Constant Modulus Waveform Estimation and Interference Suppression via Two-stage Fractional Program-based Beamforming

Junli Liang, Tao Wang, Wei Liu, H. C. So, *Fellow*, *IEEE*, Yongwei Huang, and Bo Tang

Abstract—In radar and communication systems, there exist a large class of signals with constant modulus property, including BPSK, QPSK, LFM, and phase-coded signals. In this paper, we focus on the problem of joint constant modulus waveform estimation and interference suppression from signals received at an antenna array. Instead of seeking a compromise between interference suppression and output noise power reduction by the Capon method or utilizing the interference direction (ID) prior to place perfect nulls at the IDs and subsequently minimize output noise power by the linearly constrained minimum variance (LCMV) beamformer, we devise a novel power ratio criterion, namely, interference-plus-noise-to-noise ratio (INNR) in the beamformer output to attain perfect interference nulling and minimal output noise power as in LCMV yet under the unknown ID case. A two-stage fractional program-based method is developed to jointly suppress the interferences and estimate the constant modulus waveform. In the first stage, we formulate an optimization model with a fractional objective function to minimize the INNR. Then, in the second stage, another fraction-constrained optimization problem is established to refine the weight vector from the solution space constrained by the INNR bound, to achieve approximately perfect nulls and minimum output noise power. Moreover, the solution is further extended to tackle the case with steering vector errors. Numerical results demonstrate the excellent performance of our methods.

Index Terms—Waveform estimation and interference suppression (WEIS), Interference-plus-noise-to-noise ratio (INNR), Fractional programming, Constant modulus signal, Interference direction.

I. INTRODUCTION

Waveform estimation or extraction plays an important role in the fields of array signal processing, including radar, sonar, acoustics, astronomy, communications, and medical imaging, and thus has received a lot of attention [1]-[10].

As one of the classic waveform estimation methods, the Capon beamformer [11] adaptively selects the weight vector to minimize its output power subject to a unit-gain constraint on the direction of the signal-of-interest (SOI) so that the SOI does not suffer from any distortion. Although the standard Capon beamformer may attain minimum variance distortionless response (MVDR) on the SOI, the minimum

J. Liang and T. Wang are with School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. (e-mail: liangjunli@nwpu.edu.cn); W. Liu is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, UK; H. C. So is with Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China; Y. Huang is with the School of Information Engi-neering, Guangdong University of Technology, University Town, Guangzhou, Guangdong, China; B. Tang is with College of Electronic Engineering, National University of Defense Technology, China.

output power metric-based objective function essentially seeks a compromise between the interference suppression and output noise power reduction. If the interference directions (IDs) are known, the linear null constraints at the IDs can be introduced into the Capon method to formulate the linearly constrained minimum variance (LCMV) beamformer, achieving minimum output noise power subject to perfect nulls at the IDs [4], [6]. However, in practice, the SOI steering vector may not be accurate due to direction-of-arrival (DOA) or array calibration errors, and in this case, the aforementioned beamformers may treat the SOI as an interference to suppress. To address this issue, many robust beamforming methods have been developed. For example, a robust Capon beamforming method via diagonal loading [12] is presented in [13], where an optimization problem on the steering vector is formulated to first determine the Lagrange multiplier via nonlinear equation root finding, and then obtain the accurate steering vector. Following this work, a robust Capon beamformer with two constraints is introduced in [14] to control the white noise gain; while the total weighted power output of the array is minimized in [15] under the constraint that the gain should exceed unity for all array responses in a given ellipsoid, and the weights are then calculated using the Lagrange multiplier method. In [16], the robust beamforming problem is formulated as a second-order cone program for worst-case performance optimization (WCPO). In [17], with the Capon spectral estimator integrated over a region separated from the desired signal direction, the interference covariance matrix is reconstructed to establish a more robust beamformer than the sample covariance matrix counterpart. In [18], [19], two-sided mainlobe constraints are enforced on the Capon method to enhance its robustness. The resultant beamformer has the mathematical form of a regularized semidefinite programming (SDP) [20] problem and possesses superior robustness against arbitrary array model errors, where the weight vector is determined from a high-dimensional matrix with the help of spectral factorization. In [21]-[25], quadratic or linear constraints are applied to formulate convex optimization problems [22], which can then be solved efficiently. In [26], two quadratic constraints are constructed to ensure that the magnitude responses at the arrival angles between two steering vectors exceed unity, and a closed-form solution is derived for the corresponding problem while the diagonal loading factor can be computed systematically. In [27], the DOA is modeled as a discrete random variable with known *a priori* probability density function, and a robust adaptive beamformer against DOA uncertainty is

derived using the Bayesian approach. In [28], the difference between the actual and presumed steering vectors is estimated to correct the erroneous presumed steering vector and thus the resultant method does not need any assumption on the norm of the mismatch vector or its probability. In [29], robust beamforming is achieved for both rank-one (point source) and high-rank (scattered source/fluctuating wavefront) signal models via explicit modeling of uncertainties in the desired signal array response and data covariance matrix. Furthermore, the minimum dispersion distortionless response beamformers are formulated in [30]-[32] for the non-Gaussian signals, including the l_p norm, gradient projection, and phase-only methods. The resultant non-convex quadratically constrained quadratic programming problems from the robust adaptive beamformers are tackled with the use of the quadratic matrix inequality method [33]-[34]. For online implementation, in [35]-[37], robust beamforming is realized using the recursive least squares filter or constrained Kalman filter. Recently, the deep learning (DL) techniques have been applied for beamforming [2], [38], huge numbers of sample covariance matrix-beamformer weight pairs are involved as the training data for a complex-valued convolutional neural network (CVCNN), respectively.

