SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Constant modulus sequence set design with low weighted integrated sidelobe level in spectrally crowded environments

Yi BU[1], Hui QIU[1], Tao FAN[1], Xianxiang YU[1], Guolong CUI[1*] & Lingjiang KONG[1]

[1]*School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu* 611731*, China*

Appendix A Comments on the design model $\mathcal{P}_0$

$$\mathcal{P}_0 \begin{cases} \min\limits_{\boldsymbol{s}} \ f(\boldsymbol{s}) \\ \text{s.t.} \ \ \phi_n \in \Phi, \, n \in \mathcal{N}, \\ \qquad \boldsymbol{s}(n) = 1/\sqrt{NM} e^{j\phi_n}, n \in \mathcal{N}, \\ \qquad \boldsymbol{s}^{\dagger} \left(\boldsymbol{I}_M \otimes \boldsymbol{R}_I \right) \boldsymbol{s} \leqslant E_I, \end{cases} \tag{A1}$$

The novel design framework $\mathcal{P}_0$ is raised for a cognitive radar that coexistence with other frequency overlaid emitters. The premise of using waveforms to optimize radar performance while ensuring non-interference in the spectrum is the spectrum awareness [1]. Actually, two main classes available techniques can be applied to achieve spectrum recognition [2]. One is supervised procedures that exploit intrinsic and specific emitter characteristics, and the other one relies on the statistics of the collected measurements which can be classified as unsupervised methods. By utilizing the receiver of cognitive radar, other spectral sensing strategies have been derived [3]. For instance, array signal processing techniques, two-dimensional (2-D) sensing strategies, the adaptive cyclostationary-beamforming-based spectrum sensing method, and so on.

To avoid the performance degradation in the digital signal generators that resulted by the limitation of available number of bits, the discrete phase constraint is considered in this design framework. And it is first combined with the spectral allocation in the minimization of the multi-waveform correlation sidelobes. The constant modulus constraint is also added to comply with the current amplifier technology. Therefore, this quartic optimization problem $\mathcal{P}_0$ with non-continuous and non-smooth constraints is non-convex and in general NP-hard.

Appendix B Details of the Inexact Alternating Direction Penalty Method (IADPM) framework

Next, an alternating iterative algorithm based on the IADPM framework is proposed to determine $(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}_1, \boldsymbol{u}_2, \varrho_1, \varrho_2)$. In particular, the estimates of $\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}_1, \boldsymbol{u}_2, \varrho_1, \varrho_2$ from the $(t-1)$th iteration are denoted by $\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t-1)}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_1^{(t-1)}, \boldsymbol{u}_2^{(t-1)}, \varrho_1^{(t-1)}, \varrho_2^{(t-1)}$. Then the iteration rule is given as follows:

- Given $\boldsymbol{s}^{(t-1)}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_p^{(t-1)}$, and $\varrho_p^{(t-1)}, p \in \{1, 2\}$,solve

$$\mathcal{P}_{\boldsymbol{x}^{(t)}} \begin{cases} \min\limits_{\boldsymbol{x}} \ L(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_p^{(t-1)}, \varrho_p^{(t-1)}) \\ \text{s.t.} \ \arg \boldsymbol{x}(n) \in \Phi, n \in \mathcal{N} \\ \qquad |\boldsymbol{x}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}. \end{cases} \tag{B1}$$

- Given $\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}, \boldsymbol{u}_p^{(t-1)}$, and $\varrho_p^{(t-1)}, p \in \{1, 2\}$, solve

$$\mathcal{P}_{\boldsymbol{y}^{(t)}} \begin{cases} \min\limits_{\boldsymbol{y}} \ L(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}, \boldsymbol{u}_p^{(t-1)}, \varrho_p^{(t-1)}) \\ \text{s.t.} \ \ \boldsymbol{y}^{\dagger} \left(\boldsymbol{I}_M \otimes \boldsymbol{R}_I \right) \boldsymbol{y} \leqslant E_I. \end{cases} \tag{B2}$$

- Given $\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(t-1)}$, and $\varrho_p^{(t-1)}, p \in \{1, 2\}$, solve

$$\mathcal{P}_{\boldsymbol{s}^{(t)}} \begin{cases} \min\limits_{\boldsymbol{s}} \ L(\boldsymbol{s}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(t-1)}, \varrho_p^{(t-1)}) \\ \text{s.t.} \ \ |\boldsymbol{s}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}. \end{cases} \tag{B3}$$

- $$\varrho_p^{(t)} = \begin{cases} \varrho_p^{(t-1)} & \triangle r_p^{(t)} \leqslant \delta_{1,c} \triangle r_p^{(t-1)}, \\ \varrho_p^{(t-1)} \delta_{2,c} & \text{else}, \end{cases} \tag{B4}$$

where $p \in \{1, 2\}$, $\triangle r_1^{(t)} = \|\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)}\|$, $\triangle r_2^{(t)} = \|\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)}\|$, $0 < \delta_{1,c} < 1$ and $\delta_{2,c} > 1$ but close to 1.

* Corresponding author (email: cuiguolong@uestc.edu.cn)

•

$$\boldsymbol{u}_p^{(t)}=\begin{cases}\tilde{\boldsymbol{u}}_p^{(t)} & u_{p\max}^{(t)}\leqslant\nu,\\ \tilde{\boldsymbol{u}}_p^{(t)}/u_{p\max}^{(t)} & \text{else},\end{cases}\tag{B5}$$

where $p\in\{1,2\}$, $\tilde{\boldsymbol{u}}_1^{(t)}=\boldsymbol{u}_1^{(t-1)}+\varrho_1^{(t)}(\boldsymbol{x}^{(t)}-\boldsymbol{s}^{(t)})$, $\tilde{\boldsymbol{u}}_2^{(t)}=\boldsymbol{u}_2^{(t-1)}+\varrho_2^{(t)}(\boldsymbol{y}^{(t)}-\boldsymbol{s}^{(t)})$, $u_{p\max}^{(t)}=\max[|\tilde{\boldsymbol{u}}_p^{(t)}(1)|,\cdots,|\tilde{\boldsymbol{u}}_p^{(t)}(N)|]$ and ν is a large enough positive number.

