \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right)}{\partial \phi_{mn}} = \frac{\partial \left(\mathbf{p}_{mm}^T \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right)}{\partial \phi_{mn}}$$

$$= \frac{\partial \mathbf{p}_{mm}^T}{\partial \phi_{mn}} \left(\left(\mathbf{Q} + \mathbf{Q}' \right) + \left(\mathbf{Q} + \mathbf{Q}' \right)^T \right) \mathbf{p}_{mm} \tag{20}$$

$$= 2 \frac{\partial \mathbf{p}_{mm}^T}{\partial \phi_{mn}} \mathrm{Re}\left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm},$$

where the third equality follows from the fact $\left(\mathbf{Q} + \mathbf{Q}' \right)^T = \left(\mathbf{Q} + \mathbf{Q}' \right)^*$. Let:

$$\begin{aligned}
\mathbf{y}_{mm} &= \mathrm{Re}\left(\mathbf{Q} \right) \mathbf{p}_{mm} \\
\mathbf{y}'_{mi} &= \mathrm{Re}\left(\mathbf{Q}' \right) \mathbf{p}_{mi}, \; i = 1, ..., M,
\end{aligned} \tag{21}$$

then (20) can be expressed as:

$$\frac{\partial \left(\mathbf{p}_{mm}^H \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right)}{\partial \phi_{mn}} = 2 \frac{\partial \mathbf{p}_{mm}^T}{\partial \phi_{mn}} \left(\mathbf{y}_{mm} + \mathbf{y}'_{mm} \right)$$

$$= 2 \sum_{k=1}^{2N} \frac{\partial \left(p_{mm}(k) \right)}{\partial \phi_{mn}} \left(y_{mm}(k) + y'_{mm}(k) \right). \tag{22}$$

According to (A4) and (A5), we can deduce the derivative $\partial \left(p_{mm}(k) \right) / \partial \phi_{mn}$ as follows:

$$\frac{\partial \left(p_{mm}(k) \right)}{\partial \phi_{mn}} = \frac{\partial \left(f_m(k) f_m^*(k) \right)}{\partial \phi_{mn}} = 2 \mathrm{Re}\left(\frac{\partial f_m^*(k)}{\partial \phi_{mn}} f_m(k) \right) = 2 \mathrm{Re}\left(-j x_m^*(n) e^{j \frac{2\pi}{2N} nk} f_m(k) \right). \tag{23}$$

By substituting (23) into (22), we have:

$$\frac{\partial \left(\mathbf{p}_{mm}^H \left(\mathbf{Q} + \mathbf{Q}' \right) \mathbf{p}_{mm} \right)}{\partial \phi_{mn}} = 4 \mathrm{Re}\left(-j x_m^*(n) \sum_{k=1}^{2N} f_m(k) \left(y_{mm}(k) + y'_{mm}(k) \right) e^{j \frac{2\pi}{2N} nk} \right). \tag{24}$$

To compute the second term of (19), we first deduce $\partial \left(\mathbf{p}_{mi}^H \mathrm{Re}\left(\mathbf{Q}' \right) \mathbf{p}_{mi} \right) / \partial \phi_{mn} (i \neq m)$:

$$\frac{\partial \left(\mathbf{p}_{mi}^H \mathrm{Re}\left(\mathbf{Q}' \right) \mathbf{p}_{mi} \right)}{\partial \phi_{mn}} = 2 \mathrm{Re}\left(\frac{\partial \left(\mathbf{p}_{mi}^H \right)}{\partial \phi_{mn}} \mathrm{Re}\left(\mathbf{Q}' \right) \mathbf{p}_{mi} \right) = 2 \mathrm{Re}\left(\sum_{k=1}^{2N} \frac{\partial \left(p_{mi}^*(k) \right)}{\partial \phi_{mn}} y'_{mi}(k) \right). \tag{25}$$

Similarly to the derivation of (23), it is easy to obtain that:

$$\frac{\partial\left(p_{mi}(k)\right)}{\partial\phi_{mn}} = jx_m(n)e^{-j\frac{2\pi}{2N}nk}f_i^*(k), i \neq m. \tag{26}$$

On the basis of (25) and (26), $\partial\left(\mathbf{p}_{mi}^H\mathrm{Re}\left(\mathbf{Q}'\right)\mathbf{p}_{mi}\right)/\partial\phi_{mn}$ can be further denoted as

$$\frac{\partial\left(\mathbf{p}_{mi}^H\mathrm{Re}\left(\mathbf{Q}'\right)\mathbf{p}_{mi}\right)}{\partial\phi_{mn}} = 2\mathrm{Re}\left(-jx_m^*(n)\sum_{k=1}^{2N}f_i(k)y'_{mi}(k)e^{j\frac{2\pi}{2N}nk}\right). \tag{27}$$

By substituting (24) and (27) into (19), $\partial J_T/\partial\phi_{mn}$ can be simplified as:

$$\frac{\partial J_T}{\partial\phi_{mn}} = 4\mathrm{Re}\left(-jx_m^*(n)\sum_{k=1}^{2N}\left(f_m(k)y_{mm}(k) + \sum_{i=1}^{M}f_i(k)y'_{mi}(k)\right)e^{j\frac{2\pi}{2N}nk}\right). \tag{28}$$

Let $\mathbf{z}_m$ be the inverse FFT (IFFT) of $\mathbf{f}_m \circ \mathbf{y}_{mm} + \sum_{i=1}^{M}\mathbf{f}_i \circ \mathbf{y}'_{mi}$, i.e.,

$$\begin{aligned}\mathbf{z}_m &= \mathcal{F}^{-1}\left(\mathbf{f}_m \circ \mathbf{y}_{mm} + \sum_{i=1}^{M}\mathbf{f}_i \circ \mathbf{y}'_{mi}\right)\\ z_m(n) &= \frac{1}{2N}\sum_{k=1}^{2N}\left(f_m(k)y_{mm}(k) + \sum_{i=1}^{M}f_i(k)y'_{mi}(k)\right)e^{j\frac{2\pi}{2N}nk},\end{aligned} \tag{29}$$

then $\partial J_T/\partial\phi_{mn}$ in (28) becomes:

$$\frac{\partial J_T}{\partial\phi_{mn}} = 8N\mathrm{Re}\left(-jx_m^*(n)z_m(n)\right). \tag{30}$$

By stacking (30) in a vector, the phase gradient $\nabla_{\phi_m}J_T$ is given by:

$$\nabla_{\phi_m}J_T = \left[\partial J_T/\partial\phi_{m1}, ..., \partial J_T/\partial\phi_{mN}\right]^T = 8N\mathrm{Re}\left(-j\mathbf{x}_m^* \circ \mathbf{z}_m(1:N)\right), \tag{31}$$

where $\mathbf{z}_m(1:N)$ denotes the first N elements of $\mathbf{z}_m$. It is worth noting that $\mathbf{y}_{mm}$ and $\mathbf{y}'_{mi}$ defined in (21) can be calculated by the Hadamard product and the FFT operations:

$$\begin{aligned}\mathbf{y}_{mm} &= \lambda\mathbf{D}_f\mathbf{p}_{mm} = \lambda\mathbf{w}_t \circ \mathbf{p}_{mm},\\ \mathbf{y}'_{mi} &= \mathrm{Re}\left(\frac{1}{4N^2}\left(1-\lambda\right)\mathbf{F}\mathbf{D}_t\mathbf{F}^H\right)\mathbf{p}_{mi}\\ &= \frac{1}{8N^2}\left(1-\lambda\right)\left(\mathbf{F}\mathbf{D}_t\mathbf{F}^H + \mathbf{F}^H\mathbf{D}_t\mathbf{F}\right)\mathbf{p}_{mi}\\ &= \frac{1}{4N}\left(1-\lambda\right)\left(\mathcal{F}\left(\mathbf{w}_t \circ \mathcal{F}^{-1}\left(\mathbf{p}_{mi}\right)\right) + \mathcal{F}^{-1}\left(\mathbf{w}_t \circ \mathcal{F}\left(\mathbf{p}_{mi}\right)\right)\right).\end{aligned} \tag{32}$$

From the derivation above, it is easy to see that the calculation of the phase gradient is a little bit cumbersome. In order to make the calculation process concise, we write the gradient in the form of matrix. Let $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_M]$ and $\mathbf{\Phi} = [\phi_1, ..., \phi_M]$ denote the matrix of waveform set and the corresponding phase matrix, respectively. Then, the spectrum matrix of $\mathbf{X}$ is $\mathbf{S} = \mathcal{F}_c(\mathbf{X}) = [\mathbf{f}_1, ..., \mathbf{f}_M]$. Define:

$$\begin{aligned}\mathbf{Y} &= [\mathbf{y}_{11}, ..., \mathbf{y}_{MM}] = \lambda\left(\mathbf{w}_f\mathbf{1}_M^T\right) \circ \mathbf{S} \circ \mathbf{S}^*,\\ \mathbf{Y}_1 &= \left[\sum_{i=1}^{M}\mathbf{f}_i \circ \mathbf{y}'_{1i}, ..., \sum_{i=1}^{M}\mathbf{f}_i \circ \mathbf{y}'_{Mi}\right],\end{aligned} \tag{33}$$

then the matrix form of (29) is given by:

$$\mathbf{Z} = [\mathbf{z}_1, ..., \mathbf{z}_M] = \mathcal{F}_c^{-1}\left(\mathbf{S} \circ \mathbf{Y} + \mathbf{Y}_1\right). \tag{34}$$

Thus, according to (31), the matrix of phase gradient $\mathbf{G}$ can be expressed as:

$$\mathbf{G} = \left[\nabla_{\phi_1} J_T, ..., \nabla_{\phi_M} J_T\right] = 8N\mathrm{Re}\left(-j\mathbf{X}^* \circ \mathbf{Z}(1:N,:)\right), \tag{35}$$

where $\mathbf{Z}(1:N,:)$ denotes the submatrix formed with the first N rows of $\mathbf{Z}$. By defining $\mathbf{Y}$ and $\mathbf{Y}_1$, we can easily calculate the gradient matrix. However, it is still hard to calculate $\mathbf{Y}_1$ directly by (33). To efficiently calculate $\mathbf{Y}_1$, define:

$$\mathbf{Y}' = \begin{bmatrix} \mathbf{y}'_{11} & \cdots & \mathbf{y}'_{1M} \\ \vdots & \ddots & \vdots \\ \mathbf{y}'_{M1} & \cdots & \mathbf{y}'_{MM} \end{bmatrix} \in \mathbb{C}^{2NM \times M}, \mathbf{S}_{rep} = \begin{bmatrix} \mathbf{S} \\ \vdots \\ \mathbf{S} \end{bmatrix}_{2NM \times M}. \tag{36}$$

Then it is easy to verify that:

$$\left(\mathbf{S}_{rep} \circ \mathbf{Y}'\right) \mathbf{1}_M = \mathrm{vec}\left(\mathbf{Y}_1\right), \tag{37}$$

where $\mathrm{vec}\left(\mathbf{Y}_1\right)$ denotes the column vector consisting of all of the columns of $\mathbf{Y}_1$. Thus, we can obtain $\mathbf{Y}_1$ via the inverse operation of $\mathrm{vec}\left(\cdot\right)$, i.e.,

$$\mathbf{Y}_1 = \mathrm{vec}^{-1}\left(\left(\mathbf{S}_{rep} \circ \mathbf{Y}'\right) \mathbf{1}_M\right), \tag{38}$$

where $\mathbf{Y}'$ can be calculated by (32). It is easy to see that $\mathbf{y}'_{mi}$ can be implemented by four FFT (IFFT) operations. Since $\mathbf{y}'_{mi} = \mathbf{y}'^*_{im}$, the calculation of $\mathbf{Y}'$ takes $4\frac{M(M+1)}{2} = 2M^2 + 2M$ FFT (IFFT) operations.