In radar and communication systems, there exist a large class of signals, exhibiting the constant modulus (CM) property. For example, in radar systems, besides the linear frequency modulation (LFM) signals, unimodularity or constant envelope has been a standard requirement in transmit phased-code sequence design [39]-[42] to maximize the achievable sensitivity and power efficiency through operating the power amplifiers in saturation. The BPSK $\{\pm 1\}$ or QPSK $\{\pm 1, \pm j\}$ modulated signals are widely used in communication systems. Therefore, the CM property may be utilized to solve the CM beamforming problem [43]-[45]. Although the beamformer output follows the Gaussian distribution under the Gaussian noise assumption, its magnitude (from absolute value operator) no longer follows the Gaussian distribution. As a result, the Gaussian property cannot be utilized by the CM-based beamforming methods [43]-[45]. To estimate the SOI waveform using an antenna array with uncertain steering vector, a Bayesian approach is introduced in [46] to successively derive an approximate minimum mean square error estimator and maximum *a posteriori* estimator, whereas an expected least-squares (LS) strategy is suggested in [47], resulting in a simple linear beamformer, which often performs similarly to the maximum likelihood estimator in terms of mean square error (MSE) and outperforms conventional signal-to-interference-plus-noise-ratio (SINR)-based approaches. The latter is also extended to minimize the MSE between the estimated and unknown signal waveforms via competitive beamforming approaches involving a robust MSE measure [48], [49].

Recently, we address the robust adaptive beamforming problem using the alternating direction method of multipliers (ADMM) [50], [51], where the two-sided mainlobe constraints or double constraints [52], [53] are adopted to improve the robustness of the Capon beamformer. Different from [52], [53], this paper focuses on the problem of constant modulus waveform estimation and interference suppression via two-stage fractional programming methods. The main contributions of this paper are summarized as follows:

- We devise an optimization criterion based on the Interference-plus-Noise-to-Noise Ratio (INNR) at the beamformer output to achieve perfect nulling at the IDs and minimum output noise power, which is analogous to LCMV beamformer, yet under the unknown ID case.

- An efficient two-stage approach is developed. In the first stage, we formulate an optimization model with a fractional objective function to attain the INNR bound; whereas in the second stage, another fraction-constrained optimization problem is established to refine the weight vector from the solution space constrained by the INNR bound, attaining deeper notches at the IDs and smaller output noise power than the first stage.

- Our framework is further extended to handle steering vector uncertainties, where the introduced bound variables for the numerator and denominator are utilized to breakdown the complicated problem into easily tackled subproblems.

The rest of this paper is organized as follows. The waveform estimation and interference suppression problem under known steering vector is formulated in Section II. The corresponding solution is derived in Section III, whereas its robustification against steering vector errors is presented in Section IV. In Section V, simulation results are provided to evaluate the performance of the developed algorithms. Finally, conclusions are drawn in Section VI.

Notation: Vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. $\| \cdot \|$ denotes the Frobenius norm, while $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ and $(\cdot)^{-1}$ are the transpose, conjugate, conjugate transpose and matrix inverse operators, respectively. $\mathbf{0}_{m \times n}$ and $\mathbf{I}_n$ represent the $m \times n$ zero matrix and $n \times n$ identity matrix, respectively. $\Re\{\}$ and $\Im\{\}$ return the real and imaginary parts with $j = \sqrt{-1}$. $|\cdot|$, and $\angle\{\cdot\}$ are the magnitude and phase of a complex-valued scalar, respectively.

II. PROBLEM FORMULATION

The received N-snapshot matrix $\mathbf{X} \in \mathbb{C}^{M \times N}$ by an antenna array with M sensors can be expressed as:

$$\mathbf{X} = \mathbf{a}(\theta_0)\alpha_0\mathbf{s}_0^T + \sum_{i=1}^{I} \mathbf{a}(\theta_i)\mathbf{s}_i^T + \mathbf{N}, \qquad (1)$$

where $\alpha_0\mathbf{s}_0 \in \mathbb{C}^{N \times 1}$ is a constant modulus SOI with steering vector $\mathbf{a}(\theta_0) \in \mathbb{C}^{M \times 1}$ at the direction θ_0, complex amplitude α_0, and unimodular waveform vector $\mathbf{s}_0 = [s_0(1), s_0(2), ..., s_0(N)]^T$ (i.e., $|s_0(n)| = 1, n = 1, \cdots, N$); and the ith interference signal $\mathbf{s}_i = [s_i(1), s_i(2), ..., s_i(N)]^T$ from the direction θ_i can be either constant modulus or not, for $i = 1, \cdots, I$. In addition, $\mathbf{N} \in \mathbb{C}^{M \times N}$ is the additive white Gaussian noise matrix with independent and identically distributed entries following $\mathcal{CN}(0, \sigma^2)$.

To extract the waveform $\mathbf{s}_0$, we may use the well-known Capon method (or MVDR beamformer) [11] with the constraint $\mathbf{w}^H \mathbf{a}(\theta_0) = 1$, which designs a weight vector $\mathbf{w} \in$

$\mathbb{C}^{M \times 1}$ to attain the minimum output power, i.e.,

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{R} \mathbf{w}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1, \qquad (2)$$

where $\mathbf{R} = E\{\mathbf{X}\mathbf{X}^H\}$ is the covariance matrix of $\mathbf{X}$, and is replaced by the sample covariance matrix $\frac{1}{N}\mathbf{X}\mathbf{X}^H$ in practice.