The above steps are repeated until stopping criterion is reached. Boyd et al. [4] suggested an exit condition, given by

$$\|\boldsymbol{x}^{(t)}-\boldsymbol{s}^{(t)}\|+\|\boldsymbol{y}^{(t)}-\boldsymbol{s}^{(t)}\|\leqslant\epsilon_{\text{pri}}^{(t)},\tag{B6}$$

$$\varrho_1^{(0)}\|\boldsymbol{x}^{(t-1)}-\boldsymbol{x}^{(t)}\|+\varrho_2^{(0)}\|\boldsymbol{y}^{(t-1)}-\boldsymbol{y}^{(t)}\|\leqslant\epsilon_{\text{dual}}^{(t)},\tag{B7}$$

where $\epsilon_{\text{pri}}^{(t)}>0$ and $\epsilon_{\text{dual}}^{(t)}>0$ are feasibility tolerances of primal and dual residuals at the tth iteration, respectively. These tolerances can be chosen based on an absolute and a relative critera, such as

$$\epsilon_{\text{pri}}^{(t)}=\sqrt{4NM}\epsilon_{\text{abs}}+\epsilon_{\text{rel}}\max\{2\|\boldsymbol{s}^{(t)}\|,\|\boldsymbol{x}^{(t)}\|+\|\boldsymbol{y}^{(t)}\|\},\tag{B8}$$

$$\epsilon_{\text{dual}}^{(t)}=\sqrt{2NM}\epsilon_{\text{abs}}+\epsilon_{\text{rel}}(\|\boldsymbol{u}_1^{(t)}\|+\|\boldsymbol{u}_2^{(t)}\|),\tag{B9}$$

where $\epsilon_{\text{abs}}>0$ is an absolute tolerance and $\epsilon_{\text{rel}}>0$ is a relative tolerance.

Appendix C The procedure to tackle $\mathcal{P}_{\boldsymbol{x}^{(t)}}$

The update procedure of $\boldsymbol{x}$ with the fixed $\boldsymbol{s}^{(t-1)},\boldsymbol{y}^{(t-1)},\boldsymbol{u}_p^{(t-1)}$ and $\varrho_p^{(t-1)}$, $p\in\{1,2\}$ is shown next.

Omitting the constant term irrelevant to $\boldsymbol{x}$ in the objective function $L(\boldsymbol{s}^{(t-1)},\boldsymbol{s},\boldsymbol{y}^{(t-1)},\boldsymbol{u}_p^{(t-1)},\varrho_p^{(t-1)})$, $\mathcal{P}_{\boldsymbol{x}^{(t)}}$ boils down to:

$$\mathcal{P}_{\boldsymbol{x}}\begin{cases}\min\limits_{\boldsymbol{x}} & \Re\left\{-\boldsymbol{c}^{(t)\dagger}\boldsymbol{x}\right\}\\ \text{s.t.} & \arg\boldsymbol{x}(n)\in\Phi,n\in\mathcal{N},\\ & |\boldsymbol{x}(n)|=1/\sqrt{NM},n\in\mathcal{N},\end{cases}\tag{C1}$$

where $\boldsymbol{c}^{(t)}=\boldsymbol{s}^{(t-1)}-\boldsymbol{u}_1^{(t-1)}/\varrho_1^{(t-1)}$.

It is worth noting that the objective function and the constraints in Problem (C1) are separable with respect to each $\boldsymbol{x}(n)$. Hence, it amounts to solving NM separate subproblems *in parallel*. When only focusing on the nth element of $\boldsymbol{x}$, it can be recast as

$$\max_{\theta_n\in\Phi}\cos(\theta_n-\varphi_n),\tag{C2}$$

where θ_n and φ_n denote the phases of $\boldsymbol{x}(n)$ and $\boldsymbol{c}^{(t)}(n)$, respectively. Obviously, the closed-form solution of (C2) is $\theta_n^\star=2\pi/Lp$, with

$$p=\left\lfloor\frac{\varphi_n}{\frac{2\pi}{L}}\right\rceil,\tag{C3}$$

with $\lfloor\cdot\rceil$ denotes the floor operation. Therefore, the solution to $\mathcal{P}_{\boldsymbol{x}^{(t)}}$ can be efficiently obtained by

$$\boldsymbol{x}^{(t)}=\left(1/\sqrt{NM}\right)e^{j\theta^\star},\tag{C4}$$

where, $\theta^\star=[\theta_1^\star,\theta_2^\star,\cdots,\theta_N^\star]^T$.

Appendix D The procedure to tackle $\mathcal{P}_{\boldsymbol{y}^{(t)}}$

Given $\boldsymbol{x}^{(t)},\boldsymbol{s}^{(t)},\boldsymbol{u}_p^{(t-1)},\varrho_p^{(t-1)}$, $p\in\{1,2\}$, then the Problem $\mathcal{P}_{\boldsymbol{y}^{(t)}}$ with respect to $\boldsymbol{y}$ can be rewritten as:

$$\mathcal{P}_{\boldsymbol{y}}\begin{cases}\min\limits_{\boldsymbol{y}} & \left\|\boldsymbol{y}-\boldsymbol{v}^{(t)}\right\|^2\\ \text{s.t.} & \boldsymbol{y}^\dagger\left(\boldsymbol{I}_M\otimes\boldsymbol{R}_I\right)\boldsymbol{y}\leqslant E_I,\end{cases}\tag{D1}$$

where $\boldsymbol{v}^{(t)}=\boldsymbol{s}^{(t-1)}-\boldsymbol{u}_2^{(t-1)}/\varrho_2^{(t-1)}$. Problem (D1) is a QCQP that can be solved via the ADMM algorithm [10].

To efficiently update $\boldsymbol{y}$, let the eigen-decomposition of $(\boldsymbol{I}_M\otimes\boldsymbol{R}_I)$ be $\left(\boldsymbol{U}_I\boldsymbol{\Lambda}\boldsymbol{U}_I^\dagger\right)$ with $\boldsymbol{\Lambda}$ diagonal real and $\boldsymbol{U}_I$ consists of the eigenvectors. And the Problem (D1) can be equivalently converted as

$$\mathcal{P}_{\bar{\boldsymbol{y}}}\begin{cases}\min\limits_{\bar{\boldsymbol{y}}} & \left\|\bar{\boldsymbol{y}}-\bar{\boldsymbol{v}}^{(t)}\right\|^2\\ \text{s.t.} & \bar{\boldsymbol{y}}^\dagger\boldsymbol{\Lambda}\bar{\boldsymbol{y}}\leqslant E_I,\end{cases}\tag{D2}$$

where $\bar{\boldsymbol{y}}=\boldsymbol{U}_I^\dagger\boldsymbol{y}$ and $\bar{\boldsymbol{v}}^{(t)}=\boldsymbol{U}_I^\dagger\boldsymbol{v}^{(t)}$.

The corresponding Lagrangian can be expressed as

$$L\left(\bar{\boldsymbol{y}},\eta\right)=\left\|\bar{\boldsymbol{y}}-\bar{\boldsymbol{v}}^{(t)}\right\|^2+\eta\left(\bar{\boldsymbol{y}}^\dagger\boldsymbol{\Lambda}\bar{\boldsymbol{y}}-E_I\right),\tag{D3}$$

where η denotes the Lagrangian multiplier associated with the spectral constraint.