3.2. Step Size Calculation via Taylor Series Expansion

The traditional methods for obtaining the step size are the linear search methods, which require many iterations and thus are quite time consuming. To reduce the computing expense, here we propose to calculate the step size directly. Assume that $\mathbf{\Phi}^l$ is the phase matrix of the present iteration point $\mathbf{X}^l = \left[\mathbf{x}_1^l, ..., \mathbf{x}_M^l\right] = e^{j\mathbf{\Phi}^l}$, and $\mathbf{D}^l = \left[\mathbf{d}_1^l, ..., \mathbf{d}_M^l\right]$ is the descent direction. Then, the new iteration point can be denoted as:

$$\begin{aligned} \mathbf{\Phi}^{l+1} &= \mathbf{\Phi}^l + \mu\mathbf{D}^l, \\ \mathbf{X}^{l+1} &= e^{j\mathbf{\Phi}^{l+1}} = \mathbf{X}^l \circ e^{j\mu\mathbf{D}^l}, \end{aligned} \tag{39}$$

i.e.,

$$\mathbf{x}_i^{l+1} = \mathbf{x}_i^l \circ e^{j\mu\mathbf{d}_i^l}, \ i = 1, ..., M, \tag{40}$$

where μ is the step size. Thus, the linear search problem can be formulated as the following minimization problem:

$$\min_{\mu>0} \ h(\mu) = J_T\left(\mathbf{X}^{l+1}\right) = \sum_{i=1}^{M}\left(\mathbf{p}_{ii}^{l+1}\right)^H \mathbf{Q}\mathbf{p}_{ii}^{l+1} + \sum_{i=1}^{M}\sum_{j=1}^{M}\left(\mathbf{p}_{ij}^{l+1}\right)^H \mathbf{Q}'\mathbf{p}_{ij}^{l+1}. \tag{41}$$

By taking the derivative of (41), we have:

$$\frac{\partial h}{\partial \mu} = 2\mathrm{Re}\left(\sum_{i=1}^{M}\frac{\partial\left(\mathbf{p}_{ii}^{l+1}\right)^H}{\partial \mu}\mathbf{Q}\mathbf{p}_{ii}^{l+1} + \sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\partial\left(\mathbf{p}_{ij}^{l+1}\right)^H}{\partial \mu}\mathbf{Q}'\mathbf{p}_{ij}^{l+1}\right). \tag{42}$$

To simplify (41) and (42), we first deduce $\mathbf{p}_{ij}^{l+1}$. By using the Taylor series expansion, $\mathbf{p}_{ij}^{l+1}$ can be approximated as (see Appendix C):

$$\mathbf{p}_{ij}^{l+1} \approx \mathbf{f}_i \circ \mathbf{f}_j^* + j\left(\mathbf{f}'_i \circ \mathbf{f}_j^* - \mathbf{f}_i \circ \mathbf{f}'^*_j\right)\mu - \frac{1}{2}\left(\mathbf{f}''_i \circ \mathbf{f}_j^* + \mathbf{f}_i \circ \mathbf{f}''^*_j - 2\mathbf{f}'_i \circ \mathbf{f}'^*_j\right)\mu^2, \tag{43}$$

where:

$$\mathbf{f}_i = \mathcal{F}\left(\mathbf{x}_i^l\right), \mathbf{f}'_i = \mathcal{F}\left(\mathbf{x}_i^l \circ \mathbf{d}_i^l\right), \mathbf{f}''_i = \mathcal{F}\left(\mathbf{x}_i^l \circ \mathbf{d}_i^l \circ \mathbf{d}_i^l\right). \tag{44}$$

Let

$$\begin{aligned}
\mathbf{p}_{ij}^l &= \mathbf{f}_i \circ \mathbf{f}_j^*, \\
\mathbf{c}_{ij} &= j\left(\mathbf{f}'_i \circ \mathbf{f}_j^* - \mathbf{f}_i \circ \mathbf{f}'^*_j\right), \\
\mathbf{c}'_{ij} &= -\frac{1}{2}\left(\mathbf{f}''_i \circ \mathbf{f}_j^* + \mathbf{f}_i \circ \mathbf{f}''^*_j - 2\mathbf{f}'_i \circ \mathbf{f}'^*_j\right),
\end{aligned} \tag{45}$$

then (43) can be rewritten as:

$$\mathbf{p}_{ij}^{l+1} \approx \mathbf{p}_{ij}^l + \mathbf{c}_{ij}\mu + \mathbf{c}'_{ij}\mu^2 = \bar{\mathbf{p}}_{ij}^{l+1}. \tag{46}$$

By replacing $\mathbf{p}_{ij}^{l+1}$ with $\bar{\mathbf{p}}_{ij}^{l+1}$, the approximate function of $h(\mu)$ in (41) can be denoted as:

$$h_1(\mu) = \sum_{i=1}^M \left(\bar{\mathbf{p}}_{ii}^{l+1}\right)^H \mathbf{Q} \bar{\mathbf{p}}_{ii}^{l+1} + \sum_{i=1}^M \sum_{j=1}^M \left(\bar{\mathbf{p}}_{ij}^{l+1}\right)^H \mathbf{Q}' \bar{\mathbf{p}}_{ij}^{l+1}. \tag{47}$$

Since $\mathbf{p}_{ij}^{l+1} = \bar{\mathbf{p}}_{ij}^{l+1}$ holds when $\mu = 0$, it is easy to verify that:

$$h(0) = h_1(0), \left.\frac{\partial h}{\partial \mu}\right|_{\mu=0} = \left.\frac{\partial h_1}{\partial \mu}\right|_{\mu=0}, \tag{48}$$

which indicates that $h(\mu)$ and $h_1(\mu)$ have the same function value and slope at $\mu = 0$. As $\mathbf{D}^l$ is the descent direction, the slopes of these two functions are less than zero. Thus, these two functions have at least one minimum point greater than zero. Since $h(\mu)$ is sensitive to the waveform phases, the optimal step size of $h(\mu)$ is very small and close to zero. Consequently, we can use the minimum point of $h_1(\mu)$ to approximate the optimal step size.

To calculate the minimum point of $h_1(\mu)$, we replace $\mathbf{p}_{ij}^{l+1}$ in (42) with $\bar{\mathbf{p}}_{ij}^{l+1}$, then the derivative of $h_1(\mu)$ with respective to μ is given by:

$$\begin{aligned}
\frac{\partial h_1}{\partial \mu} &= 2\mathrm{Re}\left(\sum_{i=1}^M \left(\mathbf{c}_{ii}^H + 2\mathbf{c}'^H_{ii}\mu\right)\mathbf{Q}\left(\mathbf{p}_{ii}^l + \mathbf{c}_{ii}\mu + \mathbf{c}'_{ii}\mu^2\right)\right. \\
&\quad \left. + \sum_{i=1}^M \sum_{j=1}^M \left(\mathbf{c}_{ij}^H + 2\mathbf{c}'^H_{ij}\mu\right)\mathbf{Q}'\left(\mathbf{p}_{ij}^l + \mathbf{c}_{ij}\mu + \mathbf{c}'_{ij}\mu^2\right)\right) \\
&= a\mu^3 + b\mu^2 + c\mu + d,
\end{aligned} \tag{49}$$

where:

$$a = 2\text{Re}\left(2\sum_{i=1}^{M} \mathbf{c}'^{H}_{ii}\mathbf{Q}\mathbf{c}'_{ii} + 2\sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{c}'^{H}_{ij}\mathbf{Q}'\mathbf{c}'_{ij}\right),$$

$$b = 2\text{Re}\left(\sum_{i=1}^{M}\left(2\mathbf{c}'^{H}_{ii}\mathbf{Q}\mathbf{c}_{ii} + \mathbf{c}^{H}_{ii}\mathbf{Q}\mathbf{c}'_{ii}\right) + \sum_{i=1}^{M}\sum_{j=1}^{M}\left(2\mathbf{c}'^{H}_{ij}\mathbf{Q}'\mathbf{c}_{ij} + \mathbf{c}^{H}_{ij}\mathbf{Q}'\mathbf{c}'_{ij}\right)\right),$$

$$c = 2\text{Re}\left(\sum_{i=1}^{M}\left(2\mathbf{c}'^{H}_{ii}\mathbf{Q}\mathbf{p}^{l}_{ii} + \mathbf{c}^{H}_{ii}\mathbf{Q}\mathbf{c}_{ii}\right) + \sum_{i=1}^{M}\sum_{j=1}^{M}\left(2\mathbf{c}'^{H}_{ij}\mathbf{Q}'\mathbf{p}^{l}_{ij} + \mathbf{c}^{H}_{ij}\mathbf{Q}'\mathbf{c}_{ij}\right)\right),$$

$$d = 2\text{Re}\left(\sum_{i=1}^{M}\mathbf{c}^{H}_{ii}\mathbf{Q}\mathbf{p}^{l}_{ii} + \sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{c}^{H}_{ij}\mathbf{Q}'\mathbf{p}^{l}_{ij}\right). \tag{50}$$

To simplify the calculation, let:

$$\mathbf{h}_{ij} = \mathcal{F}^{-1}\left(\mathbf{c}_{ij}\right) = \frac{1}{2N}\mathbf{F}^{H}\mathbf{c}_{ij},$$

$$\mathbf{h}'_{ij} = \mathcal{F}^{-1}\left(\mathbf{c}'_{ij}\right) = \frac{1}{2N}\mathbf{F}^{H}\mathbf{c}'_{ij}. \tag{51}$$

Then (50) can be rewritten as:

$$a = 4\text{Re}\left(\lambda\mathbf{w}_{f}^{T}\sum_{i=1}^{M}\left(\mathbf{c}'^{*}_{ii}\circ\mathbf{c}'_{ii}\right) + (1-\lambda)\,\mathbf{w}_{t}^{T}\sum_{i=1}^{M}\sum_{j=1}^{M}\left(\mathbf{h}'^{*}_{ij}\circ\mathbf{h}'_{ij}\right)\right),$$

$$b = 2\text{Re}\left(\lambda\mathbf{w}_{f}^{T}\sum_{i=1}^{M}\left(2\mathbf{c}'^{*}_{ii}\circ\mathbf{c}_{ii} + \mathbf{c}^{*}_{ii}\circ\mathbf{c}'_{ii}\right) + (1-\lambda)\,\mathbf{w}_{t}^{T}\sum_{i=1}^{M}\sum_{j=1}^{M}\left(2\mathbf{h}'^{*}_{ij}\circ\mathbf{h}_{ij} + \mathbf{h}^{*}_{ij}\circ\mathbf{h}'_{ij}\right)\right),$$

$$c = 2\text{Re}\left(\lambda\mathbf{w}_{f}^{T}\sum_{i=1}^{M}\left(2\mathbf{c}'^{*}_{ii}\circ\mathbf{p}^{l}_{ii} + \mathbf{c}^{*}_{ii}\circ\mathbf{c}_{ii}\right) + (1-\lambda)\,\mathbf{w}_{t}^{T}\sum_{i=1}^{M}\sum_{j=1}^{M}\left(2\mathbf{h}'^{*}_{ij}\circ\mathbf{r}^{l}_{ij} + \mathbf{h}^{*}_{ij}\circ\mathbf{h}_{ij}\right)\right),$$

$$d = 2\text{Re}\left(\lambda\mathbf{w}_{f}^{T}\sum_{i=1}^{M}\left(\mathbf{c}^{*}_{ii}\circ\mathbf{p}^{l}_{ii}\right) + (1-\lambda)\,\mathbf{w}_{t}^{T}\sum_{i=1}^{M}\sum_{j=1}^{M}\left(\mathbf{h}^{*}_{ij}\circ\mathbf{r}^{l}_{ij}\right)\right), \tag{52}$$

where the first term of each equality follows from the fact $\mathbf{a}^{H}\mathbf{Q}\mathbf{b} = \lambda\mathbf{a}^{H}\mathbf{D}_{f}\mathbf{b} = \mathbf{w}_{f}^{T}(\mathbf{a}^{*}\circ\mathbf{b})$ ($\mathbf{a}, \mathbf{b}$ are the arbitrary vectors), and the second term of each equality follows from the fact $\mathbf{a}^{H}\mathbf{Q}'\mathbf{b} = \frac{1}{4N^{2}}(1-\lambda)\,\mathbf{a}^{H}\mathbf{F}\mathbf{D}_{t}\mathbf{F}^{H}\mathbf{b} = (1-\lambda)\,\mathbf{w}_{t}^{T}\left((\mathcal{F}^{-1}(\mathbf{a}))^{*}\circ\mathcal{F}^{-1}(\mathbf{b})\right)$. Let $\frac{\partial h_{1}}{\partial\mu} = 0$, then the minimum point of $h_{1}(\mu)$ can be obtained by solving the following cubic equation:

$$a\mu^{3} + b\mu^{2} + c\mu + d = 0 \tag{53}$$

It is well known that a cubic equation with real coefficients has three roots, in which there is at least a real root. Therefore, we can choose the positive root that is closest to zero as the precise estimation of the step size.