The power of the Capon beamformer output $\mathbf{w}^H \mathbf{X}$ with the constraint $\mathbf{w}^H \mathbf{a}(\theta_0) = 1$ can be written as:

$$\mathbf{w}^H \mathbf{R} \mathbf{w} = |\alpha_0|^2 E\{|s_0(n)|^2\}$$
$$+ \sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\} + \sigma^2 \mathbf{w}^H \mathbf{w}, \qquad (3)$$

where we can see that the first term $|\alpha_0|^2 E\{|s_0(n)|^2\}$ is unrelated to $\mathbf{w}$. However, the second and third components depend on $\mathbf{w}$. As a result, the minimum output power metric-based Capon beamformer actually tries to reach a compromise between $\sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}^T$ (interference suppression) and $\sigma^2 \mathbf{w}^H \mathbf{w}$ (output noise power reduction), and thus it may not effectively suppress the interferences.

If we know the IDs $\{\theta_i\}_{i=1}^I$, the constraints $\{\mathbf{w}^H \mathbf{a}(\theta_i) = 0\}_{i=1}^I$ can be introduced into the Capon beamformer where the perfect nulls are placed at the IDs, corresponding to the LCMV beamformer [4], [6]. That is, both minimum output noise power and maximum SNR subject to the constraints of perfect nulls at the IDs and unit gain at the direction of SOI are attained via the following formulation:

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{R} \mathbf{w}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1,$$
$$\mathbf{w}^H \mathbf{a}(\theta_i) = 0, i = 1, \cdots, I. \qquad (4)$$

However, in practice, the IDs may be unknown but only the direction of SOI is available. To address this challenge, this paper first utilizes the INNR in the beamformer output:

$$\frac{\sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\} + \sigma^2 \mathbf{w}^H \mathbf{w}}{\sigma^2 \mathbf{w}^H \mathbf{w}}$$
$$= \frac{\sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}}{\sigma^2 \mathbf{w}^H \mathbf{w}} + 1. \qquad (5)$$

Minimizing (5) is equivalent to minimizing the interference power to the noise power, i.e.,

$$\frac{\sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}}{\sigma^2 \mathbf{w}^H \mathbf{w}}, \qquad (6)$$

where the minimum is attained at $\mathbf{w}^H \mathbf{a}(\theta_i) = 0$ for $i = 1, \cdots, I$. In other words, the INNR minimization criterion implies $\mathbf{w}^H \mathbf{a}(\theta_i) = 0$. Through an intuition, we adopt INNR to approximate perfect interference nulling of the LCMV beamformer even when the IDs are unknown.

To realize this criterion, the expected value is approximated based on finite array snapshots. On the one hand, under Gaussian noise assumption, we may still use $\sigma^2 \mathbf{w}^H \mathbf{w}$ to represent the denominator. On the other hand, the numerator of (5) is replaced by the empirical power of the difference of the beamformer output and the constant modulus SOI:

$$\frac{1}{N}\|\sum_i \mathbf{w}^H \mathbf{a}(\theta_i)\mathbf{s}_i^T + \mathbf{w}^H \mathbf{N}\|^2 = \frac{1}{N}\|\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T\|^2, \qquad (7)$$

where the beamformer output

$$\mathbf{w}^H \mathbf{X} = \alpha_0 \mathbf{s}_0^T + \sum_i \mathbf{w}^H \mathbf{a}(\theta_i)\mathbf{s}_i^T + \mathbf{w}^H \mathbf{N}, \qquad (8)$$

and the unit-gain constraint $\mathbf{w}^H \mathbf{a}(\theta_0) = 1$ are applied.

Thus, the practical realization of (5) becomes:

$$\frac{(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)^H}{\sigma^2 \mathbf{w}^H \mathbf{w}}, \qquad (9)$$

which is equivalent to

$$\frac{(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)^H}{\mathbf{w}^H \mathbf{w}}. \qquad (10)$$

The expected value of (10) is given by:

$$E\left\{\frac{(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)^H}{\mathbf{w}^H \mathbf{w}}\right\}$$
$$= \sigma^2 + \frac{\sum_i |\mathbf{w}^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}}{\mathbf{w}^H \mathbf{w}}, \qquad (11)$$

where the minimum is still achieved at $\mathbf{w}^H \mathbf{a}(\theta_i) = 0$ for $i = 1, \cdots, I$. Hence we also refer (10) to as the INNR criterion.

Based on the aforementioned discussion, we formulate the following optimization problem via INNR minimization to estimate the constant modulus SOI and suppress the interferences:

$$\min_{\mathbf{w}, \alpha_0, \mathbf{s}_0} \quad \frac{(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)(\mathbf{w}^H \mathbf{X} - \alpha_0 \mathbf{s}_0^T)^H}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1$$
$$|s_0(n)| = 1, n = 1, \cdots, N. \qquad (12)$$

From (12), it can be seen that α_0 depends on $\{\mathbf{w}, \mathbf{s}_0\}$. When $\{\mathbf{w}, \mathbf{s}_0\}$ is known, α_0 is given by:

$$\alpha_0 = \frac{\mathbf{w}^H \mathbf{X} \mathbf{s}_0^*}{N}, \qquad (13)$$

which implies that $\frac{\mathbf{w}^H \mathbf{X} \mathbf{s}_0^* \mathbf{s}_0^T \mathbf{X}^H \mathbf{w}}{N^2}$ is the power estimate at the direction θ_0. Substituting (13) into (12) yields an optimization problem with only the variables $\{\mathbf{w}, \mathbf{s}_0\}$ (see Appendix A for the detailed derivation of the objective function):

$$\min_{\mathbf{w}, \mathbf{s}_0} \quad \frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^* \mathbf{s}_0^T)\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1$$
$$|s_0(n)| = 1, n = 1, \cdots, N. \qquad (14)$$