According to the Karush-Kuhn-Tucker(KKT) condition, the optimal solution should satisfies

$$\nabla L(\bar{\boldsymbol{y}},\eta) = 2\left(\bar{\boldsymbol{y}} - \bar{\boldsymbol{v}}^{(t)}\right) + 2\eta\bar{\boldsymbol{y}}^{\dagger}\boldsymbol{\Lambda} = \boldsymbol{0}. \tag{D4}$$

Therefore, it can be derived that $\bar{\boldsymbol{y}} = \boldsymbol{E}^{-1}\bar{\boldsymbol{v}}^{(t)}$, where, $\boldsymbol{E}=(\boldsymbol{I}_{NM} + \eta\boldsymbol{\Lambda})$. Then, to get the solution of η, a nonlinear equation should be solved as

$$\bar{\boldsymbol{v}}^{(t)\dagger}\boldsymbol{E}^{-1}\boldsymbol{\Lambda}\boldsymbol{E}^{-1}\bar{\boldsymbol{v}}^{(t)} = E_I, \tag{D5}$$

which can be rewritten as

$$\sum_{k=1}^{NM}\frac{\lambda_k}{(1+\eta\lambda_k)^2}\left|\bar{v}_k\right|^2 = E_I, \tag{D6}$$

while λ_k and $\bar{v}_k, k \in \mathcal{N}$ are the eigenvalues of the $(\boldsymbol{I}_M \otimes \boldsymbol{R}_I)$ and the elements of $\bar{\boldsymbol{v}}^{(t)}$. Herein, η can be efficiently obtained by handling the Problem (D6) via bisection or Newton's method.

To this end, plugging η back to get $\bar{\boldsymbol{y}}$, the solution to Problem (D1) can be simply given by $\boldsymbol{y} = \boldsymbol{U}_I\bar{\boldsymbol{y}}$.

Appendix E The procedure to tackle $\mathcal{P}_{\boldsymbol{S}^{(t)}}$

Given $\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(t-1)}, \varrho_p^{(t-1)}$, $p \in \{1,2\}$, then Problem $\mathcal{P}_{\boldsymbol{S}^{(t)}}$ with respect to $\boldsymbol{s}$ can be recast as

$$\mathcal{P}_{\boldsymbol{S}}\begin{cases}\min\limits_{\boldsymbol{s}} & f(\boldsymbol{s}) + 1/2\|\boldsymbol{s} - \boldsymbol{q}^{(t)}\|^2\\ \text{s.t.} & |\boldsymbol{s}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}.\end{cases} \tag{E1}$$

where $\boldsymbol{q}^{(t)} = \boldsymbol{u}_1^{(t-1)} + \boldsymbol{u}_2^{(t-1)} + \varrho_1^{(t-1)}\boldsymbol{x}^{(t)} + \varrho_2^{(t-1)}\boldsymbol{y}^{(t)}$.

The aperiodic cross-correlation of $\boldsymbol{s}_{m1}$ and $\boldsymbol{s}_{m2}$ at the kth lag is

$$r_{m_1m_2}(k) = \boldsymbol{s}_{m_2}^{\dagger}\boldsymbol{J}_k\boldsymbol{s}_{m_1}, \tag{E2}$$

where $k \in \{0,1,\ldots N-1\}$, and $\boldsymbol{J}_k \in \mathbb{R}^{N\times N}$ is a shift matrix[1]. Suppose γ_k and w_k are symmetry with respect to the origin, then the WISL can be rewritten as

$$f(\boldsymbol{s}) = f_0(\boldsymbol{s}) + f_1(\boldsymbol{s}) + f_2(\boldsymbol{s}) - w_0MN^2, \tag{E3}$$

where

$$f_0(\boldsymbol{s}) = 2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\left|\boldsymbol{s}_m^{\dagger}\boldsymbol{J}_k\boldsymbol{s}_m\right|^2, \tag{E4}$$

$$f_1(\boldsymbol{s}) = 2\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\sum_{k=1}^{N-1}w_k\left|\boldsymbol{s}_{m_1}^{\dagger}\boldsymbol{J}_k\boldsymbol{s}_{m_2}\right|^2, \tag{E5}$$

$$f_2(\boldsymbol{s}) = \sum_{m_1=1}^{M}\sum_{m_2=1}^{M}w_0\left|\boldsymbol{s}_{m_1}^{\dagger}\boldsymbol{s}_{m_2}\right|^2. \tag{E6}$$

Proposition 1. As the first step toward the development of an effective procedure to handle $\mathcal{P}_{\boldsymbol{S}}$, the $f_0(\mathbf{s})$, $f_1(\mathbf{s})$ and $f_2(\mathbf{s})$ can be equivalently recast as

$$f_0(\boldsymbol{s}) = \text{vec}\left(\boldsymbol{s}\boldsymbol{s}^{\dagger}\right)^{\dagger}\boldsymbol{L}\text{vec}\left(\boldsymbol{s}\boldsymbol{s}^{\dagger}\right), \tag{E7}$$

$$f_1(\mathbf{s}) = \boldsymbol{s}^{\dagger}\boldsymbol{M}(\boldsymbol{S})\boldsymbol{s}, \tag{E8}$$

$$f_2(\boldsymbol{s}) = \sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\gamma_0\boldsymbol{s}_{m_1}^{\dagger}\boldsymbol{s}_{m_2}\mathbf{s}_{m_2}^{\dagger}\boldsymbol{s}_{m_1} = w_0\boldsymbol{s}^{\dagger}(\boldsymbol{I}_M \otimes \boldsymbol{S}\boldsymbol{S}^{\dagger})\boldsymbol{s}, \tag{E9}$$

where

$$\boldsymbol{L}=2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\text{vec}\left(\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\right)\text{vec}\left(\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\right)^{\dagger}, \tag{E10}$$

$$\boldsymbol{M}(\boldsymbol{S}) = \boldsymbol{I}_M \otimes (\boldsymbol{R}_1(\boldsymbol{S}) + \boldsymbol{R}_2(\boldsymbol{S})). \tag{E11}$$

$\boldsymbol{U}_m \in \mathbb{R}^{N\times NM}$ is the block selection matrix defined as

$$\boldsymbol{U}_m = \left[\boldsymbol{0}_{N\times(m-1)N}, \boldsymbol{I}_N, \boldsymbol{0}_{N\times(M-m)N}\right], \tag{E12}$$

with $m \in \mathcal{M}$, and

$$\boldsymbol{R}_1(\boldsymbol{S}) = \sum_{k=1}^{N-1}w_k\boldsymbol{J}_k\boldsymbol{S}\boldsymbol{S}^{\dagger}\boldsymbol{J}_k^T, \tag{E13}$$

$$\boldsymbol{R}_2(\boldsymbol{S}) = \sum_{k=1}^{N-1}w_k\boldsymbol{J}_k^T\boldsymbol{S}\boldsymbol{S}^{\dagger}\boldsymbol{J}_k. \tag{E14}$$

■

1) $\boldsymbol{J}_k(a,b) = 1$ if $b - a = k$, else $\boldsymbol{J}_k(a,b) = 0$, $a, b \in \{1, 2, \cdots, N\}$

Proof. By utilizing the block selection matrix $\boldsymbol{U}_m, m \in \mathcal{M}$, $f_0(\boldsymbol{s})$ can be derived as

$$\begin{aligned} f_0(\boldsymbol{s}) =& 2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\left|\boldsymbol{s}^\dagger\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\boldsymbol{s}\right|^2 \\ =& 2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\left|\text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right)^\dagger\text{vec}\left(\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\right)\right|^2 \\ =& \text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right)^\dagger\boldsymbol{L}\text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right), \end{aligned} \tag{E15}$$

and when $\gamma_k = w_k, k \in \{1\cdots N-1\}$, $f_0(\boldsymbol{s}) = 0$.