In order to facilitate the calculation, we write the above derivation in the form of matrix. Define:

$$\mathbf{S} = [\mathbf{f}_{1},...,\mathbf{f}_{M}]\,,\mathbf{S}' = [\mathbf{f}'_{1},...,\mathbf{f}'_{M}]\,,\ \mathbf{S}'' = [\mathbf{f}''_{1},...,\mathbf{f}''_{M}]\,, \tag{54}$$

then according to (44) we have:

$$\mathbf{S} = \mathcal{F}_{c}\left(\mathbf{X}^{l}\right), \mathbf{S}' = \mathcal{F}_{c}\left(\mathbf{X}^{l}\circ\mathbf{D}^{l}\right), \mathbf{S}'' = \mathcal{F}_{c}\left(\mathbf{X}^{l}\circ\mathbf{D}^{l}\circ\mathbf{D}^{l}\right). \tag{55}$$

Let:

$$
\begin{aligned}
\mathbf{P} &= \left[\mathbf{p}^l_{11}, ..., \mathbf{p}^l_{M1}, \mathbf{p}^l_{12}, ..., \mathbf{p}^l_{M2}, ..., \mathbf{p}^l_{1M}, ..., \mathbf{p}^l_{MM} \right], \\
\mathbf{C} &= \left[\mathbf{c}_{11}, ..., \mathbf{c}_{M1}, \mathbf{c}_{12}, ..., \mathbf{c}_{M2}, ..., \mathbf{c}_{1M}, ..., \mathbf{c}_{MM} \right], \\
\mathbf{C}' &= \left[\mathbf{c}'_{11}, ..., \mathbf{c}'_{M1}, \mathbf{c}'_{12}, ..., \mathbf{c}'_{M2}, ..., \mathbf{c}'_{1M}, ..., \mathbf{c}'_{MM} \right], \\
\mathbf{P}_d &= \left[\mathbf{p}^l_{11}, \mathbf{p}^l_{22}, ..., \mathbf{p}^l_{MM} \right], \\
\mathbf{C}_d &= \left[\mathbf{c}_{11}, \mathbf{c}_{22}, ..., \mathbf{c}_{MM} \right], \\
\mathbf{C}'_d &= \left[\mathbf{c}'_{11}, \mathbf{c}'_{22}, ..., \mathbf{c}'_{MM} \right],
\end{aligned}
\tag{56}
$$

Then the inverse Fourier transforms of $\mathbf{P}, \mathbf{C}, \mathbf{C}'$ can be denoted as:

$$
\mathbf{R} = \mathcal{F}_c^{-1}\left(\mathbf{P} \right), \mathbf{H} = \mathcal{F}_c^{-1}\left(\mathbf{C} \right), \mathbf{H}' = \mathcal{F}_c^{-1}\left(\mathbf{C}' \right).
\tag{57}
$$

According to (52), (56) and (57), the coefficients of the cubic equation can be reformulated as:

$$
\begin{aligned}
a &= 4\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(\mathbf{C}'^*_d \circ \mathbf{C}'_d \right) \mathbf{1}_M + (1-\lambda)\, \mathbf{w}_t^T \left(\mathbf{H}'^* \circ \mathbf{H}' \right) \mathbf{1}_{M^2} \right), \\
b &= 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(2\mathbf{C}'^*_d \circ \mathbf{C}_d + \mathbf{C}^*_d \circ \mathbf{C}'_d \right) \mathbf{1}_M + (1-\lambda)\, \mathbf{w}_t^T \left(2\mathbf{H}'^* \circ \mathbf{H} + \mathbf{H}^* \circ \mathbf{H}' \right) \mathbf{1}_{M^2} \right), \\
c &= 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(2\mathbf{C}'^*_d \circ \mathbf{P}_d + \mathbf{C}^*_d \circ \mathbf{C}_d \right) \mathbf{1}_M + (1-\lambda)\, \mathbf{w}_t^T \left(2\mathbf{H}'^* \circ \mathbf{R} + \mathbf{H}^* \circ \mathbf{H} \right) \mathbf{1}_{M^2} \right), \\
d &= 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(\mathbf{C}^*_d \circ \mathbf{P}_d \right) \mathbf{1}_M + (1-\lambda)\, \mathbf{w}_t^T \left(\mathbf{H}^* \circ \mathbf{R} \right) \mathbf{1}_{M^2} \right).
\end{aligned}
\tag{58}
$$

3.3. Algorithm Summary

After deducing the phase gradient and the step size, it is easy to solve the unconstrained problem (17) by using the conjugate gradient algorithm (CGA). Here, we apply the classical Polak–Ribiere–Polyak CGA (PRP-CGA) to deal with the waveform design problem. The searching direction of the PRP-CGA can be expressed as:

$$
\mathbf{d}^{l+1} = -\mathbf{g}^{l+1} + \frac{\left(\mathbf{g}^{l+1} - \mathbf{g}^l \right)^T \mathbf{g}^{l+1}}{\left\| \mathbf{g}^l \right\|^2} \mathbf{d}^l,
\tag{59}
$$

where $\mathbf{g}^l$ and $\mathbf{d}^l$ are the gradient vector and the direction vector, respectively. For the problem here, $\mathbf{g}^l$ and $\mathbf{d}^l$ are defined as:

$$
\mathbf{g}^l = \left[\left(\mathbf{g}^l_1 \right)^T, ..., \left(\mathbf{g}^l_M \right)^T \right]^T, \mathbf{d}^l = \left[\left(\mathbf{d}^l_1 \right)^T, ..., \left(\mathbf{d}^l_M \right)^T \right]^T,
\tag{60}
$$

where $\mathbf{g}^l_m = \nabla_{\phi_m} J_T\left(\mathbf{X}^l \right), m = 1, ..., M$. Since the step size is an approximate value, $\left(\mathbf{g}^{l+1} \right)^T \mathbf{d}^l$ is not equal to zero. Thus, $\mathbf{d}^{l+1}$ may not be descendent, i.e.,

$$
\left(\mathbf{g}^{l+1} \right)^T \mathbf{d}^{l+1} = -\left\| \mathbf{g}^{l+1} \right\|^2 + \frac{\left(\mathbf{g}^{l+1} - \mathbf{g}^l \right)^T \mathbf{g}^{l+1}}{\left\| \mathbf{g}^l \right\|^2} \left(\mathbf{g}^{l+1} \right)^T \mathbf{d}^l < 0
\tag{61}
$$

is not always satisfied. For guaranteeing that the searching direction is descendant, we adopt the following modified direction:

$$
\tilde{\mathbf{d}}^{l+1} = -\mathbf{g}^{l+1} + \frac{\left(\mathbf{g}^{l+1} - \mathbf{g}^l\right)^T \mathbf{g}^{l+1}}{\left\|\mathbf{g}^l\right\|^2} \mathbf{d}^l,
$$

$$
\mathbf{d}^{l+1} = \begin{cases} \tilde{\mathbf{d}}^{l+1}, & \left(\mathbf{g}^{l+1}\right)^T \tilde{\mathbf{d}}^{l+1} < 0 \\ -\mathbf{g}^{l+1}, & \left(\mathbf{g}^{l+1}\right)^T \tilde{\mathbf{d}}^{l+1} \geq 0 \end{cases}.
$$

$$(62)$$

Actually, (59) can also be expressed as:

$$
\mathbf{D}^{l+1} = -\mathbf{G}^{l+1} + \frac{\left(\mathbf{g}^{l+1} - \mathbf{g}^l\right)^T \mathbf{g}^{l+1}}{\left\|\mathbf{g}^l\right\|^2} \mathbf{D}^l.
\tag{63}
$$

Since

$$
\left\|\mathbf{g}^l\right\|^2 = \sum_{m=1}^M \left\|\mathbf{g}_m^l\right\|^2 = \left\|\mathbf{G}^l\right\|_F^2,
$$

$$
\left(\mathbf{g}^{l+1} - \mathbf{g}^l\right)^T \mathbf{g}^{l+1} = \left\|\mathbf{G}^{l+1}\right\|_F^2 - \mathbf{1}_N^T \left(\mathbf{G}^{l+1} \circ \mathbf{G}^l\right) \mathbf{1}_M,
$$

$$(64)$$

and (63) can be rewritten as:

$$
\mathbf{D}^{l+1} = -\mathbf{G}^{l+1} + \frac{\left\|\mathbf{G}^{l+1}\right\|_F^2 - \mathbf{1}_N^T \left(\mathbf{G}^{l+1} \circ \mathbf{G}^l\right) \mathbf{1}_M}{\left\|\mathbf{G}^l\right\|_F^2} \mathbf{D}^l.
\tag{65}
$$

Thus, (62) can be expressed as the following matrix form:

$$
\tilde{\mathbf{D}}^{l+1} = -\mathbf{G}^{l+1} + \frac{\left\|\mathbf{G}^{l+1}\right\|_F^2 - \mathbf{1}_N^T \left(\mathbf{G}^{l+1} \circ \mathbf{G}^l\right) \mathbf{1}_M}{\left\|\mathbf{G}^l\right\|_F^2} \mathbf{D}^l,
$$

$$
\mathbf{D}^{l+1} = \begin{cases} \tilde{\mathbf{D}}^{l+1}, & \mathbf{1}_N^T \left(\mathbf{G}^{l+1} \circ \tilde{\mathbf{D}}^{l+1}\right) \mathbf{1}_M < 0 \\ -\mathbf{G}^{l+1}, & \mathbf{1}_N^T \left(\mathbf{G}^{l+1} \circ \tilde{\mathbf{D}}^{l+1}\right) \mathbf{1}_M \geq 0 \end{cases}.
$$

$$(66)$$

On the basis of the above derivation, the gradient-based algorithm, which we call Gra-WeCorr-SFW, is summarized in Algorithm 1.

4. Simplified Algorithm for the Design Problem without Considering the Correlation Weights

In Section 3, we present a gradient-based algorithm to handle Problem (17). From (9), we can see that the criterion J_{CF} can be simplified when $w_t(k) = 1, k = 1 - N, ..., N - 1$ (i.e., the correlation weights are not taken into account). Thus, in this section, we derive a simplified algorithm for the design problem without considering the correlation weights. Let $w_t(k) = 1, k = 1 - N, ..., N - 1$, then the criterion J_{CF} can be simplified as:

$$
J_{CF} = \frac{1}{2N} \sum_{i=1}^M \sum_{j=1}^M \mathbf{p}_{ij}^H \mathbf{p}_{ij} = \frac{1}{2N} \sum_{i=1}^M \sum_{j=1}^M \mathbf{p}_{ii}^H \mathbf{p}_{jj} = \frac{1}{2N} \sum_{i=1}^M \mathbf{p}_{ii}^H \sum_{j=1}^M \mathbf{p}_{jj}.
\tag{67}
$$

Thus we can rewrite the problem (17) as:

$$J_T = \lambda \sum_{i=1}^{M} \mathbf{p}_{ii}^H \mathbf{D}_f \mathbf{p}_{ii} + \lambda_1 \sum_{i=1}^{M} \mathbf{p}_{ii}^H \sum_{j=1}^{M} \mathbf{p}_{jj}, \tag{68}$$

where $\lambda_1 = \frac{1-\lambda}{2N}$. Like the derivation in Section 3, we deduce the gradient and the step size.

Algorithm 1: Gra-WeCorr-SFW.

Initialization: $l = 0, \lambda, M, N, \mathbf{w}_t, \mathbf{w}_f, \mathbf{X}^0 = \left[\mathbf{x}_1^0, ..., \mathbf{x}_M^0\right]$,
$\qquad \mathbf{S} = \mathcal{F}_c\left(\mathbf{X}^0\right), \mathbf{G}^0, \mathbf{D}^0 = -\mathbf{G}^0$.

Repeat

1: $\mathbf{S}' = \mathcal{F}_c\left(\mathbf{X}^l \circ \mathbf{D}^l\right), \mathbf{S}'' = \mathcal{F}_c\left(\mathbf{X}^l \circ \mathbf{D}^l \circ \mathbf{D}^l\right)$.

2: Compute $\mathbf{P}, \mathbf{C}, \mathbf{C}', \mathbf{P}_d, \mathbf{C}_d, \mathbf{C}'_d$ according to (45).

3: $\mathbf{R} = \mathcal{F}_c^{-1}\left(\mathbf{P}\right), \mathbf{H} = \mathcal{F}_c^{-1}\left(\mathbf{C}\right), \mathbf{H}' = \mathcal{F}_c^{-1}\left(\mathbf{C}'\right)$.

4: Compute coefficients a, b, c, d according to (58).

5: Solve the cubic equation (53), and then choose the
$\qquad$ positive root that is closest to zero as the step size μ^l.

6: $\mathbf{X}^{l+1} = \mathbf{X}^l \circ e^{j\mu^l \mathbf{D}^l}, \mathbf{S} = \mathcal{F}_c(\mathbf{X}^{l+1})$.