However, if any vector $\mathbf{w}$ satisfies

$$\mathbf{w}^H \mathbf{a}(\theta_0) = 1$$
$$\mathbf{w}^H \mathbf{a}(\theta_i) = 0, \ i = 1, \cdots, I, \qquad (15)$$

from the mathematical expectation view, it will be the solution of (14) and can recover the waveform $\mathbf{s}_0$ and attain the objective function bound (i.e., INNR bound). Since $I+1 < M$, i.e., the number of linear constraints on $\mathbf{w}$ is less than the dimensionality of $\mathbf{w}$, there exists many solutions to (14),

all of which can recover the SOI accurately, and suppress the interferences effectively. Therefore, similar to the LCMV beamformer, which minimizes the output noise power under the constraints of perfect nulls on the IDs and unit-gain on the SOI, we consider the following fraction-constrained optimization problem (**i.e., the second stage**) to minimize the beamformer output noise power ($\sigma^2 \mathbf{w}^H \mathbf{w}$, where the unknown constant σ^2 is removed) subject to the inherited INNR bound from (14) (**i.e., the first stage**):

$$\min_{\mathbf{w}, \mathbf{s}_0} \quad \mathbf{w}^H \mathbf{w}$$
$$\text{s.t.} \quad \frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^* \mathbf{s}_0^T)\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \leq \hat{\sigma}^2$$
$$\mathbf{w}^H \mathbf{a}(\theta_0) = 1$$
$$|s_0(n)| = 1, n = 1, \cdots, N, \qquad (16)$$

where $\hat{\sigma}^2$, i.e., the objective function value of (14), is provided as the INNR bound of $\frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^* \mathbf{s}_0^T)\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}}$ in the second stage. Thus, with the INNR bound constraint it is ensured that the second stage essentially inherits the minimal INNR from the first stage. According to the mathematical expectation, (16) may further refine the weight vector from the solution space constrained by the INNR bound $\hat{\sigma}^2$. Therefore, the synergy of both stages may achieve perfect interference nulling and minimum output noise power as the LCMV beamformer yet under the unknown ID case.

Now we discuss another physical meaning of (16). Let $\mathbf{w}_1$ and $\mathbf{w}_2$ denote the obtained weight vectors by (14) and (16), respectively. According to the objective function of (14) and the constraint in (16), we have $\frac{\sum_i |\mathbf{w}_2^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}}{\mathbf{w}_2^H \mathbf{w}_2} \leq \frac{\sum_i |\mathbf{w}_1^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}}{\mathbf{w}_1^H \mathbf{w}_1}$. Moreover, according to the objective function in (16) and because there is no energy constraint on $\mathbf{w}$ (i.e., $\mathbf{w}^H \mathbf{w}$) in (14), we have $\mathbf{w}_2^H \mathbf{w}_2 \leq \mathbf{w}_1^H \mathbf{w}_1$, and thus $\sum_i |\mathbf{w}_2^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\} \leq \sum_i |\mathbf{w}_1^H \mathbf{a}(\theta_i)|^2 E\{|s_i(n)|^2\}$. In other words, (16) will generate deeper notch levels, and higher output SINR, signal-to-interference-ratio (SIR) and SNR than (14).

Consequently, our task is to determine the waveform $\mathbf{s}_0$, and weight vector $\mathbf{w}$ with a two-stage procedure, i.e., implement (14) to compute $\hat{\sigma}^2$ and then solve (16) to find $\mathbf{w}$ and $\mathbf{s}_0$.

III. SOLUTIONS TO (14) AND (16)

Both (14) and (16) form the complete Waveform Estimation and Interference Suppression (WEIS) algorithm with known accurate steering vector. In this section, we first derive the solution to (14), and then (16) for the WEIS algorithm.

A. Stage 1: Solution to (14)

For (14), variables $\mathbf{s}_0$ and $\mathbf{w}$ can be updated via an alternating optimization (AO) procedure, i.e., update $\mathbf{s}_0(k)$ with gotten $\mathbf{w}(k-1)$ from the latest $(k-1)$th iteration, and then update $\mathbf{w}(k)$ with the newly obtained $\mathbf{s}_0(k)$ in the kth iteration, i.e.,

$$\mathbf{s}_0(k) = \arg \min_{\mathbf{s}_0} \quad \frac{\mathbf{w}^H(k-1)\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^* \mathbf{s}_0^T)\mathbf{X}^H \mathbf{w}(k-1)}{\mathbf{w}^H(k-1)\mathbf{w}(k-1)}$$

$$= \arg \min_{\mathbf{s}_0} \quad \frac{\mathbf{w}^H(k-1)\mathbf{X}\mathbf{X}^H \mathbf{w}(k-1)}{\mathbf{w}^H(k-1)\mathbf{w}(k-1)}$$
$$+ \frac{-\frac{1}{N}\mathbf{w}^H(k-1)\mathbf{X}\mathbf{s}_0^* \mathbf{s}_0^T \mathbf{X}^H \mathbf{w}(k-1))}{\mathbf{w}^H(k-1)\mathbf{w}(k-1)}$$

$$= \arg \min_{\mathbf{s}_0} \quad \frac{-\frac{1}{N}\|\mathbf{w}^H(k-1)\mathbf{X}\mathbf{s}_0^*\|^2}{\mathbf{w}^H(k-1)\mathbf{w}(k-1)}$$

$$= \arg \max_{\mathbf{s}_0} \quad \|\mathbf{w}^H(k-1)\mathbf{X}\mathbf{s}_0^*\|^2$$
$$\text{s.t.} \quad |s_0(n)| = 1, n = 1, \cdots, N$$
$$= e^{j\angle(\mathbf{w}^H(k-1)\mathbf{X})}, \qquad (17)$$

$$\mathbf{w}(k) = \arg \min_{\mathbf{w}} \quad \frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*(k)\mathbf{s}_0^T(k))\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}}$$

$$= \arg \min_{\mathbf{w}} \quad \frac{\mathbf{w}^H \mathbf{X}\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}} - \frac{\frac{1}{N}\|\mathbf{w}^H \mathbf{X}\mathbf{s}_0^*(k)\|^2}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1. \qquad (18)$$

The AO procedure is repeated until a given Stopping Criterion (SC) (see Exp. 1 for detailed discussion) is reached.