For an arbitrary matrix $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots \boldsymbol{s}_M] \in \mathbb{C}^{N\times M}$, it satisfies

$$\sum_{m=1}^{M}\boldsymbol{s}_m^\dagger\boldsymbol{A}\boldsymbol{s}_m = \boldsymbol{s}^\dagger\left(\boldsymbol{I}_M \otimes \boldsymbol{A}\right)\boldsymbol{s}. \tag{E16}$$

According to (E16) and $\sum_{m=1}^{M}\boldsymbol{s}_m\boldsymbol{s}_m^\dagger = \boldsymbol{S}\boldsymbol{S}^\dagger$, $f_2(\boldsymbol{s})$ can be derived as $f_2(\boldsymbol{s}) = w_0\boldsymbol{s}^\dagger\left(\boldsymbol{I}_M \otimes \boldsymbol{S}\boldsymbol{S}^\dagger\right)\boldsymbol{s}$, and $f_1(\boldsymbol{s})$ in (E8) can be expressed as

$$\begin{aligned} f_1(\boldsymbol{s}) =& \boldsymbol{s}^\dagger(\boldsymbol{I}_M \otimes (\sum_{k=1}^{N-1}w_k\boldsymbol{J}_k\boldsymbol{S}\boldsymbol{S}^\dagger\boldsymbol{J}_k^T))\boldsymbol{s} + \boldsymbol{s}^\dagger(\boldsymbol{I}_M \otimes (\sum_{k=1}^{N-1}w_k\boldsymbol{J}_k^T\mathbf{S}\boldsymbol{S}^\dagger\boldsymbol{J}_k))\boldsymbol{s} \\ =& \boldsymbol{s}^\dagger(\boldsymbol{I}_M \otimes (\boldsymbol{R}_1(\boldsymbol{S}) + \boldsymbol{R}_2(\boldsymbol{S})))\boldsymbol{s} \\ =& \boldsymbol{s}^\dagger\boldsymbol{M}(\boldsymbol{S})\boldsymbol{s}. \end{aligned} \tag{E17}$$

■

To this end, the objective function $f(\mathbf{s})$ can be expressed as

$$\begin{aligned} f(\boldsymbol{s}) =& f_0(\boldsymbol{s}) + f_1(\boldsymbol{s}) + f_2(\boldsymbol{s}) - w_0MN^2 \\ =& \text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right)^\dagger\boldsymbol{L}\text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right) + \boldsymbol{s}^\dagger\mathbf{Q}(\boldsymbol{S})\boldsymbol{s} - w_0MN^2, \end{aligned} \tag{E18}$$

where

$$\boldsymbol{Q}(\boldsymbol{S}) = \boldsymbol{M}(\boldsymbol{S}) + w_0\boldsymbol{I}_M \otimes \boldsymbol{S}\boldsymbol{S}^\dagger. \tag{E19}$$

Inspection on (E18) exhibits it is a quartic optimization problem with constant modulus restriction. As ISQO algorithm [11] enjoys a fast convergence speed, it is introduced to solve Problem (E1) in this subsection.

In particular, the high polynomial Problem (E1) can be solved by ISQO according to the following proposition.

Proposition 2. Problem (E1) can be handled by sequentially solving the following approximation

$$\mathcal{P}_{(i+1)}\begin{cases} \min\limits_{\boldsymbol{s}} \; l\left(\boldsymbol{s}, \boldsymbol{s}_{(i)}\right) \\ \text{s.t.} \;\; |\boldsymbol{s}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}, \end{cases} \tag{E20}$$

where $\boldsymbol{s}_{(i)}$ denotes the ith iteration solution of ISQO and

$$\begin{aligned} l\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right) =& f_0\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right) + f_1\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right) + f_2\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right) \\ &+ \Re\left\{\frac{1}{2}(\boldsymbol{s}_{(i)} - \boldsymbol{q}^{(t)})^\dagger(\boldsymbol{s} - \boldsymbol{q}^{(t)})\right\} \\ =& \text{vec}\left(\boldsymbol{s}\boldsymbol{s}^\dagger\right)^\dagger\boldsymbol{L}\text{vec}\left(\boldsymbol{s}_{(i)}{\boldsymbol{s}_{(i)}}^\dagger\right) + \boldsymbol{s}^\dagger\boldsymbol{Q}(\boldsymbol{S}_{(i)})\boldsymbol{s} \\ &+ \Re\left\{\frac{1}{2}(\boldsymbol{s}_{(i)} - \boldsymbol{q}^{(t)})^\dagger(\boldsymbol{s} - \boldsymbol{q}^{(t)})\right\}. \end{aligned} \tag{E21}$$

■

Some insights of the proposition are given next.

Remark 1. The iterative procedure based on ISQO is yielded to solve problem (E1) more easily by fixing a part of the variable $\boldsymbol{s}$ in the quartic function, then the non-convex quartic objective function reduces to a quadratic convex one at each iteration. Herein, its convergence is proved and a good solution would be provided under some conditions (Please refer to Chapter 9 in [12]).

Appendix E.1 The solution to Problem (E20)

For a given $\boldsymbol{s}_{(i)}$, $\mathcal{P}_{(i+1)}$ is a quadratic problem with respect to $\boldsymbol{s}$. More specifically, $f_0\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right)$ can be converted as

$$\begin{aligned} & f_0\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right) \\ =& 2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\boldsymbol{s}^\dagger\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\boldsymbol{s}\left({\boldsymbol{s}_i}^\dagger\boldsymbol{U}_m^T{\boldsymbol{J}_k}^\dagger\boldsymbol{U}_m\boldsymbol{s}_{(i)}\right) \\ =& 2\sum_{m=1}^{M}\sum_{k=1}^{N-1}(\gamma_k - w_k)\,r_{mm}^{(i)}(-k)\,\boldsymbol{s}^\dagger\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m\boldsymbol{s} \\ =& \boldsymbol{s}^\dagger\boldsymbol{G}_{(i)}\boldsymbol{s}, \end{aligned} \tag{E22}$$

where

$$\boldsymbol{G}_{(i)}=2\sum_{m=1}^{M}\sum_{k=1}^{N-1}\left(\gamma_k-w_k\right)r_{mm}^{(i)}\left(-k\right)\boldsymbol{U}_m^T\boldsymbol{J}_k\boldsymbol{U}_m. \tag{E23}$$

Then, Problem (E20) can be equivalently expressed as

$$\mathcal{P}_{\boldsymbol{s}_1}\begin{cases}\max\limits_{\boldsymbol{s}} & \lambda\boldsymbol{s}^\dagger\boldsymbol{s}-l\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right)\\ \text{s.t.} & |\boldsymbol{s}(n)|=1/\sqrt{NM}, n\in\mathcal{N},\end{cases} \tag{E24}$$

where $\lambda\geqslant\lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{S}_{(i)})\right)+\lambda_{\max}\left(\boldsymbol{G}_{(i)}\right)$.