7: $\mathbf{Y} = \lambda \left(\mathbf{w}_f \mathbf{1}_M^T\right) \circ \mathbf{S} \circ \mathbf{S}^*$.

8: Compute $\mathbf{Y}'$ according to (32).

9: $\mathbf{S}_{rep} = \left[\mathbf{S}^T, ..., \mathbf{S}^T\right]_{2NM \times M}^T, \mathbf{Y}_1 = \mathrm{vec}^{-1}\left(\left(\mathbf{S}_{rep} \circ \mathbf{Y}'\right)\mathbf{1}_M\right)$.

10: $\mathbf{Z} = \mathcal{F}_c^{-1}\left(\mathbf{S} \circ \mathbf{Y} + \mathbf{Y}_1\right)$.

11: $\mathbf{G}^{l+1} = 8N\mathrm{Re}\left(-j\left(\mathbf{X}^{l+1}\right)^* \circ \mathbf{Z}(1:N,:)\right)$.

12: Compute the searching direction $\mathbf{D}^{l+1}$ according to (66).

13: $l = l + 1$.

Until convergence

4.1. Phase Gradient

According to (68), we deduce the derivative $\partial J_T / \partial \phi_{mn}$ as:

$$\frac{\partial J_T}{\partial \phi_{mn}} = 2\mathrm{Re}\left(\frac{\partial \mathbf{p}_{mm}^H}{\partial \phi_{mn}}\left(\lambda \mathbf{w}_f \circ \mathbf{p}_{mm} + \lambda_1 \sum_{j=1}^{M} \mathbf{p}_{jj}\right)\right). \tag{69}$$

Let:

$$\mathbf{y}''_m = \lambda \mathbf{w}_f \circ \mathbf{p}_{mm} + \lambda_1 \sum_{j=1}^{M} \mathbf{p}_{jj}, \tag{70}$$

then according to (23), (69) can be rewritten as:

$$\frac{\partial J_T}{\partial \phi_{mn}} = 2\mathrm{Re}\left(\sum_{k=1}^{2N} \frac{\partial\left(p_{mm}^*(k)\right)}{\partial \phi_{mn}} y''_m(k)\right) = 4\mathrm{Re}\left(-jx_{mm}^* \sum_{k=1}^{2N} f_m(k) y''_m(k) e^{j\frac{2\pi}{2N}nk}\right). \tag{71}$$

By stacking (71) in a vector, the gradient is given by:

$$\nabla_{\phi_m} J_T = 8N\mathrm{Re}\left(-j\mathbf{x}_m^* \circ \mathbf{z}'_m\left(1:N\right)\right), \tag{72}$$

where $\mathbf{z}'_m = \mathcal{F}^{-1}(\mathbf{f}_m \circ \mathbf{y}''_m)$. Define $\mathbf{Y}_2 = [\mathbf{y}''_1, ..., \mathbf{y}''_M]$, $\mathbf{Z}_1 = [\mathbf{z}'_1, ..., \mathbf{z}'_M]$, then it is easy to verify that:

$$\mathbf{Y}_2 = \lambda \left(\mathbf{w}_f \mathbf{1}_M^T\right) \circ \mathbf{S} \circ \mathbf{S}^* + \lambda_1 \left(\mathbf{S} \circ \mathbf{S}^*\right) \mathbf{1}_M \mathbf{1}_M^T,$$
$$\mathbf{Z}_1 = \mathcal{F}_c^{-1} \left(\mathbf{S} \circ \mathbf{Y}_2\right). \tag{73}$$

Thus, the gradient matrix can be expressed as:

$$\mathbf{G} = 8N\mathrm{Re}\left(-j\mathbf{X}^* \circ \mathbf{Z}_1 \left(1:N,:\right)\right). \tag{74}$$

4.2. Step Size

Similarly to (41), the linear search problem here can be denoted as:

$$\min_{\mu > 0} h(\mu) = \lambda \sum_{i=1}^{M} \left(\mathbf{p}_{ii}^{l+1}\right)^H \mathbf{D}_f \mathbf{p}_{ii}^{l+1} + \lambda_1 \sum_{i=1}^{M} \left(\mathbf{p}_{ii}^{l+1}\right)^H \sum_{j=1}^{M} \mathbf{p}_{jj}^{l+1}. \tag{75}$$

By taking the derivative of $h(\mu)$, we have:

$$\frac{\partial h}{\partial \mu} = 2\mathrm{Re}\left(\lambda \sum_{i=1}^{M} \left(\frac{\partial \mathbf{p}_{ii}^{l+1}}{\partial \mu}\right)^H \mathbf{D}_f \mathbf{p}_{ii}^{l+1} + \lambda_1 \left(\sum_{i=1}^{M} \frac{\partial \mathbf{p}_{ii}^{l+1}}{\partial \mu}\right)^H \sum_{j=1}^{M} \mathbf{p}_{jj}^{l+1}\right). \tag{76}$$

Then, the approximate derivative can be obtained by replacing $\mathbf{p}_{pq}^{l+1}$ with $\bar{\mathbf{p}}_{pq}^{l+1}$ defined in (46):

$$\frac{\partial h}{\partial \mu} \approx 2\mathrm{Re}\left(\lambda \sum_{i=1}^{M} \left(\mathbf{c}_{ii} + 2\mathbf{c}'_{ii}\mu\right)^H \mathbf{D}_f \left(\mathbf{p}_{ii}^l + \mathbf{c}_{ii}\mu + \mathbf{c}'_{ii}\mu^2\right)\right.$$
$$\left. + \lambda_1 \sum_{i=1}^{M} \left(\mathbf{c}_{ii} + 2\mathbf{c}'_{ii}\mu\right)^H \sum_{j=1}^{M} \left(\mathbf{p}_{jj}^l + \mathbf{c}_{jj}\mu + \mathbf{c}'_{jj}\mu^2\right)\right). \tag{77}$$

According the definition (56), we have:

$$\sum_{i=1}^{M} \mathbf{p}_{ii}^l = \mathbf{P}_d \mathbf{1}_M, \sum_{i=1}^{M} \mathbf{c}_{ii} = \mathbf{C}_d \mathbf{1}_M, \sum_{i=1}^{M} \mathbf{c}'_{ii} = \mathbf{C}'_d \mathbf{1}_M, \tag{78}$$

where $\mathbf{P}_d$, $\mathbf{C}_d$ and $\mathbf{C}'_d$ can be expressed as the following equalities according to (45):

$$\mathbf{P}_d = \mathbf{S} \circ \mathbf{S}^*, \mathbf{C}_d = -2\mathrm{Im}\left(\mathbf{S}' \circ \mathbf{S}^*\right), \mathbf{C}'_d = -\left(\mathrm{Re}\left(\mathbf{S}'' \circ \mathbf{S}^*\right) - \mathbf{S}' \circ \mathbf{S}'^*\right). \tag{79}$$

Thus, (77) can be written more compactly as:

$$\frac{\partial h}{\partial \mu} \approx a_1\mu^3 + b_1\mu^2 + c_1\mu + d_1, \tag{80}$$

where:

$$a_1 = 4\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(\mathbf{C}'^*_d \circ \mathbf{C}'_d\right) \mathbf{1}_M + \lambda_1 (\mathbf{C}'_d \mathbf{1}_M)^H (\mathbf{C}'_d \mathbf{1}_M)\right),$$
$$b_1 = 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(2\mathbf{C}'^*_d \circ \mathbf{C}_d + \mathbf{C}^*_d \circ \mathbf{C}'_d\right) \mathbf{1}_M + \lambda_1 \left(2(\mathbf{C}'_d \mathbf{1}_M)^H (\mathbf{C}_d \mathbf{1}_M) + (\mathbf{C}_d \mathbf{1}_M)^H (\mathbf{C}'_d \mathbf{1}_M)\right)\right),$$
$$c_1 = 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(2\mathbf{C}'^*_d \circ \mathbf{P}_d + \mathbf{C}^*_d \circ \mathbf{C}_d\right) \mathbf{1}_M + \lambda_1 \left(2(\mathbf{C}'_d \mathbf{1}_M)^H (\mathbf{P}_d \mathbf{1}_M) + (\mathbf{C}_d \mathbf{1}_M)^H (\mathbf{C}_d \mathbf{1}_M)\right)\right),$$
$$d_1 = 2\mathrm{Re}\left(\lambda \mathbf{w}_f^T \left(\mathbf{C}^*_d \circ \mathbf{P}_d\right) \mathbf{1}_M + \lambda_1 (\mathbf{C}_d \mathbf{1}_M)^H (\mathbf{P}_d \mathbf{1}_M)\right). \tag{81}$$

By solving the cubic equation $a_1\mu^3 + b_1\mu^2 + c_1\mu + d_1 = 0$, the step size can be easily obtained.

4.3. Algorithm Summary

In the previous two subsections, the gradient and step size for Problem (68) are derived. Then, the simplified algorithm (which we call Gra-Corr-SFW) for the design problem without considering the correlation weights is summarized in Algorithm 2. It is easy to observe that both Algorithm 1 and Algorithm 2 can be easily implemented by the (I)FFT operations and the Hadamard product. Since the Hadamard product is more efficient than the (I)FFT operation, we mainly use the number of (I)FFT operations to measure the time complexity of the algorithms. From Section 3, we know that the calculation of $\mathbf{Y}'$ needs $2M^2 + 2M$ (I)FFT operations. Thus, the Gra-WeCorr-SFW requires $5M^2 + 6M$ (I)FFT operations at each iteration. Compared to the Gra-WeCorr-SFW, the Gra-Corr-SFW is simpler and only needs $4M$ (I)FFT operations at each iteration. It is worth noting that the calculation of the cubic function is very simple and needs only a small amount of computation.

Algorithm 2: Gra-Corr-SFW.

Initialization: $l = 0, \lambda, \lambda_1, M, N, \mathbf{w}_t, \mathbf{w}_f, \mathbf{X}^0 = \left[\mathbf{x}_1^0, ..., \mathbf{x}_M^0\right]$,

$\qquad\qquad \mathbf{S} = \mathcal{F}_c\left(\mathbf{X}^0\right), \mathbf{G}^0, \mathbf{D}^0 = -\mathbf{G}^0$.

Repeat

1: $\mathbf{S}' = \mathcal{F}_c\left(\mathbf{X}^l \circ \mathbf{D}^l\right), \mathbf{S}'' = \mathcal{F}_c\left(\mathbf{X}^l \circ \mathbf{D}^l \circ \mathbf{D}^l\right)$.

2: $\mathbf{P}_d = \mathbf{S} \circ \mathbf{S}^*, \mathbf{C}_d = -2\text{Im}\left(\mathbf{S}' \circ \mathbf{S}^*\right), \mathbf{C}'_d = -\left(\text{Re}\left(\mathbf{S}'' \circ \mathbf{S}^*\right) - \mathbf{S}' \circ \mathbf{S}'^*\right)$.

3: Compute the coefficients a_1, b_1, c_1, d_1 according to (81).

4: Solve the cubic equation (80), and then choose the positive root
$\qquad$ which is closest to zero as the step size μ^l.

5: $\mathbf{X}^{l+1} = \mathbf{X}^l \circ e^{j\mu^l \mathbf{D}^l}, \mathbf{S} = \mathcal{F}_c\left(\mathbf{X}^{l+1}\right)$

6: $\mathbf{Y}_2 = \lambda\left(\mathbf{w}_f \mathbf{1}_M^T\right) \circ \mathbf{S} \circ \mathbf{S}^* + \lambda_1\left(\mathbf{S} \circ \mathbf{S}^*\right)\mathbf{1}_M \mathbf{1}_M^T$.

7: $\mathbf{Z}_1 = \mathcal{F}_c^{-1}\left(\mathbf{S} \circ \mathbf{Y}_2\right)$.

8: $\mathbf{G}^{l+1} = 8N\text{Re}\left(-j\left(\mathbf{X}^{l+1}\right)^* \circ \mathbf{Z}_1\left(1:N,:\right)\right)$.

9: Compute the searching direction $\mathbf{D}^{l+1}$ according to (66).

10: $l = l + 1$.

Until convergence

Table 1. The per iteration computational complexities of different algorithms.