According to the equivalent *max* operation in the objective function of (17), it is essentially to pursuit the maximum correlation between $\mathbf{w}^H \mathbf{X}$ and $\mathbf{s}_0^*$. Therefore, in (18) the first component essentially minus a maximum (second) component. Therefore, the iterations with (17) and (18) enable the objective function in (14) to be nonincremental with the iterations, and thus the AO procedure is convergent [58] (or see Appendix E for details).

As (17) has a closed-form solution for $\mathbf{s}_0$, we focus on deriving the solution to (18) here. First we simplify (18) as:

$$\min_{\mathbf{w}} \quad \frac{\mathbf{w}^H \mathbf{R}_1(k)\mathbf{w}}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1, \qquad (19)$$

where the matrix:

$$\mathbf{R}_1(k) = \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*(k)\mathbf{s}_0^T(k))\mathbf{X}^H \in \mathbb{C}^{M \times M}, \qquad (20)$$

is introduced for presentation simplicity purpose.

Introducing a scalar variable η to ensure that $(\eta\mathbf{w})^H(\eta\mathbf{w}) = 1$ and thus eliminate the fraction, we rewrite (19) as:

$$\min_{\mathbf{w}, \eta} \quad (\eta\mathbf{w})^H \mathbf{R}_1(k)(\eta\mathbf{w})$$
$$\text{s.t.} \quad (\eta\mathbf{w})^H(\eta\mathbf{w}) = 1$$
$$(\eta\mathbf{w})^H \mathbf{a}(\theta_0) = \eta. \qquad (21)$$

To facilitate the solution derivation, we define two augmented vectors:

$$\bar{\mathbf{w}} = [(\eta\mathbf{w})^T \quad \eta]^T \in \mathbb{C}^{(M+1) \times 1}, \qquad (22)$$
$$\bar{\mathbf{a}}(\theta_0) = [\mathbf{a}^T(\theta_0) \quad -1]^T \in \mathbb{C}^{(M+1) \times 1}, \qquad (23)$$

to represent (21) with one vector variable $\bar{\mathbf{w}}$:

$$\min_{\bar{\mathbf{w}}} \quad \bar{\mathbf{w}}^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}$$
$$\text{s.t.} \quad \bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1$$
$$\bar{\mathbf{w}}^H \bar{\mathbf{a}}(\theta_0) = 0, \tag{24}$$

where the augmented matrices are:

$$\bar{\mathbf{R}}_1(k) = \begin{bmatrix} \mathbf{R}_1(k) & \mathbf{0}_{M \times 1} \\ \mathbf{0}_{1 \times M} & 0 \end{bmatrix} \in \mathbb{C}^{(M+1) \times (M+1)}, \tag{25}$$

$$\bar{\mathbf{I}}_{M+1} = \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times 1} \\ \mathbf{0}_{1 \times M} & 0 \end{bmatrix} \in \mathbb{C}^{(M+1) \times (M+1)}. \tag{26}$$

Note that the challenging fraction is removed at the cost of introduction of the optimization variable η. Although the dimensionality increase of both the augmented optimization variable $\bar{\mathbf{w}}$ and matrices $\{\bar{\mathbf{R}}_1(k), \bar{\mathbf{I}}_{M+1}\}$ enlarges the scale of the optimization problem, (24) is equivalent to (19). Therefore, we may solve the simpler (24) to find the solution to (19).

To tackle the nonconvex constraint $\bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1$, we consider introducing a consensus variable $\bar{\mathbf{v}} \in \mathbb{C}^{(M+1) \times 1}$:

$$\bar{\mathbf{w}} = \bar{\mathbf{v}}, \tag{27}$$

to separate the two constraints into two groups, where only the constraint $\bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1$ associates with the objective function so that the resultant subproblem may be readily tackled, i.e.,

$$\min_{\bar{\mathbf{w}}, \bar{\mathbf{v}}} \quad \bar{\mathbf{w}}^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}$$
$$\text{s.t.} \quad \bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1$$
$$\bar{\mathbf{w}} = \bar{\mathbf{v}}$$
$$\bar{\mathbf{v}}^H \bar{\mathbf{a}}(\theta_0) = 0. \tag{28}$$

With (28), we construct an augmented Lagrangian as:

$$\mathcal{L}_1(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \boldsymbol{\lambda}) = \bar{\mathbf{w}}^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}} + \Re(\boldsymbol{\lambda}^H(\bar{\mathbf{w}} - \bar{\mathbf{v}})) + \frac{\rho}{2} \|\bar{\mathbf{w}} - \bar{\mathbf{v}}\|^2$$
$$\text{s.t.} \quad \bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1$$
$$\bar{\mathbf{v}}^H \bar{\mathbf{a}}(\theta_0) = 0, \tag{29}$$

where $\boldsymbol{\lambda} \in \mathbb{C}^{(M+1) \times 1}$ is the Lagrange multiplier vector corresponding to the consensus constraint in (27), and $\rho > 0$, called (user-defined) step size or augmented Lagrangian parameter in ADMM [50], is the penalty parameter on the consensus error (CE). Then, we apply the ADMM [50], [51] to solve (29) via the following iterative steps.