Since the objective function in Problem $\mathcal{P}_{\boldsymbol{s}_1}$ is convex, its first-order Taylor expansion is

$$\begin{aligned}&\lambda\boldsymbol{s}^\dagger{}_{(i)}\boldsymbol{s}_{(i)}-l\left(\boldsymbol{s}_{(i)},\boldsymbol{s}_{(i)}\right)\\&\quad+\Re\left\{\nabla^\dagger\left(\lambda\boldsymbol{s}^\dagger\boldsymbol{s}-l\left(\boldsymbol{s},\boldsymbol{s}_{(i)}\right)\right)|_{\boldsymbol{s}=\boldsymbol{s}_{(i)}}\left(\boldsymbol{s}-\boldsymbol{s}_{(i)}\right)\right\}\\=&\lambda\boldsymbol{s}^\dagger{}_{(i)}\boldsymbol{s}_{(i)}-l\left(\boldsymbol{s}_{(i)},\boldsymbol{s}_{(i)}\right)\\&\quad+2\Re\left\{\left[\left(\lambda\boldsymbol{I}_{NM}-\boldsymbol{Q}(\boldsymbol{S}_{(i)})-\boldsymbol{G}_{(i)}\right)\boldsymbol{s}_{(i)}\right]^\dagger\left(\boldsymbol{s}-\boldsymbol{s}_{(i)}\right)\right\}\\&\quad+\frac{1}{2}\Re\left\{\left[\left(\boldsymbol{s}_{(i)}-\boldsymbol{q}^{(t)}\right)\right]^\dagger\left(\boldsymbol{s}-\boldsymbol{s}_{(i)}\right)\right\}.\end{aligned} \tag{E25}$$

Omitting the constant terms, $\mathcal{P}_{(i+1)}$ can be solved by

$$\mathcal{P}_{\boldsymbol{s}_{(i+1)}}\begin{cases}\max\limits_{\boldsymbol{s}} & \Re\left\{\boldsymbol{z}_{(i)}^\dagger\boldsymbol{s}\right\}\\ \text{s.t.} & |\boldsymbol{s}(n)|=1/\sqrt{NM}, n\in\mathcal{N},\end{cases} \tag{E26}$$

where

$$\boldsymbol{z}_{(i)}=\left(\lambda\boldsymbol{I}_{NM}-\boldsymbol{Q}(\boldsymbol{S}_{(i)})-\boldsymbol{G}_{(i)}\right)\boldsymbol{s}_{(i)}-\frac{1}{4}\left(\boldsymbol{s}_{(i)}-\boldsymbol{q}^{(t)}\right). \tag{E27}$$

Therefore, the closed-form solution of $\mathcal{P}_{\boldsymbol{s}_{(i+1)}}$ is derived as

$$\boldsymbol{s}_{(i+1)}=\left(1/\sqrt{NM}\right)e^{j\arg\left(\boldsymbol{z}_{(i)}\right)}. \tag{E28}$$

Observing the above equations, the computational cost to solve Problem $\mathcal{P}_{(i+1)}$ is dominated by the computation of $\boldsymbol{z}_{(i)}$. Precisely, it refers to the compute of $\boldsymbol{Q}(\boldsymbol{S}_{(i)})\boldsymbol{s}_{(i)}$ and $\boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)}$. Next, two special structures are explored to obtain an efficient implementation.

The calculation of $\boldsymbol{Q}(\boldsymbol{S})$ involves the computation of $\boldsymbol{R}_1\left(\boldsymbol{S}\right)$ and $\boldsymbol{R}_2\left(\boldsymbol{S}\right)$. To save the computational time, they can be simplified by using a fixed sparse matrix $\boldsymbol{\Omega}$ as follows

$$\begin{aligned}\boldsymbol{R}_1\left(\boldsymbol{S}\right)&=\text{mat}\left\{\text{vec}\left(\sum_{k=1}^{N-1}w_k\boldsymbol{J}_k\boldsymbol{S}\boldsymbol{S}^\dagger\boldsymbol{J}_k^T\right)\right\}\\&=\text{mat}\left\{\boldsymbol{\Omega}\text{vec}\left(\boldsymbol{S}\boldsymbol{S}^\dagger\right)\right\},\end{aligned} \tag{E29}$$

$$\begin{aligned}\boldsymbol{R}_2\left(\boldsymbol{S}\right)&=\text{mat}\left\{\text{vec}\left(\sum_{k=1}^{N-1}w_k\boldsymbol{J}_k^T\boldsymbol{S}\boldsymbol{S}^\dagger\boldsymbol{J}_k\right)\right\}\\&=\text{mat}\left\{\boldsymbol{\Omega}^T\text{vec}\left(\boldsymbol{S}\boldsymbol{S}^\dagger\right)\right\},\end{aligned} \tag{E30}$$

where

$$\boldsymbol{\Omega}=\sum_{k=1}^{N-1}w_k(\boldsymbol{J}_k\otimes\boldsymbol{J}_k). \tag{E31}$$

Proposition 3. Fast Fourier transform(FFT) operations can be applied to the matrix vector multiplication $\boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)}$. ■

Proof. It is worth noting that $\boldsymbol{G}_{(i)}$ can be written as the block matrix

$$\boldsymbol{G}_{(i)}=\begin{bmatrix}\boldsymbol{G}_{1(i)} & \boldsymbol{0}_{N\times N} & \cdots & \boldsymbol{0}_{N\times N}\\ \boldsymbol{0}_{N\times N} & \boldsymbol{G}_{2(i)} & \cdots & \boldsymbol{0}_{N\times N}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{0}_{N\times N} & \boldsymbol{0}_{N\times N} & \cdots & \boldsymbol{G}_{M(i)}\end{bmatrix}, \tag{E32}$$

where each block is given by

$$\boldsymbol{G}_{m(i)}=2\sum_{k=1}^{N-1}\left(\gamma_k-w_k\right)r_{mm}^{(i)}\left(-k\right)\boldsymbol{J}_k, m\in\mathcal{M}. \tag{E33}$$

Obviously, $\boldsymbol{G}_{m(i)}, m\in\mathcal{M}$ are Toeplitz matrices.