Algorithm	Number of (I)FFT	Complexity
MM-WeCorr-acc	$6M^2 + 8M$	$\mathcal{O}((12M + 16)MN \log 2N)$
MM-Corr-acc	$4M$	$\mathcal{O}(8MN \log 2N)$
MDISAA-SFW	$3M$	$\mathcal{O}(6MN \log 2N)$
Gra-WeCorr-SFW	$5M^2 + 6M$	$\mathcal{O}((10M + 12)MN \log 2N)$
Gra-Corr-SFW	$4M$	$\mathcal{O}(8MN \log 2N)$

To analyze the convergence speed, Table 1 presents the per iteration computational complexities of the proposed and existing algorithms. As shown in Table 1, the proposed Gra-WeCorr-SFW requires fewer (I)FFT operations than the MM-WeCorr-acc (MM-WeCorr-acceleration) at each iteration. Compared to these two algorithms, the per iteration computational complexities of the rest of the three algorithms (MM-Corr-acc (MM-Corr-acceleration), MDISAA-SFW and Gra-Corr-SFW), which do not consider the correlation weights, are much smaller. In addition to the per iteration computational complexity, the iteration number is also an important factor affecting the convergence speed. Thus, it is difficult to compare the convergence performance of the algorithms via the per iteration computational

complexity. In the following section, several numerical experiments are provided to show the convergence performance of the proposed algorithms.

5. Numerical Experiments

To illustrate the effectiveness and superiority of the proposed algorithms, several numerical experiments are presented in this section. We first validate the monotonicity of the proposed algorithms and then assess the performance of the algorithms by designing three different waveform sets. The proposed algorithms are compared with the MM-Corr-acc [14], MM-WeCorr-acc [14] and MDISAA-SFW [37] algorithms, where MM-Corr-acc and MM-WeCorr-acc are the state-of-the-art algorithms for designing waveform sets with a good correlation property, and MDISAA-SFW is the state-of-the-art algorithm for designing the orthogonal waveform set (OWS) with the stopband constraint.

All of the experiments are performed on a PC with a 3.60-GHz i7-4790 CPU and 8GB RAM. The software environment is MATLAB 2012b. In the following experiments, all of the algorithms are initialized by the unimodular waveform sets with random phases.

5.1. Verification of The Monotonicity

In this subsection, we investigate the monotonicity of the proposed algorithms in three different waveform design problems, which are respectively complementary sets of sequences (CSS) design, waveform set design with zero correlation zone and orthogonal waveform set (OWS) design with the stopband constraint. In order to measure the monotonicity, define the relative error of the l-th iteration as follows:

$$\varepsilon_r^l = \frac{h(\mu^l) - h(\dot{\mu}^l)}{h(0) - h(\dot{\mu}^l)}, \tag{82}$$

where μ^l denotes the approximate step size of the l-th iteration, which is obtained via the Taylor series expansion, and $\dot{\mu}^l$ denotes the optimal step size obtained by the searching method. The smaller the value of ε_r^l is, the closer the approximate step size is to the optimal step size. Thus, ε_r^l can be used to express the accuracy of the approximate step size. When $\varepsilon_r^l < 1$, we have $h(\mu^l) < h(0)$, which means the objective function is decreasing at the l-th iteration. Therefore, we can measure the monotonicity of the algorithms by using the following peak relative error P_{re}:

$$P_{re} = \max \left(\varepsilon_r^l, l = 1, ..., N_I \right), \tag{83}$$

where N_I is the iteration number.

Table 2. Simulation parameters of different design problems.

	λ	Correlation weights $\{w_k\}_{k=1-N}^{N-1}$	Stopband Ω_f
Case 1[a]	0	$w_k = 1, k = 1 - N, ..., N - 1$	
Case 2[b]	0	$w_k = \begin{cases} 1, & k = 0 \\ 1, & k \in [-80, -51] \cup [51, 80] \\ 0, & \text{otherwise} \end{cases}$	
Case 3[c]	0.9	$w_k = 1, k = 1 - N, ..., N - 1$	$(0.04, 0.21), (0.23, 0.25), (0.28, 0.37),$ $(0.39, 0.49), (0.52, 0.56)$

a: CSS design; b: waveform design with zero correlation zone; c: OWS design with stopband constraint.

To simulate ε_r^l and P_{re}, we use the Gra-WeCorr-SFW to design waveform sets with $M = 3$ waveforms and each waveform of length $N = 256$. The simulation parameters of different design problems are shown in Table 2. Additionally, we stop the algorithm after 200 iterations. Figure 1 shows the evolution curves of the relative error with respect to the iteration number. From Figure 1, we can see that the relative error is a little larger at the initial iterations. However, it decreases rapidly and is substantially

below 10^{-4} after several iterations. In the whole iteration process, the relative error is less than one, which means that the objective function is monotonically decreasing. Figure 2 shows the peak relative error of 100 random trials. It is easy to see that the peak relative error of all three cases is very small and less than one. This indicates that the monotonicity of the Gra-WeCorr-SFW is not affected by the initial iteration point. Since the Gra-Corr-SFW is a simplified algorithm of the Gra-WeCorr-SFW, it can also guarantee that the objective function is monotonically decreasing at each iteration.
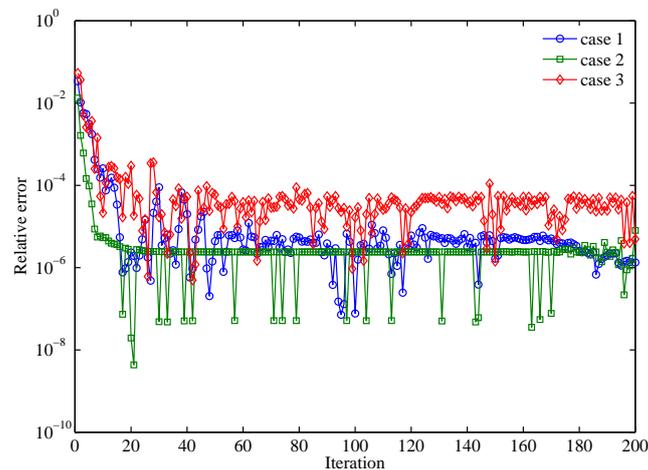
Figure 1. Evolution of the relative error with respect to the iteration number.
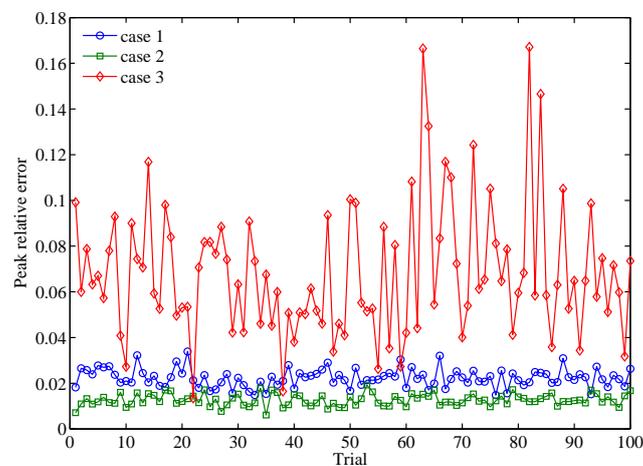
Figure 2. The peak relative errors of 100 random trials.

5.2. CSS Design or OWS Design

In Appendix A, we have demonstrated that the CSS design is equivalent to the OWS design. In this subsection, we apply the Gra-Corr-SFW to design CSS (or OWS). Since the lower bound of the CISL is zero, we choose CISL $\leq \varepsilon$ as the stopping criterion for all of the algorithms in this experiment. The weight coefficient λ and the correlation weights are the same as Case 1 in Table 2. Figure 3 shows the normalized autocorrelation sum of the waveform sets designed by the Gra-Corr-SFW, where the normalized autocorrelation sum is defined as:

$$r_{sum}(k) = 20\log_{10}\frac{\left|\sum\limits_{m=1}^{M} r_{mm}(k)\right|}{\sum\limits_{m=1}^{M} r_{mm}(0)}, k = 1 - N, ..., N - 1. \tag{84}$$
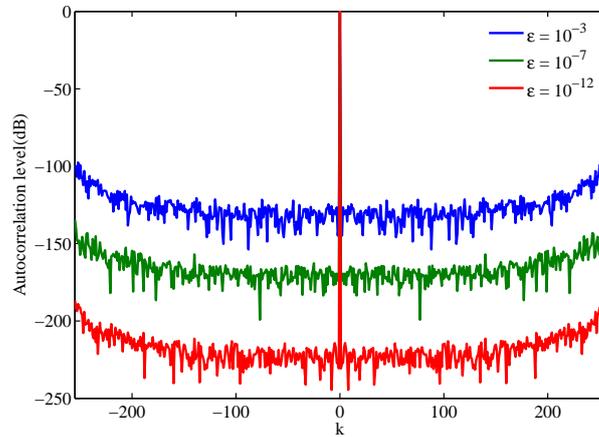

Figure 3. The autocorrelation sum of the waveform sets designed by the Gra-Corr-SFW with different ε. ($N = 256, M = 3$)

From Figure 3, we can see that the autocorrelation sum of the CSS is a Delta function. Additionally, with the decrease of ε, the sidelobes of the CSS are getting lower and lower.

Next, we compare the performance between the proposed Gra-Corr-SFW and the state-of-the-art MM-Corr-acc by designing waveform sets with different waveform number M and waveform length N. For both algorithms, we choose M and N as follows:

$$M = 3k, N = 256k, \quad k = 1, ..., 5. \tag{85}$$
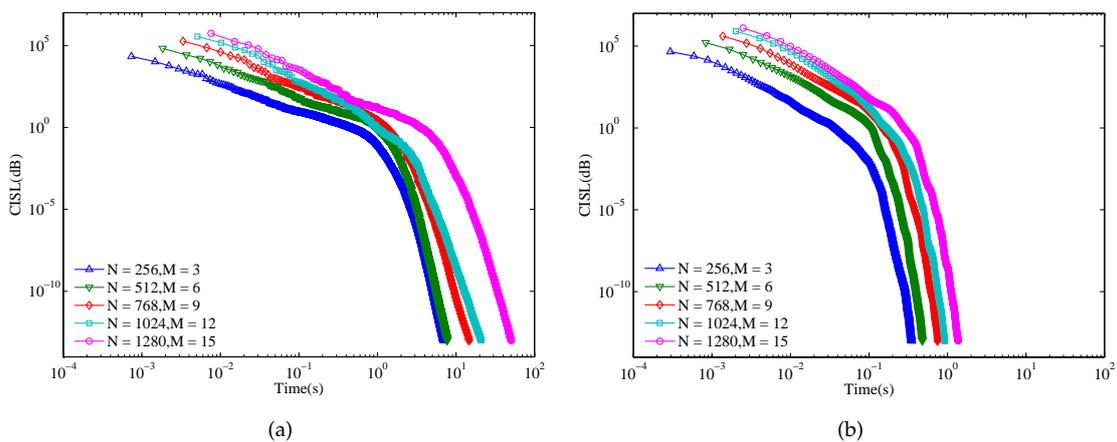
Figure 4. Evolution of the CISL with respect to the running time: (**a**) MM-Corr-acc; (**b**) Gra-Corr-SFW. ($\varepsilon = 10^{-13}$).

The evolution curves of MM-Corr-acc and Gra-Corr-SFW are respectively shown in Figure 4a,b. From these two subfigures, we observe that the convergence speed of the proposed Gra-Corr-SFW

is much faster than that of the MM-Corr-acc. Even for $(N = 1280, M = 15)$, the proposed algorithm takes 1.35 s to coverage to CISL $\leq 10^{-13}$, while the MM-Corr-acc takes 48.11 s.

Further, to eliminate the randomness, we repeat the algorithms 100 times for each (M, N) pair and record the average iteration number N_I, the average running time t and the average peak sidelobe level of the autocorrelation sum P_{sum}, where P_{sum} is defined as:

$$P_{sum} = \max\left(r_{sum}(k), k = 1, ..., N - 1\right). \tag{86}$$

Then, we choose $\varepsilon = 10^{-1}$ for the stopping criterion. The performance parameters (N_I, t, P_{sum}) of these two algorithms are provided in Table 3. As can be seen from this Table, since the algorithms stop the iteration when CISL $\leq 10^{-1}$, the P_{sum} of these two algorithms are basically the same. At the same time, the average running time in Table 3 shows that the Gra-Corr-SFW is an order of magnitude faster than the MM-Corr-acc. From Table 1, we can see that the per iteration computational complexities of both the MM-Corr-acc and the Gra-Corr-SFW are $\mathcal{O}(8MN \log 2N)$. Therefore, the reason for the slower convergence of the MM-Corr-acc may be that the MM strategy in this algorithm makes the objective function loose, so that the iteration number of the MM-Corr-acc is larger than that of the Gra-Corr-SFW (as shown in Table 3).

Table 3. The comparison of the performance parameters between MM-Corr-acc and Gra-Corr-SFW.