Step A1: Update $\bar{\mathbf{v}}(t+1)$ with given $\{\bar{\mathbf{w}}(t), \boldsymbol{\lambda}(t)\}$:

$$\bar{\mathbf{v}}(t+1) = \arg\min_{\bar{\mathbf{v}}} \quad \mathcal{L}_1(\bar{\mathbf{w}}(t), \bar{\mathbf{v}}, \boldsymbol{\lambda}(t))$$
$$\text{s.t.} \quad \bar{\mathbf{v}}^H \bar{\mathbf{a}}(\theta_0) = 0, \tag{30}$$

where t denotes the iteration index.

Completing the square and ignoring the constant terms, we simplify (30) into:

$$\min_{\bar{\mathbf{v}}} \quad \|\bar{\mathbf{v}} - \tilde{\mathbf{v}}(t+1)\|^2$$
$$\text{s.t.} \quad \bar{\mathbf{v}}^H \bar{\mathbf{a}}(\theta_0) = 0, \tag{31}$$

where

$$\tilde{\mathbf{v}}(t+1) = \bar{\mathbf{w}}(t) + \frac{\boldsymbol{\lambda}(t)}{\rho} \in \mathbb{C}^{(M+1) \times 1}, \tag{32}$$

is introduced for presentation simplicity. Then, the solution to (31) is given by:

$$\bar{\mathbf{v}}(t+1) = -\frac{\bar{\mathbf{a}}^H(\theta_0) \tilde{\mathbf{v}}(t+1)}{\bar{\mathbf{a}}^H(\theta_0) \bar{\mathbf{a}}(\theta_0)} \bar{\mathbf{a}}(\theta_0) + \tilde{\mathbf{v}}(t+1). \tag{33}$$

Step A2: Update $\bar{\mathbf{w}}(t+1)$ with given $\{\bar{\mathbf{v}}(t+1), \boldsymbol{\lambda}(t)\}$ via solving (see Appendix B for details):

$$\bar{\mathbf{w}}(t+1) = \arg\min_{\bar{\mathbf{w}}} \quad \mathcal{L}_1(\bar{\mathbf{w}}, \bar{\mathbf{v}}(t+1), \boldsymbol{\lambda}(t)) \tag{34}$$
$$\text{s.t.} \quad \bar{\mathbf{w}}^H \bar{\mathbf{I}}_{M+1} \bar{\mathbf{w}} = 1.$$

Step A3: Update $\boldsymbol{\lambda}(t+1)$ with given $\{\bar{\mathbf{w}}(t+1), \bar{\mathbf{v}}(t+1)\}$:

$$\boldsymbol{\lambda}(t+1) = \boldsymbol{\lambda}(t) + \rho(\bar{\mathbf{w}}(t+1) - \bar{\mathbf{v}}(t+1)). \tag{35}$$

Steps A1) to A3) are repeated until a given SC is reached. Once $\bar{\mathbf{w}}(t+1)$ is computed from the ADMM procedure, we obtain $\mathbf{w}(k)$ via throwing out the last element of $\bar{\mathbf{w}}(t+1)$, i.e.,

$$\mathbf{w}(k) = \frac{1}{\bar{w}_{M+1}} [\bar{w}_1, \bar{w}_2, ..., \bar{w}_M]^T, \tag{36}$$

where $\bar{w}_m$ is the mth element of $\bar{\mathbf{w}}(t+1)$ for $m = 1, \cdots, M+1$. Completing AO procedure, we insert $\mathbf{w}(k)$ and $\mathbf{s}_0(k)$ into the objective function of (14) to yield the INNR bound $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{\mathbf{w}^H(k) \mathbf{X}(\mathbf{I}_M - \frac{1}{N} \mathbf{s}_0^*(k) \mathbf{s}_0^T(k)) \mathbf{X}^H \mathbf{w}(k)}{\mathbf{w}^H(k) \mathbf{w}(k)}, \tag{37}$$

as the input of Stage 2.

We summarize the aforementioned procedure for INNR bound computation in **Stage 1 of WEIS** (WEIS-S1) in Fig. 1.

B. Stage 2: Solution to (16)

As in Stage 1, we solve (16) with an AO procedure: update $\mathbf{s}_0(k)$ using (17), and then update $\mathbf{w}(k)$ via solving the following subproblem:

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{w}$$
$$\text{s.t.} \quad \frac{\mathbf{w}^H \mathbf{R}_1(k) \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \leq \hat{\sigma}^2$$
$$\mathbf{w}^H \mathbf{a}(\theta_0) = 1, \tag{38}$$

where $\mathbf{R}_1(k)$ is defined in (20). The AO procedure is repeated until a given SC is reached. Similarly, here we focus on the solution derivation of (38).