Lemma 1. Let $\boldsymbol{T}$ be an $N \times N$ Toeplitz matrix defined as

$$\boldsymbol{T} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{N-1} \\ t_{-1} & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{1-N} & \cdots & t_{-1} & t_0 \end{bmatrix}, \tag{E34}$$

and $\boldsymbol{F} \in \mathbb{C}^{2N\times 2N}$ be a FFT matrix with $F(m,n) = e^{-j\frac{2mn\pi}{2N}}, 0 \leqslant m,n \leqslant 2N$. Then, $\boldsymbol{T}$ can be decomposed as $\boldsymbol{T} = \frac{1}{2N}\boldsymbol{H}^{\dagger}\mathrm{Diag}\left(\boldsymbol{F}\boldsymbol{c}\right)\boldsymbol{H}$, where

$$\boldsymbol{H} = \boldsymbol{F} \cdot \begin{bmatrix} \boldsymbol{I}_N \\ \boldsymbol{0}_N \end{bmatrix}, \tag{E35}$$

and $\boldsymbol{c} = [t_0, t_{-1}, \cdots t_{1-N}, 0, t_{N-1}, \cdots t_1]^T$. ■

According to **Lemma 1**, $\boldsymbol{G}_{m(i)}$ can be recast as

$$\boldsymbol{G}_{m(i)} = \frac{1}{2N}\boldsymbol{H}^{\dagger}\mathrm{Diag}\left(\boldsymbol{F}\boldsymbol{g}_{m(i)}\right)\boldsymbol{H}, \tag{E36}$$

where

$$\begin{aligned} \boldsymbol{g}_{m(i)} = 2\Big[&\boldsymbol{0}_{1\times(N+1)}, \left(\gamma_{N-1} - w_{N-1}\right) r_{mm}^{(i)}\left(1-N\right), \\ &\cdots\left(\gamma_1 - w_1\right) r_{mm}^{(i)}\left(-1\right)\Big]^T. \end{aligned} \tag{E37}$$

Therefore, the matrix vector multiplication $\boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)}$ can be performed as

$$\boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)} = \left[\left(\boldsymbol{G}_{1(i)}\boldsymbol{U}_1\boldsymbol{s}_{(i)}\right)^T, \ldots, \left(\boldsymbol{G}_{m(i)}\boldsymbol{U}_m\boldsymbol{s}_{(i)}\right)^T, \ldots, \left(\boldsymbol{G}_{M(i)}\boldsymbol{U}_M\boldsymbol{s}_{(i)}\right)^T\right]^T. \tag{E38}$$

■

For computational efficiency, suppose $\lambda = \lambda_1 + \lambda_2$, where

$$\lambda_1 \geqslant \lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{S}_{(i)})\right), \tag{E39}$$

$$\lambda_2 \geqslant \lambda_{\max}\left(\boldsymbol{G}_{(i)}\right). \tag{E40}$$

Hence, on the basis of [13], λ_1 can be set as a constant

$$\lambda_1 = NM\sum_{k=-N+1}^{N-1} w_k. \tag{E41}$$

While, λ_2 can be assigned to $\left\|\boldsymbol{G}_{(i)}\right\|_{\infty} = \max\limits_{m=1,\cdots NM}\sum\limits_{n=1}^{NM}\left|\boldsymbol{G}_{(i)}\left(m,n\right)\right|$ avoiding the eigenvalue decomposition of $\boldsymbol{G}_{(i)}$ to find $\lambda_{\max}\left(\boldsymbol{G}_{(i)}\right)$.

Similarly, $\left\|\boldsymbol{G}_{(i)}\right\|_{\infty}$ can be computed efficiently via FFT operations since it can equivalently expressed as

$$\left\|\boldsymbol{G}_{(i)}\right\|_{\infty} = \max\left(\bar{\boldsymbol{G}}_{(i)}\boldsymbol{1}_{N\times 1}\right), \tag{E42}$$

where $\boldsymbol{1}_{N\times 1}$ is the vector with all element values be 1, and

$$\bar{\boldsymbol{G}}_{(i)}(a,b) = \left|\boldsymbol{G}_{(i)}(a,b)\right|, a,b \in \{1,2,\cdots,N\}. \tag{E43}$$

Appendix E.2 ISQO procedure for solving $\mathcal{P}_{\boldsymbol{S}^{(t)}}$

The procedure of solving $\mathcal{P}_{\boldsymbol{S}^{(t)}}$ is summarized in **Algorithm E1**.

Since the objective function of $\mathcal{P}_{(i+1)}$ is not optimized directly but is surrogated by its first-order Taylor expansion, the optimization process suffers from a slow convergence speed. To get a solution to Problem (E1) with a fast convergence speed, the Squared Iterative Method (SQUAREM) [14] which is originally proposed to accelerate any Expectation Maximization(EM) algorithms can be applied. It is an acceleration method via a two-point iteration strategy that no extra parameter is required, except possibly the computationally projection to the feasibility set, and it is guaranteed to converge.

Appendix F Comments about the IADPM framework

Without theoretical convergence guarantee, the Alternating Direction Method of Multipliers (ADMM) requires a suitable choice of the penalty parameter and initialization to ensure convergence (see Chapter 9 in [4]) for the non-convex optimization. Motivated by the method of multipliers [5–9], the IADPM framework is proposed to avoid the tuning of the penalty factor and initialization in ADMM. Unlike ADMM with fixed penalty factor ϱ_p and multiplier vector $\boldsymbol{u}_p$, the IADPM framework updates $\boldsymbol{u}_p^{(t)}$ and $\varrho_p^{(t)}$ based on the primal residual $\triangle r_p^{(t)}$, $p \in \{1,2\}$ at each iteration. More specifically, to find a feasible point with a higher possibility, a larger penalty factor $\varrho_p^{(t)}$ would be utilized to make $\triangle r_p^{(t)}$ approach zero if it does not decrease with the iteration number. Otherwise, $\varrho_p^{(t)}$ remains unchanged. Thus the proposed IADPM provides a higher probability to guarantee the convergence starting from a small value $\varrho_p^{(0)}$, which makes it possible for providing a better objective value than ADMM.

However, some limitations of the IADPM framework are still existed. Starting from small valus $\varrho_p^{(0)}$, the proposed IADPM framework needs more iterations to find the feasible penalty parameters ϱ_p and multiplier vectors $\boldsymbol{u}_p$ with $p \in \{1,2\}$. Therefore, the presented IADPM framework is more suitable for high computing performance platform but has limited value to some applications that requires high real-time performance. Meanwhile, for the case of unbounded penalty parameters, the convergence of the IADPM cannot be proved theoretically.

Algorithm E1 : ISQO algorithm for solving $\mathcal{P}_{\boldsymbol{s}^{(t)}}$.