(M, N)	**MM-Corr-acc**			**Gra-Corr-SFW**		
	$t(s)$	N_I	$P_{sum}(\text{dB})$	$t(s)$	N_I	$P_{sum}(\text{dB})$
$M = 3, N = 256$	1.249	1687	-63.8	0.071	239	-65.9
$M = 6, N = 512$	0.914	496	-68.7	0.110	132	-71.4
$M = 9, N = 768$	1.558	460	-72.0	0.178	131	-74.7
$M = 12, N = 1024$	2.531	484	-74.4	0.265	132	-77.5
$M = 15, N = 1280$	4.483	587	-76.1	0.343	138	-79.1

5.3. Waveform Set Design with Zero Correlation Zone

To verify the effectiveness of the Gra-WeCorr-SFW, we consider the problem of suppressing the correlation sidelobes at the specified intervals, i.e., designing the waveform set with the zero correlation zone (ZCZ), and compare the performance with the MM-WeCorr-acc. The experimental parameters are the same as Case 2 in Table 2. For both algorithms, we choose the value of the objective function ψ in (4) as the stopping criterion, i.e., $\psi \leq \varepsilon$. The auto- and cross-correlations of the waveform sets designed by MM-WeCorr-acc and Gra-WeCorr-SFW are shown in Figure 5. From this figure, we can see that the proposed algorithm can also generate a waveform set with correlation sidelobes that are almost zero at the specified intervals.

Similar to Figure 4, for each (M, N) pair, we simulate the evolution curves of the objective function ψ with respect to the running time. Additionally, the results of these two algorithms are presented in Figure 6. It is easy to see that the Gra-WeCorr-SFW is faster than the MM-WeCorr-acc. In addition, we can also find that when the objective function ψ is less than 10^{-11}, the convergence speed of the MM-WeCorr-acc decreases, especially for a large (M, N) pair.
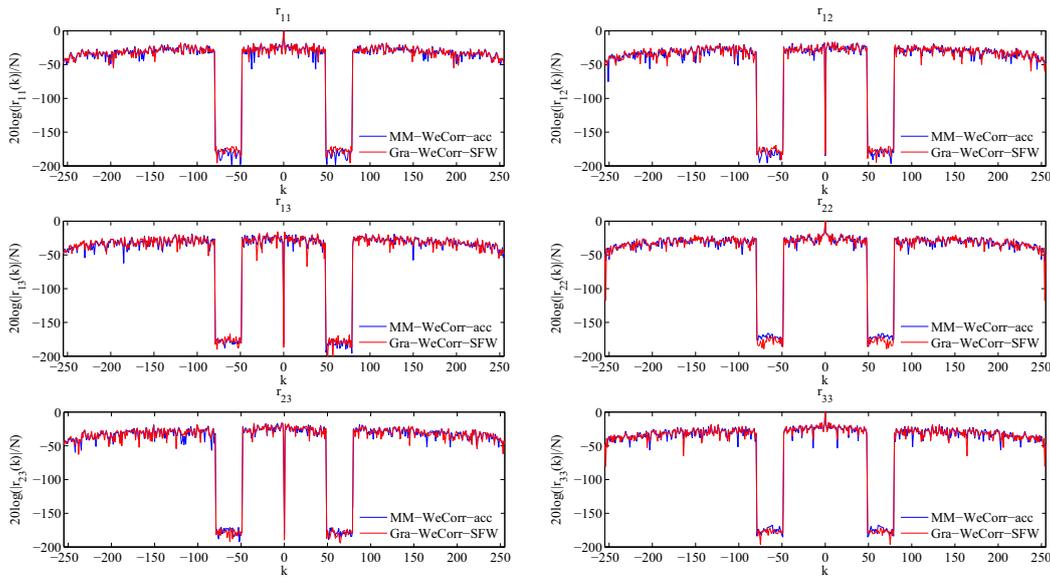
Figure 5. The normalized auto- and cross-correlations of the waveform sets designed by MM-WeCorr-acc and Gra-WeCorr-SFW. ($\varepsilon = 10^{-10}$).
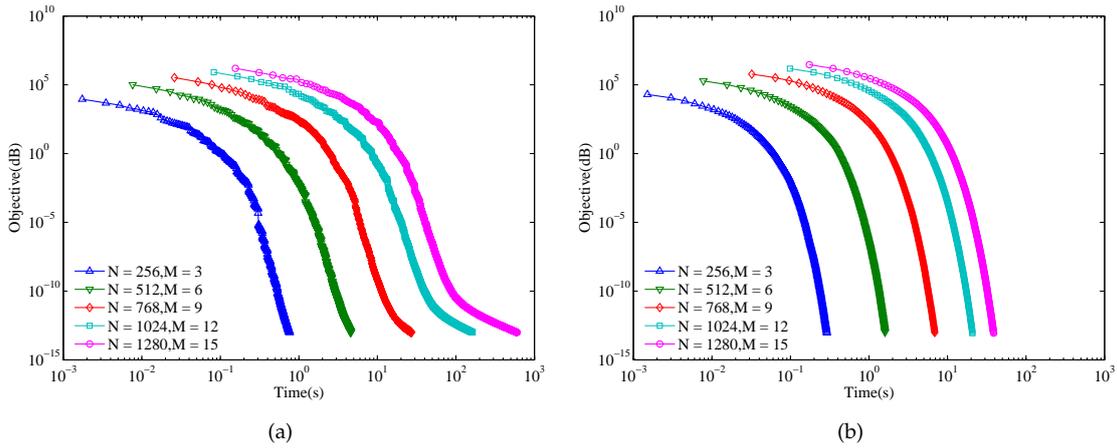

Figure 6. Evolution of the objective function with respect to the running time: (**a**) MM-WeCorr-acc; (**b**) Gra-WeCorr-SFW ($\varepsilon = 10^{-13}$).

Further, Table 4 presents the comparison of performance parameters between MM-WeCorr-acc and Gra-WeCorr-SFW, where P_{zcz} is the peak sidelobe level at the zero correlation zone defined as:

$$P_{zcz} = \max\left(20\log_{10}\left|\frac{r_{ij}(k)}{N}\right|, i, j = 1, ..., M, k \in [51, 80]\right). \tag{87}$$

Similarly, for each (M, N) pair, the algorithms are repeated 100 times. As can be seen from Table 4, the P_{zcz} of the Gra-WeCorr-SFW is a little bit lower than that of the MM-WeCorr-acc. Moreover, due to fewer iterations, the average running time of the Gra-WeCorr-SFW is about 50% of that of the MM-WeCorr-acc, which indicates that the former is more computationally efficient.

Table 4. The comparison of the performance parameters between MM-WeCorr-acc and Gra-WeCorr-SFW ($\varepsilon = 10^{-10}$).

(M, N)	MM-WeCorr-acc			Gra-WeCorr-SFW		
	$t(s)$	N_I	$P_{zcz}(dB)$	$t(s)$	N_I	$P_{zcz}(dB)$
$M = 3, N = 256$	0.584	335	-167.5	0.248	164	-168.2
$M = 6, N = 512$	2.691	407	-178.1	1.348	176	-179.5
$M = 9, N = 768$	10.167	445	-184.5	5.149	178	-186.3
$M = 12, N = 1024$	37.760	486	-189.7	15.106	179	-191.5
$M = 15, N = 1280$	84.759	549	-193.5	30.683	182	-195.1

5.4. OWS Design with the Stopband Constraint

In this subsection, we consider the problem of designing OWS with the stopband constraint. Since the Gra-Corr-SFW is a simplified version of the Gra-WeCorr-SFW and needs much fewer (I)FFT operations, it is more efficient than the Gra-WeCorr-SFW. Thus, we only investigate the performance of the Gra-Corr-SFW and compare it with the MDISAA-SFW. Here, $\left\| \mathbf{X}^{l+1} - \mathbf{X}^l \right\|_F \leq 10^{-4}$ is employed as the stopping criterion, and the experimental parameters are the same as Case 3 in Table 2. First, we apply these two algorithms to design OWS with the stopband constraint. Suppose the waveform number M is three and the waveform length N is 256. The correlation levels and spectral power of the waveform sets designed by MDISAA-SFW and Gra-Corr-SFW are shown in Figure 7 and Figure 8. From Figure 7, we can see that the waveform set designed by the MDISAA-SFW has better autocorrelation performance, while the waveform set designed by the Gra-Corr-SFW has better cross-correlation performance. This is because the MDISAA-SFW optimizes the autocorrelation explicitly and, thus, places more emphasis on suppressing the autocorrelation sidelobes. Figure 8 indicates that the stopband spectral power of the waveform set designed by the Gra-Corr-SFW is less than that of the waveform set designed by the MDISAA-SFW.
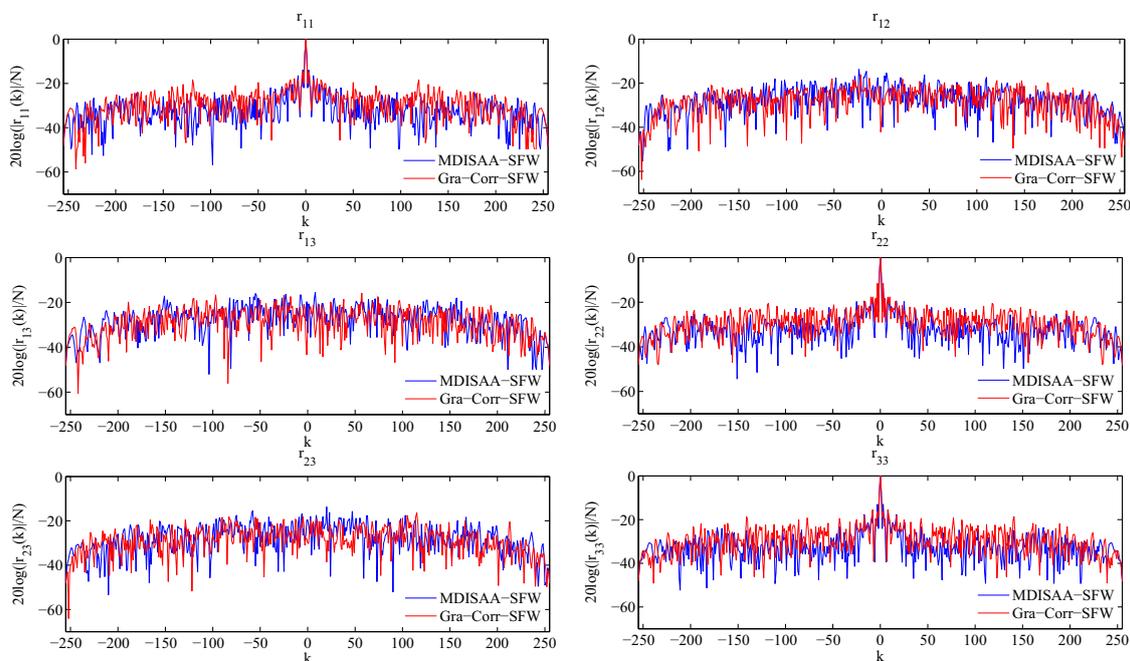

**Figure 7.** The correlation levels of the waveform sets designed by MDISAA-SFW and Gra-Corr-SFW.
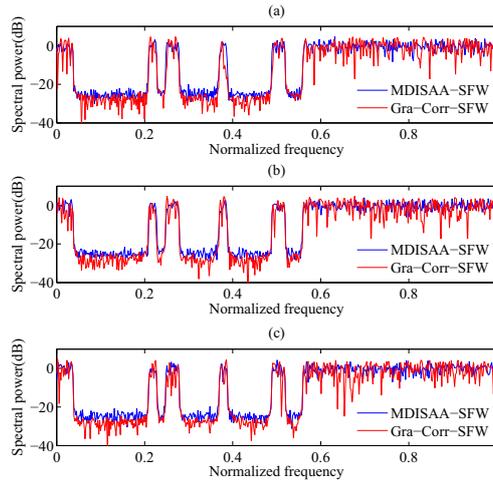
Figure 8. The spectral power of the waveform sets designed by MDISAA-SFW and Gra-Corr-SFW.