Since the matrix $\mathbf{R}_1(k)$ is semidefinite positive (see Appendix A for details), it can be decomposed as $\mathbf{R}_1(k) = \mathbf{U}_k \mathbf{U}_k^H$. Then, introducing two auxiliary vectors:

$$\mathbf{g} = \mathbf{w}^H \mathbf{U}_k \in \mathbb{C}^{M \times 1}, \tag{39}$$
$$\mathbf{h} = \mathbf{w} \in \mathbb{C}^{M \times 1}, \tag{40}$$

we transform (38) into the following equivalent problem:

$$\min_{\mathbf{w}, \mathbf{g}, \mathbf{h}} \quad \mathbf{w}^H \mathbf{w}$$
$$\text{s.t.} \quad \frac{\|\mathbf{g}\|^2}{\|\mathbf{h}\|^2} \leq \hat{\sigma}^2$$
$$\mathbf{g} = \mathbf{w}^H \mathbf{U}_k$$
$$\mathbf{h} = \mathbf{w}$$
$$\mathbf{w}^H \mathbf{a}(\theta_0) = 1. \tag{41}$$

With (41), we construct another augmented Lagrangian:

$$\begin{aligned}
&\mathcal{L}_2\left(\mathbf{w}, \mathbf{g}, \mathbf{h}, \boldsymbol{\lambda}_g, \boldsymbol{\lambda}_h\right) \\
&= \mathbf{w}^H \mathbf{w} + \Re(\boldsymbol{\lambda}_g^H\left(\mathbf{g} - \mathbf{w}^H \mathbf{U}_k\right)) + \frac{\rho}{2}\|\mathbf{g} - \mathbf{w}^H \mathbf{U}_k\|^2 \\
&\quad + \Re(\boldsymbol{\lambda}_h^H\left(\mathbf{h} - \mathbf{w}\right)) + \frac{\rho}{2}\|\mathbf{h} - \mathbf{w}\|^2
\end{aligned}$$

$$\text{s.t.} \quad \frac{\|\mathbf{g}\|^2}{\|\mathbf{h}\|^2} \le \hat{\sigma}^2$$
$$\mathbf{w}^H \mathbf{a}(\theta_0) = 1, \tag{42}$$

where $\boldsymbol{\lambda}_g \in \mathbb{C}^{M \times 1}$ and $\boldsymbol{\lambda}_h \in \mathbb{C}^{M \times 1}$ are the Lagrange multiplier vectors corresponding to the equality constraints $\mathbf{g} = \mathbf{w}^H \mathbf{U}_k$ and $\mathbf{h} = \mathbf{w}$, respectively, and $\rho > 0$ is a user-defined penalty parameter on the CE.

Then, based on the ADMM [50], [51], we solve (42) via the following iterative steps.

Step B1: Update $\mathbf{w}(t + 1)$ with given $\{\mathbf{g}(t), \mathbf{h}(t), \boldsymbol{\lambda}_g(t), \boldsymbol{\lambda}_h(t)\}$ via solving:

$$\min_{\mathbf{w}} \quad \mathcal{L}_2\left(\mathbf{w}, \mathbf{g}(t), \mathbf{h}(t), \boldsymbol{\lambda}_g(t), \boldsymbol{\lambda}_h(t)\right)$$
$$\text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1, \tag{43}$$

the solution to which is given by:

$$\mathbf{w}(t+1) = ((\frac{2}{\rho} + 1)\mathbf{I}_M + \mathbf{U}_k \mathbf{U}_k^H)^{-1}(\mathbf{r} + \frac{\gamma^*}{2}\mathbf{a}(\theta_0)), \tag{44}$$

where

$$\mathbf{r} = \mathbf{U}_k(\mathbf{g}(t) + \frac{\boldsymbol{\lambda}_g(t)}{\rho})^H + \mathbf{h}(t) + \frac{\boldsymbol{\lambda}_h(t)}{\rho} \in \mathbb{C}^{M \times 1}, \tag{45}$$

$$\gamma = \frac{2 - 2\mathbf{r}^H((\frac{2}{\rho} + 1)\mathbf{I}_M + \mathbf{U}_k \mathbf{U}_k^H)^{-1}\mathbf{a}(\theta_0)}{\mathbf{a}^H(\theta_0)((\frac{2}{\rho} + 1)\mathbf{I}_M + \mathbf{U}_k \mathbf{U}_k^H)^{-1}\mathbf{a}(\theta_0)} \in \mathbb{C}. \tag{46}$$

Step B2: Update $\mathbf{g}(t + 1)$ and $\mathbf{h}(t + 1)$ with given $\{\mathbf{w}(t + 1), \boldsymbol{\lambda}_g(t), \boldsymbol{\lambda}_h(t)\}$ via solving the following subproblem (see Appendix C for details):

$$\min_{\mathbf{g}, \mathbf{h}} \quad \mathcal{L}_2\left(\mathbf{w}(t+1), \mathbf{g}, \mathbf{h}, \boldsymbol{\lambda}_g(t), \boldsymbol{\lambda}_h(t)\right),$$
$$\text{s.t.} \quad \frac{\|\mathbf{g}\|^2}{\|\mathbf{h}\|^2} \le \hat{\sigma}^2. \tag{47}$$

Step B3: Update $\boldsymbol{\lambda}_g(t + 1)$ and $\boldsymbol{\lambda}_h(t + 1)$ with given $\{\mathbf{w}(t + 1), \mathbf{g}(t + 1), \mathbf{h}(t + 1)\}$:

$$\boldsymbol{\lambda}_g(t + 1) = \boldsymbol{\lambda}_g(t) + \rho\left(\mathbf{g}(t + 1) - \mathbf{w}^H(t + 1)\mathbf{U}_k\right) \tag{48}$$
$$\boldsymbol{\lambda}_h(t + 1) = \boldsymbol{\lambda}_h(t) + \rho\left(\mathbf{h}(t + 1) - \mathbf{w}(t + 1)\right). \tag{49}$$

Steps B1) to B3) are repeated until a given SC is reached. Now, the complete derivation of the WEIS method is finished. Here we give a block diagram in Fig. 1 for the complete procedure to show the inputs and outputs of each block in all stages (see detailed SC1-SC4 discussion in Exp.1).