Require: N, M, w_k, γ_k, $\boldsymbol{x}^{(t)}$, $\boldsymbol{y}^{(t)}$, $\varrho_p^{(t-1)}$, $\boldsymbol{u}_p^{(t-1)}$, $\boldsymbol{s}^{(t-1)}$;
Ensure: An optimal solution $\boldsymbol{s}^{(t)}$;
  $i=0$, initialize $\boldsymbol{s}_{(0)} = \boldsymbol{s}^{(t-1)}$;
  Get $\boldsymbol{\Omega}$ via (E31);
  Compute λ_1 according to (E41);
  repeat
    Get $\boldsymbol{Q}(\boldsymbol{S}_{(i)}) = \boldsymbol{I}_M \otimes (\boldsymbol{R}_1(\boldsymbol{S}_{(i)}) + \boldsymbol{R}_2(\boldsymbol{S}_{(i)}) + w_0 \boldsymbol{S}_{(i)} \boldsymbol{S}_{(i)}^{\dagger})$ through (E29) and (E30);
    Compute $\boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)}$ according to (E38);
    Get λ_2 based on (E42), and set $\lambda = \lambda_1 + \lambda_2$;
    Compute $\boldsymbol{q}^{(t)} = \boldsymbol{u}_1^{(t-1)} + \boldsymbol{u}_2^{(t-1)} + \varrho_1^{(t-1)}\boldsymbol{x}^{(t)} + \varrho_2^{(t-1)}\boldsymbol{y}^{(t)}$;
    $\boldsymbol{z}_{(i)} = (\lambda \boldsymbol{I}_{NM} - \boldsymbol{Q}(\boldsymbol{S}_{(i)}))\boldsymbol{s}_{(i)} - \boldsymbol{G}_{(i)}\boldsymbol{s}_{(i)} - \frac{1}{4}(\boldsymbol{s}_{(i)} - \boldsymbol{q}^{(t)})$;
    $\boldsymbol{s}_{(i+1)} = (1/\sqrt{NM}) e^{j \arg(\boldsymbol{z}_{(i)})}$;
    $i \leftarrow i+1$
  until convergence

Appendix G Convergence Analysis

The proposed IADPM is guaranteed to converge for arbitrary initializations $\boldsymbol{s}^{(0)}, \boldsymbol{u}_p^{(0)}$ provided that $\varrho_p^{(0)} > 0$ and ϱ_p are bounded, $p \in \{1,2\}$, $0 < \delta_{1,c} < 1$, $\delta_{2,c} > 1$ and a large enough number ν. Next, we will prove the convergence of the IADPM framework.

As can be seen from the definition of $\varrho_p^{(t)}$, $\{\varrho_p^{(t)}\}_{t=0}^{\infty}$ is non-decreasing sequences. Therefore, the two cases are $\varrho_p^{(t)}$ are bounded and $\varrho_p^{(t)}$ grow to infinity with iteration number t.

When the penalty parameters $\varrho_p^{(t)}$ are bounded, there exists $t \geqslant \tilde{t}$ such that $\varrho_p^{(t)}$ are constants, where $\tilde{t}$ is a sufficiently large number. In particular,

$$\varrho_p^{(t)} = \varrho_p^{(t-1)} = \varrho_p^{(*)} \neq \infty, \text{for } t \geqslant \tilde{t}. \tag{G1}$$

From the definition of $\varrho_p^{(t)}$, it is required to satisfy $\triangle r^{(t)} \leqslant \delta_{1,c} \triangle r^{(t-1)}$ when $t \geqslant \tilde{t}$. Hence, it can be derived that

$$\triangle r_p^{(t)} \leqslant \delta_{1,c} \triangle r_p^{(t-1)} \leqslant \cdots \leqslant \delta_{1,c}^{t-1} \triangle r_p^{(1)}, \text{for } t \geqslant \tilde{t}. \tag{G2}$$

Since $\delta_{1,c} < 1$, then

$$\lim_{t\to\infty} \delta_{1,c}^{t-1} = 0. \tag{G3}$$

As $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{s}$ are constant modulus sequences, $\triangle r_1^{(1)} = \|\boldsymbol{x}^{(1)} - \boldsymbol{s}^{(1)}\| < \infty$, and $\triangle r_2^{(1)} = \|\boldsymbol{y}^{(1)} - \boldsymbol{s}^{(1)}\| < \infty$. Hence, combining (G2) and (G3), it can be obtained that

$$\lim_{t\to\infty} \triangle r_p^{(t)} = 0. \tag{G4}$$

It follows that

$$\begin{aligned} &\lim_{t\to\infty} \left\| \boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)} \right\| = 0 \\ &\lim_{t\to\infty} \left(\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)} \right) = \boldsymbol{0}_{N\times 1}, \end{aligned} \tag{G5}$$

and

$$\begin{aligned} &\lim_{t\to\infty} \left\| \boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)} \right\| = 0 \\ &\lim_{t\to\infty} \left(\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)} \right) = \boldsymbol{0}_{N\times 1}. \end{aligned} \tag{G6}$$

According to the definition of $\boldsymbol{u}_p^{(t)}$, if $u_{p\max}^{(t)} \leqslant \nu$, $\boldsymbol{u}_1^{(t)} - \boldsymbol{u}_1^{(t-1)} = \varrho_1^{(t)}(\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)})$ and $\boldsymbol{u}_2^{(t)} - \boldsymbol{u}_2^{(t-1)} = \varrho_2^{(t)}(\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)})$. Based on (G1), (G5) and (G6), it can be obtained that

$$\lim_{t\to\infty} (\boldsymbol{u}_1^{(t-1)} - \boldsymbol{u}_1^{(t)}) = \lim_{t\to\infty} \varrho_1^{(t)}(\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)}) = \varrho_1^{(*)} \times \boldsymbol{0}_{N\times 1} = \boldsymbol{0}_{N\times 1}. \tag{G7}$$

$$\lim_{t\to\infty} (\boldsymbol{u}_2^{(t-1)} - \boldsymbol{u}_2^{(t)}) = \lim_{t\to\infty} \varrho_1^{(t)}(\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)}) = \varrho_2^{(*)} \times \boldsymbol{0}_{N\times 1} = \boldsymbol{0}_{N\times 1}. \tag{G8}$$

Besides, according to (G7) and (G8), it is shown that there exists $t \geqslant \bar{t} \geqslant \tilde{t}$ and two very small positive constant ϵ_1 and ϵ_2 such that $\varrho_1^{(*)}|\boldsymbol{x}^{(t)}(n) - \boldsymbol{s}^{(t)}(n)| < \epsilon_1$, $\varrho_2^{(*)}|\boldsymbol{y}^{(t)}(n) - \boldsymbol{s}^{(t)}(n)| < \epsilon_2, n \in \mathcal{N}$, where $\bar{t}$ is a sufficiently large number. Hence, when $\boldsymbol{u}_1^{(t)} = [\boldsymbol{u}_1^{(t-1)} + \varrho_1^{(*)}(\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)})]/u_{\max}^{(t)}$, it can be derived that

$$\begin{aligned} |\boldsymbol{u}_1^{(t+1)}(n)| &= |\boldsymbol{u}_1^{(t)}(n) + \varrho_1^{(*)}(\boldsymbol{x}^{(t+1)}(n) - \boldsymbol{s}^{(t+1)}(n))| \\ &\leqslant 1 + \varrho_1^{(*)}\epsilon_1 < \nu, n \in \mathcal{N}, \end{aligned} \tag{G9}$$

for $t \geqslant \bar{t} \geqslant \tilde{t}$. Therefore, $\boldsymbol{u}_1^{(t+1)} = \boldsymbol{u}_1^{(t)} + \varrho_1^{(*)}(\boldsymbol{x}^{(t+1)} - \boldsymbol{s}^{(t+1)})$ can be easily get for the $(t+1)$-th iteration. Similarly, $\boldsymbol{u}_2^{(t+1)} = \boldsymbol{u}_2^{(t)} + \varrho_2^{(*)}(\boldsymbol{y}^{(t+1)} - \boldsymbol{s}^{(t+1)})$. Thus, the same conclusion can be get that $\boldsymbol{u}_p^{(t)}$ are convergent.