In order to fully compare the algorithms, we define the peak autocorrelation sidelobe P_{ac}, peak cross-correlation P_{cc}, peak stopband power P_{sp} and integrated sidelobe level (ISL) [16] as follows:

$$
\begin{aligned}
P_{ac} &= \max\left(20\log_{10}\left|\frac{r_{ii}(k)}{N}\right|, i=1,...,M, k=2,...,N-1\right), \\
P_{cc} &= \max\left(20\log_{10}\left|\frac{r_{ij}(k)}{N}\right|, i,j=1,...,M, i\neq j, k=0,...,N-1\right), \\
P_{sp} &= \max\left(P_{stop}^{i}, i=1,...,M\right), \\
ISL &= \sum_{i=1}^{M}\sum_{\substack{k=-N+1\\k\neq 0}}^{N-1}|r_{ii}(k)|^2 + \sum_{i=1}^{M}\sum_{\substack{j=1\\j\neq i}}^{M}\sum_{k=-N+1}^{N-1}|r_{ij}(k)|^2,
\end{aligned}
\tag{88}
$$

where P_{stop}^{i} denotes the peak stopband power of the waveform $\mathbf{x}_i$. In the above four performance parameters, P_{ac} and P_{cc} are used to measure the autocorrelation and cross-correlation performances, respectively; P_{sp} is the parameter related to the stopband performance; and ISL indicates the overall sidelobe performance. Since the decrease of the bandwidth leads to the broadening of the main lobe, $r_{ii}(1), i=1,...,M$ can be regarded as the autocorrelation main lobe. Thus, the definition of P_{ac} does not take $k=1$ into account. Here, we consider the waveform sets with $M=3$ waveforms and each waveform of length $N \in \{256, 512, 1024\}$. For each (M, N) pair, we repeat these two algorithms 100 times and record the average values of the performance parameters. The comparison of the performance parameters between MDISAA-SFW and Gra-Corr-SFW is shown in Table 5. In addition to the parameters defined in (88), the average iteration number N_I and the running time t are provided in Table 5. From this table, we obtain three findings, as follows. Above all, the P_{cc} and P_{sp} of the Gra-Corr-SFW are lower than that of the MDISAA-SFW, which means that the cross-correlation and stopband performances of the Gra-Corr-SFW are better than that of the MDISAA-SFW. However, in terms of autocorrelation performance, the Gra-Corr-SFW is inferior to the MDISAA-SFW. From the ISL in Table 5, it is easy to see that the overall sidelobe performance of the Gra-Corr-SFW is better than that of the MDISAA-SFW. Furthermore, although the per iteration computational complexity of the Gra-Corr-SFW is slightly higher than that of the MDISAA-SFW (see Table 1), the Gra-Corr-SFW needs fewer iterations due to the fast convergence of the gradient algorithm, which makes the proposed algorithm more efficient. Finally, with the increase of the waveform length, both the correlation and stopband performances of these two algorithms are improved, especially the cross-correlation performance.

Table 5. The comparison of the performance parameters between MDISAA-SFW and Gra-Corr-SFW.

Algorithm	N	P_{ac}(dB)	P_{cc}(dB)	P_{sp}(dB)	ISL	N_I	t(s)
	$N = 256$	-13.6	-13.9	-17.5	893,573	11932	2.512
MDISAA-SFW	$N = 512$	-13.9	-16.6	-17.9	3,532,447	18,602	6.045
	$N = 1024$	-14.1	-19.3	-18.2	14,013,417	27,268	14.080
	$N = 256$	-12.5	-15.9	-20.3	832,482	1216	0.274
Gra$-$Corr$-$SFW	$N = 512$	-12.9	-18.6	-22.2	3,321,729	2100	1.013
	$N = 1024$	-13.3	-21.1	-23.8	13,237,903	3577	3.033

Next, we investigate the performance of the algorithms under different λ. The average values (100 trials) of P_{ac}, P_{cc} and P_{sp} are shown in Figure 9. From this figure, we can see that with the increase of λ, the P_{ac} and P_{cc} of these two algorithms change slowly, while the P_{sp} is sensitive to the change of λ. It can also be observed that except the autocorrelation performance, the cross-correlation and stopband performances of the Gra-Corr-SFW are better than that of the MDISAA-SFW. This means that the stopband property can be easily obtained by using the Gra-Corr-SFW. Moreover, since the correlation performance changes slowly with the increase of λ, we can choose a relatively large λ (e.g., $\lambda \in (0.85, 0.95)$) for the Gra-Corr-SFW to obtain both good correlation and stopband properties.
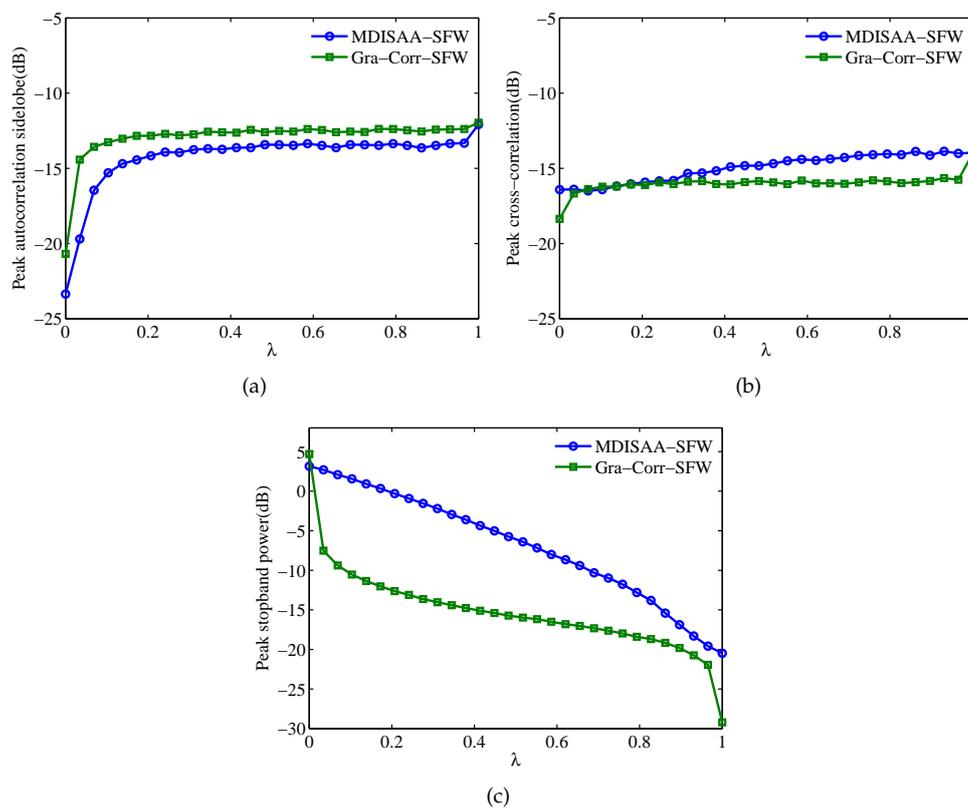
Figure 9. The comparison of P_{ac}, P_{cc} and P_{sp} between MDISAA-SFW and Gra-Corr-SFW versus λ from 0–1. (**a**) P_{ac}; (**b**) P_{cc}; (**c**) P_{sp} ($M = 3, N = 256$).

6. Conclusions

In this paper, we propose an efficient algorithm named Gra-WeCorr-SFW for designing the waveform set with a good correlation and stopband properties. The algorithm optimizes the objective function directly and can guarantee that the objective function is monotonically decreasing at each iteration. By changing the design parameters, the Gra-WeCorr-SFW can be used to generate different waveform sets, such as CSS, the waveform set with ZCZ and OWS with the stopband constraint.

As the main steps can be implemented by the FFT operations and the Hadamard product, the proposed algorithm is computationally efficient and can be used to design waveform sets with large M and N. Besides, the simplified version of the Gra-WeCorr-SFW (named Gra-Corr-SFW) is proposed for the design problem without considering the correlation weights. Compared to the Gra-WeCorr-SFW, the simplified algorithm requires fewer FFT operations and is faster. Numerical experiments show that the proposed algorithms are faster than the state-of-the-art algorithms (MM-WeCorr-acc and MM-Corr-acc) when designing CSS or the waveform set with ZCZ. In the case of designing OWS with the stopband constraint, the simplified algorithm has better stopband performance and computational efficiency compared with the MDISAA-SFW.

Author Contributions: Liang Tang conceived of and devised the idea. Liang Tang and Yongfeng Zhu designed and performed the experiments. Yongfeng Zhu and Qiang Fu analyzed the experiment results. All authors contributed to writing and editing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MIMO	multiple-input multiple-output
(I)FFT	(inverse) fast Fourier transform
CSS	complementary sets of sequences
OWS	orthogonal waveform set
CAN	cyclic algorithm-new
WeCAN	weighted cyclic algorithm-new
LBFGS	limited-memory Broyden, Fletcher, Goldfarb and Shanno
MM	majorization-minimization
MM-Corr	MM-correlation
MM-Corr-acc	MM-correlation-acceleration
MM-WeCorr	MM-weighted correlation
MM-WeCorr-acc	MM-weighted correlation-acceleration
CDMA	code-division multiple-access
OFDM	orthogonal frequency division multiplexing
SFW	sparse frequency waveforms
HFSWR	high frequency surface wave radar
UWB	ultra-wide bandwidth
MDISAA-SFW	multi-dimensional iterative spectral approximation algorithm-SFW
PSD	power spectrum density
Gra-WeCorr-SFW	gradient-weighted correlation-SFW
Gra-Corr-SFW	gradient-correlation-SFW
CISL	complementary integrated sidelobe level

Appendix A

Proof. Due to the unimodular constraint, the energy of each waveform is N, i.e., $r_{ii}(0) = N, i = 1,...,M$. Then, (3) can be rewritten as:

$$\text{CISL} = \sum_{k=1-N}^{N-1} \left| \sum_{m=1}^{M} r_{mm}(k) \right|^2 - \left| \sum_{m=1}^{M} r_{mm}(0) \right|^2 = \sum_{k=1-N}^{N-1} \left| \sum_{m=1}^{M} r_{mm}(k) \right|^2 - M^2 N^2. \tag{A1}$$

Since $\left| \sum\limits_{m=1}^{M} r_{mm}(k) \right|^2 = \sum\limits_{i=1}^{M} \sum\limits_{j=1}^{M} r_{ii}^*(k) r_{jj}(k)$, we have:

$$\text{CISL} = \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k=1-N}^{N-1} r_{ii}^*(k) r_{jj}(k) - M^2 N^2 = \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{r}_{ii}^H \mathbf{r}_{jj} - M^2 N^2. \tag{A2}$$

by substituting (6) into (A2), CISL can be expressed as:

$$\text{CISL} = \frac{1}{2N} \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ii}^H \mathbf{p}_{jj} - M^2 N^2. \tag{A3}$$

Assume $\mathbf{f}_m$ is the $2N$-point discrete Fourier transform (DFT) vector of the waveform $\mathbf{x}_m, m = 1, ..., M$, i.e.,

$$f_m(i) = \sum_{n=1}^{N} x_m(n) e^{-jn\omega_i}, \omega_i = \frac{2\pi}{2N} i, i = 1, ..., 2N, \tag{A4}$$

then the PSD vector $\mathbf{p}_{ij}$ can be written as:

$$\mathbf{p}_{ij} = \mathbf{f}_i \circ \mathbf{f}_j^*. \tag{A5}$$

According to (A5), it is easy to verify that $\mathbf{p}_{ii}^H \mathbf{p}_{jj} = \mathbf{p}_{ij}^H \mathbf{p}_{ij}$. Thus, (A3) can be denoted as:

$$\text{CISL} = \frac{1}{2N} \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ij}^H \mathbf{p}_{ij} - M^2 N^2. \tag{A6}$$

Let $w_t(k) = 1, k = 1 - N, ..., N - 1$, then we have $\mathbf{D}_t = \text{Diag}(\mathbf{1}_{2N}) = \mathbf{I}_{2N}$. Thus, (8) can be simplified as:

$$\psi_1 = \frac{1}{2N} \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{p}_{ij}^H \mathbf{p}_{ij} - MN^2. \tag{A7}$$

From (A6) and (A7), we can observe that the only different between CISL and ψ_1 is the constant term, i.e., CISL and ψ_1 are equivalent. The proof is complete. □

Appendix B

Proof. According to (A5), it is easy to verify that:

$$\mathbf{p}_{im} = \mathbf{p}_{mi}^*. \tag{A8}$$

by substituting (A8) into $\sum\limits_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{im}^H \mathbf{Q}' \mathbf{p}_{im}$, we have:

$$\sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{im}^H \mathbf{Q}' \mathbf{p}_{im} = \sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{mi}^T \mathbf{Q}' \mathbf{p}_{mi}^* = \left(\sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{mi}^H \mathbf{Q}'^T \mathbf{p}_{mi} \right)^T = \sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{mi}^H \mathbf{Q}'^T \mathbf{p}_{mi}. \tag{A9}$$

Since:

$$\mathbf{Q}'^H = \frac{1}{4N^2} (1 - \lambda) \left(\mathbf{F} \mathbf{D}_t \mathbf{F}^H \right)^H = \mathbf{Q}', \tag{A10}$$

we have:

$$\mathbf{Q}'^T = \left(\mathbf{Q}'^H \right)^* = \mathbf{Q}'^*. \tag{A11}$$

Thus, (A9) can be rewritten as:

$$\sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{im}^{H} \mathbf{Q}' \mathbf{p}_{im} = \sum_{\substack{i=1 \\ i \neq m}}^{M} \mathbf{p}_{mi}^{H} \mathbf{Q}'^{*} \mathbf{p}_{mi}. \tag{A12}$$

The proof is complete. □

Appendix C

Proof. Here, we temporarily replace the subscripts i, j with p, q (i.e., $\mathbf{p}_{pq}^{l+1}$) to avoid the confusion of the number j and the imaginary unit j. According to (A4), (A5) and (40), the k-th element of $\mathbf{p}_{pq}^{l+1}$ can be written as:

$$p_{pq}^{l+1}(k) = f_p^{l+1}(k) \left(f_q^{l+1}(k) \right)^{*} = \sum_{n=1}^{N} \sum_{m=1}^{N} x_p^{l}(n) \left(x_q^{l}(m) \right)^{*} \cdot e^{j\mu \left(d_p^{l}(n) - d_q^{l}(m) \right)} \cdot e^{-j\frac{2\pi}{2N}(n-m)k}, \tag{A13}$$

where $x_p^{l}(n), d_p^{l}(n)$ denote the n-th elements of $\mathbf{x}_p^{l}$ and $\mathbf{d}_p^{l}$. By using the Taylor series, the expansion of $e^{j\mu \left(d_p^{l}(n) - d_q^{l}(m) \right)}$ in (A13) (which we keep the first three terms) is given by:

$$e^{j\mu \left(d_p^{l}(n) - d_q^{l}(m) \right)} \approx 1 + j\mu \left(d_p^{l}(n) - d_q^{l}(m) \right) + \left(j\mu \left(d_p^{l}(n) - d_q^{l}(m) \right) \right)^{2} \Big/ 2. \tag{A14}$$

Thus, $p_{pq}^{l+1}(k)$ can be approximated as:

$$\begin{aligned} p_{pq}^{l+1}(k) \approx &\sum_{n=1}^{N} \sum_{m=1}^{N} x_p^{l}(n) \left(x_q^{l}(m) \right)^{*} \cdot e^{-j\frac{2\pi}{2N}(n-m)k} \\ &+ j\mu \sum_{n=1}^{N} \sum_{m=1}^{N} x_p^{l}(n) \left(x_q^{l}(m) \right)^{*} \cdot \left(d_p^{l}(n) - d_q^{l}(m) \right) \cdot e^{-j\frac{2\pi}{2N}(n-m)k} \\ &- \frac{\mu^2}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} x_p^{l}(n) \left(x_q^{l}(m) \right)^{*} \cdot \left(d_p^{l}(n) - d_q^{l}(m) \right)^{2} \cdot e^{-j\frac{2\pi}{2N}(n-m)k} \end{aligned} \tag{A15}$$

Let:
$$\mathbf{f}_i = \mathcal{F}\left(\mathbf{x}_i^{l} \right), \mathbf{f}'_i = \mathcal{F}\left(\mathbf{x}_i^{l} \circ \mathbf{d}_i^{l} \right), \mathbf{f}''_i = \mathcal{F}\left(\mathbf{x}_i^{l} \circ \mathbf{d}_i^{l} \circ \mathbf{d}_i^{l} \right). \tag{A16}$$

Then, (A15) can be rewritten as:

$$\begin{aligned} p_{pq}^{l+1}(k) \approx &f_p(k) f_q^{*}(k) + j\mu \left(f'_p(k) f_q^{*}(k) - f_p(k) \left(f'_q(k) \right)^{*} \right) \\ &- \frac{\mu^2}{2} \left(f''_p(k) f_q^{*}(k) + f_p(k) \left(f''_q(k) \right)^{*} - 2f'_p(k) \left(f'_q(k) \right)^{*} \right), \end{aligned} \tag{A17}$$

where $f_p(k), f'_p(k), f''_p(k)$ are the k-th elements of $\mathbf{f}_p, \mathbf{f}'_p, \mathbf{f}''_p$. By representing (A17) as a vector, we have:

$$\begin{aligned} \mathbf{p}_{pq}^{l+1} \approx &\mathbf{f}_p \circ \mathbf{f}_q^{*} + j \left(\mathbf{f}'_p \circ \mathbf{f}_q^{*} - \mathbf{f}_p \circ \mathbf{f}'^{*}_q \right) \mu \\ &- \frac{1}{2} \left(\mathbf{f}''_p \circ \mathbf{f}_q^{*} + \mathbf{f}_p \circ \mathbf{f}''^{*}_q - 2\mathbf{f}'_p \circ \mathbf{f}'^{*}_q \right) \mu^2. \end{aligned} \tag{A18}$$

The proof is complete. □

References

1. Blunt, S.D.; Mokole, E.L. Overview of radar waveform diversity. *IEEE Aerosp. Electron. Syst. Mag.* **2016**, *31*, 2–42.
2. Cheng, X.; Aubry, A.; Ciuonzo, D.; De Maio, A.; Wang, X. Robust waveform and filter bank design of polarimetric radar. *IEEE Trans. Aerosp. Electron. Syst.* in press.
3. He, H.; Li, J.; Stoica, P. *Waveform Design for Active Sensing Systems: A Computational Approach*; Cambridge University Press: Cambridge, UK, 2012.
4. Chen, C.Y.; Vaidyanathan, P.P. MIMO radar waveform optimization with prior information of the extended target and clutter. *IEEE Trans. Signal Process.* **2009**, *57*, 3533–3544.
5. Ciuonzo, D.; De Maio, A.; Foglia, G.; Piezzo, M. Intrapulse radar-embedded communications via multiobjective optimization. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 2960–2974.
6. Li, J.; Stoica, P. *MIMO Radar Signal Processing*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
7. Li, J.; Stoica, P.; Zheng, X.Y. Signal synthesis and receiver design for MIMO radar imaging. *IEEE Trans. Signal Process.* **2008**, *56*, 3959–3968.
8. Wang, Z.J.; Babu, P.; Palomar, D.P. Design of PAR-constrained sequences for MIMO channel estimation via majorization-minimization. *IEEE Trans. Signal Process.* **2016**, *64*, 6132–6144.
9. Deng, H. Polyphase code design for orthogonal netted radar systems. *IEEE Trans. Signal Process.* **2004**, *52*, 3126–3135.
10. Khan, H.A.; Zhang, Y.Y.; Ji, C.L.; Stevens, C.J.; Edwards, D.J.; O'Brien, D. Optimizing polyphase sequences for orthogonal netted radar. *IEEE Signal Process. Lett.* **2006**, *13*, 589–592.
11. He, H.; Stoica, P.; Li, J. Designing unimodular sequence sets with good correlations-including an application to MIMO radar. *IEEE Trans. Signal Process.* **2009**, *57*, 4391–4405.
12. Wang, Y.C.; Dong, L.; Xue, X.; Yi, K.C. On the design of constant modulus sequences with low correlation sidelobes levels. *IEEE Commun. Lett.* **2012**, *16*, 462–465.
13. Wu, H.; Fu, Q.; Li, Y.X.; Xia, Y. Designing sequence sets with good correlation properties based on PSD fitting. *Electron. Lett.* **2015**, *51*, 1193–1195.
14. Song, J.; Babu, P.; Palomar, D.P. Sequence set design with good correlation properties via majorization-minimization. *IEEE Trans. Signal Process.* **2016**, *64*, 2866–2879.
15. O'Donnell, B.; Baden, J.M. Fast gradient descent for multi-objective waveform design. In Proceedings of the 2016 IEEE Radar Conference, Philadelphia, PA, USA, 2–6 May 2016.
16. He, H.; Stoica, P.; Li, J. On aperiodic-correlation bounds. *IEEE Signal Process. Lett.* **2010**, *17*, 253–256.
17. Bomer, L.; Antweiler, M. Periodic complementary binary sequences. *IEEE Trans. Inf. Theory* **1990**, *36*, 1487–1494.
18. Levanon, N. Noncoherent radar pulse compression based on complementary sequences. *IEEE Trans. Aerosp. Electron. Syst.* **2009**, *45*, 742–747.
19. Wu, H.; Song, Z.; Liu, Z.; Fu, Q. Complete complementary sequence for MIMO radar waveform design with low range sidelobes. In Proceedings of the 2015 IET International Radar Conference, Hangzhou, China, 14–16 October 2015.
20. Li, F.; Chen, J.; Zhang, L. Optimisation of complete complementary codes in MIMO radar system. *Electron. Lett.* **2010**, *46*, 1157–1159.
21. Tseng, S.M.; Bell, M.R. Asynchronous multicarrier DS-CDMA using mutually orthogonal complementary sets of sequences. *IEEE Trans. Commun.* **2000**, *48*, 53–59.
22. Spasojevic, P.; Georghiades, C.N. Complementary sequences for ISI channel estimation. *IEEE Trans. Inf. Theory* **2001**, *47*, 1145–1152.
23. Schmidt, K.U. Complementary sets, generalized Reed-Muller codes, and power control for OFDM. *IEEE Trans. Inf. Theory* **2007**, *53*, 808–814.
24. Soltanalian, M.; Stoica, P. Computational design of sequences with good correlation properties. *IEEE Trans. Signal Process.* **2012**, *60*, 2180–2193.
25. Soltanalian, M.; Naghsh, M.M.; Stoica, P. A fast algorithm for designing complementary sets of sequences. *Signal Process.* **2013**, *93*, 2096–2102.
26. Stoica, P.; He, H.; Li, J. New algorithms for designing unimodular sequences with good correlation properties. *IEEE Trans. Signal Process.* **2009**, *57*, 1415–1425.

27.   Lindenfeld, M.J. Sparse frequency transmit and receive waveform design. *IEEE Trans. Aerosp. Electron. Syst.* **2004**, *40*, 851–861.

28.   Kassab, R.; Lesturgie, M.; Fiorina, J. Performance analysis of interrupted sparse HFSWR waveform coded with successive fast Fourier transforms. *Electron. Lett.* **2009**, *45*, 525–526.

29.   Liu, W.X.; Lesturgie, M.; Lu, Y.L. Real-time sparse frequency waveform design for HFSWR system. *Electron. Lett.* **2007**, *43*, 1387–1389.

30.   Haykin, S. Cognitive radar: A way of the future. *IEEE Signal Process. Mag.* **2006**, *23*, 30–40.

31.   Wang, G.H.; Mai, C.Y.; Sun, J.P.; Lu, Y.L. Sparse frequency waveform analysis and design based on ambiguity function theory. *IET Radar Sonar Navig.* **2016**, *10*, 707–717.

32.   Liang, J.L.; So, H.C.; Li, J.; Farina, A. Unimodular Sequence Design Based on Alternating Direction Method of Multipliers. *IEEE Trans. Signal Process.* **2016**, *64*, 5367–5381.

33.   Feng, X.; Song, Y.C.; Zhou, Z.Q.; Zhao, Y.N. Designing unimodular waveform with low range sidelobes and stopband for cognitive radar via relaxed alternating projection. *Int. J. Antenna. Propag.* **2016**, *2016*. doi:org/10.1155/2016/6280508.

34.   Liang, J.; So, H.C.; Leung, C.S.; Li, J.; Farina, A. Waveform design with unit modulus and spectral shape constraints via lagrange programming neural network. *IEEE J. Sel. Topics Signal Process.* **2015**, *9*, 1377–1386.

35.   Rowe, W.; Stoica, P.; Li, J. Spectrally constrained waveform design. *IEEE Signal Process. Mag.* **2014**, *31*, 157–162.

36.   Wang, G.; Lu, Y. Designing single/multiple sparse frequency waveforms with sidelobe constraint. *IET Radar Sonar Navig.* **2011**, *5*, 32–38.

37.   Zhao, Y.N.; Li, F.C.; Zhang, T.; Zhou, Z.Q. Computational design of optimal waveforms for MIMO radar via multi-dimensional iterative spectral approximation. *Multidim. Syst. Signal Process.* **2016**, *27*, 43–60.

38.   Yue, W.; Zhang, Y.; Liu, Y.; Xie, J. Radar constant-modulus waveform design with prior information of the extended target and clutter. *Sensors* **2016**, *16*, doi:10.3390/s16060889.