To this end, the multiplier vector $\boldsymbol{u}^{(t)}$ converge to $\boldsymbol{u}_p^{(*)}$, i.e.,

$$\lim_{t\to\infty} \boldsymbol{u}_p^{(t-1)} = \lim_{t\to\infty} \boldsymbol{u}_p^{(t)} = \boldsymbol{u}_p^{(*)}. \tag{G10}$$

For $t \geqslant \bar{t} \geqslant \tilde{t}$, the penalty parameter $\varrho_p^{(t)}$ and multiplier vector $\boldsymbol{u}_p^{(t)}$ are bounded to $\varrho_p^{(*)}$ and $\boldsymbol{u}_p^{(*)}$ separately. Hence, the updates of $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{s}$ can be computed through handling the following two problems

$$\begin{aligned} \min_{\boldsymbol{x}} \ & L(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}) \\ \text{s.t.} \ & \arg \boldsymbol{x}(n) \in \Phi, n \in \mathcal{N} \\ & |\boldsymbol{x}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}, \end{aligned} \tag{G11}$$

$$\begin{aligned} \min_{\boldsymbol{y}} \ & L(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}) \\ \text{s.t.} \ & \boldsymbol{y}^\dagger \left(\boldsymbol{I}_M \otimes \boldsymbol{R}_I\right) \boldsymbol{y} \leqslant E_I, \end{aligned} \tag{G12}$$

and

$$\begin{aligned} \min_{\boldsymbol{s}} \ & L(\boldsymbol{s}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}) \\ \text{s.t.} \ & |\boldsymbol{s}(n)| = 1/\sqrt{NM}, n \in \mathcal{N}, \end{aligned} \tag{G13}$$

The objective function decreases monotonically as

$$\begin{aligned} & L\left(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t-1)}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right) \geqslant L\left(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t-1)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right) \geqslant L\left(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right) \\ & \geqslant L\left(\boldsymbol{s}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right). \end{aligned} \tag{G14}$$

Combing (G5) and (G6), it can be derived that

$$\begin{aligned} & \lim_{t\to\infty} L\left(\boldsymbol{s}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right) \\ & = \lim_{t\to\infty} f\left(\boldsymbol{s}^{(t)}\right) + \Re\left\{\boldsymbol{u}_1^{(*)\dagger}\left(\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)}\right)\right\} + \Re\left\{\boldsymbol{u}_2^{(*)\dagger}\left(\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)}\right)\right\} + \varrho_1^{(*)}/2\left\|\boldsymbol{x}^{(t)} - \boldsymbol{s}^{(t)}\right\|^2 + \varrho_2^{(*)}/2\left\|\boldsymbol{y}^{(t)} - \boldsymbol{s}^{(t)}\right\|^2 \\ & = \lim_{t\to\infty} f\left(\boldsymbol{s}^{(t)}\right). \end{aligned} \tag{G15}$$

$f\left(\boldsymbol{s}\right)$ is greater than 0, which means the objective function value has lower bound. Hence, based on (G14) and (G15), it can be concluded that the objective function converges to a finite value,

$$\lim_{t\to\infty} L\left(\boldsymbol{s}^{(t)}, \boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)}, \boldsymbol{u}_p^{(*)}, \varrho_p^{(*)}\right) = \lim_{t\to\infty} f\left(\boldsymbol{s}^{(t)}\right) = f^*, \tag{G16}$$

where f^* is the lower bound of $f\left(\boldsymbol{s}^{(t)}\right)$.

While for the case of unbounded penalty parameter, the convergence of the IADPM cannot be proved theoretically.

References

1 Aubry A, Carotenuto V, D Maio A, Farina A, Pallotta L. Optimization theory-based radar waveform design for spectrally dense environments. IEEE Aerospace and Electronic Systems Magazine, 2016, 31: 14-25

2 Aubry A, Carotenuto V, D Maio A, Govoni M A. Multi-Snapshot Spectrum Sensing for Cognitive Radar via Block-Sparsity Exploitation. IEEE Transaction on signal processing, 2019, 67: 1396-1406

3 Aubry A, Carotenuto V, D Maio A, Govoni M. A, Farina A. Experimental Analysis of Block-Sparsity-Based Spectrum Sensing Techniques for Cognitive Radar. IEEE Transactions on Aerospace and Electronic Systems. 2021, 57: 355-370

4 Boyd S, Parikh N, Chu E, et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. 2010

5 Erseghe T. A distributed and maximum-likelihood sensor network localization algorithm based upon a nonconvex problem formulation. IEEE Transaction on Signal and Information Processing over Networks, 2015, 1: 247-258

6 Bertsekas D. Constrained Optimization and Lagrange Multiplier Methods. New York, NY, USA : Academic, 1982.

7 Magnússon S, Weeraddana P. C, Rabbat M G, Fischione C. On the convergence of alternating direction lagrangian methods for nonconvex structured optimization problems. IEEE Transaction on Control of Network Systems, 2016, 3: 296-309

8 Geißler B, Morsi A, Schewe L, Schmidt M. Penalty alternating direction methods for mixed-integer optimization: A new view on feasibility pumps. SIAM Journal on Optimization, 2017, 27: 1611-1636

9 Yu X X, Cui G L, Yang J, Li J, Kong L J. Quadratic optimization for unimodular sequence design via an ADPM framework. IEEE Transaction on Signal Processing, 2020, 68: 3619-3634

10 Huang K, Sidiropoulos N. D. Consensus-ADMM for General Quadratically Constrained Quadratic Programming. IEEE Trans. Signal Process, 2016, 64: 5297-5310

11 Yang J, Cui G L, Yu X X, Xiao Y, Kong L J. Cognitive local ambiguity function shaping with spectral coexistence. IEEE Access, 2018, 6: 50077-50086

12 Cui G L , De Maio A , Farina A, Li J. Radar Waveform Design Based on Optimization Theory, Hertfordshire, UK: Institution of Engineering and Technology, 2020

13 G. A. F. Seber, A Matrix Handbook for Statisticians. Hoboken, NJ, USA: Wiley, 2008

14 Varadhan R, Roland C. Simple and globally convergent methods for accelerating the convergence of any EM algorithm. Scandinavian Journal of Statistic. 2008, 35: 335-353