IV. EXTENSION TO INACCURATE STEERING VECTOR

In this section, we tackle the Waveform Estimation and Interference Suppression problem (under inaccurate steering vector case) via Robust beamforming (named as WEIS-R algorithm). Similar to [15], we assume that the uncertainty of the steering vector $\mathbf{a}(\theta_0)$ is contained in an ellipsoid, i.e., $\{\mathbf{a}(\theta_0) \in \mathbf{A}\mathbf{u} + \bar{\mathbf{a}}, \|\mathbf{u}\| \le 1\}$, where $\bar{\mathbf{a}}$ and $\mathbf{A}$ are the
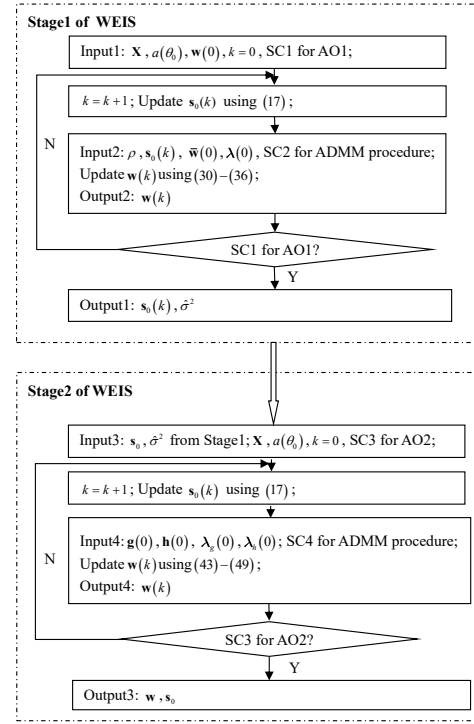

Fig. 1: Block diagram of WEIS algorithm.

ellipsoid center and configuration matrix, respectively. Thus, the constraint on $\mathbf{w}$ is given by [15]:

$$\|\mathbf{A}^H \mathbf{w}\| \le |\bar{\mathbf{a}}^H \mathbf{w}| - 1, \tag{50}$$

and the waveform estimation and interference suppression problem under the inaccurate steering vector case is formulated as:

$$\min_{\mathbf{w}, \mathbf{s}_0} \quad \frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^* \mathbf{s}_0^T)\mathbf{X}^H \mathbf{w}}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \|\mathbf{A}^H \mathbf{w}\| \le |\bar{\mathbf{a}}^H \mathbf{w}| - 1$$
$$|s_0(n)| = 1, n = 1, \cdots, N. \tag{51}$$

Similar to (14), $\mathbf{s}_0$ and $\mathbf{w}$ in (51) are optimized alternately. Since the subproblem on $\mathbf{s}_0$ can be solved in a closed form (see (17) for details), we focus on the subproblem on $\mathbf{w}$:

$$\min_{\mathbf{w}} \quad \frac{\mathbf{w}^H \mathbf{R}_1(k)\mathbf{w}}{\mathbf{w}^H \mathbf{w}}$$
$$\text{s.t.} \quad \|\mathbf{A}^H \mathbf{w}\| \le \Re(\bar{\mathbf{a}}^H \mathbf{w}) - 1$$
$$\Im(\bar{\mathbf{a}}^H \mathbf{w}) = 0, \tag{52}$$

where $\mathbf{R}_1(k)$ is defined in (20), and $\Re\{.\}$ and $\Im\{.\}$ operators are introduced to avoid the phase ambiguity of the solutions (see [15] for details).

Note that both the numerator and denominator in the objective function of (52) are functions of $\mathbf{w}$. To decouple them, the following equality constraints with auxiliary variables

$\{\mathbf{c} \in \mathbb{C}^{M \times 1}, \mathbf{d} \in \mathbb{C}^{M \times 1}, q \in \Re, p \in \Re\}$ are introduced:

$$\mathbf{c} = \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}, \tag{53}$$

$$\mathbf{d} = \mathbf{w}, \tag{54}$$

$$q = \|\mathbf{c}\|^2, \tag{55}$$

$$p = \|\mathbf{d}\|^2. \tag{56}$$

Employing $q = \|\mathbf{c}\|^2$ and $p = \|\mathbf{d}\|^2$, the objective function of the optimization problem becomes $\frac{q}{p}$ with the constraints $q = \|\mathbf{c}\|^2$ and $p = \|\mathbf{d}\|^2$. In this case, although the introduced constraints simplify the objective function, the resultant problem from them is still hard to solve. Moreover, we relax the two constraints as $\|\mathbf{c}\|^2 \leq q$ and $\|\mathbf{d}\|^2 \geq p$. It is easily found that the minimum of $\frac{q}{p}$ is reached at the lower and upper bounds of the numerator and denominator, respectively, i.e., $q = \|\mathbf{c}\|^2$ and $p = \|\mathbf{d}\|^2$. Therefore, the constraint relaxation still ensures the equivalence with the original problem. Finally, we introduce $q = \bar{q}$ and $p = \bar{p}$ to decouple the objective function $\frac{q}{p}$ and the constraints $\|\mathbf{c}\|^2 \leq q$ and $\|\mathbf{d}\|^2 \geq p$. In doing so, (52) is transformed to:

$$\begin{aligned} \min_{\mathbf{w}, p, q, \bar{p}, \bar{q}, \mathbf{c}, \mathbf{d}} \quad & \frac{\bar{q}}{\bar{p}} \\ \text{s.t.} \quad & \mathbf{c} = \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w} \\ & \mathbf{d} = \mathbf{w} \\ & \|\mathbf{c}\|^2 \leq q \\ & \|\mathbf{d}\|^2 \geq p \\ & q = \bar{q} \\ & p = \bar{p} \\ & \|\mathbf{A}^H \mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H \mathbf{w}) - 1 \\ & \Im(\bar{\mathbf{a}}^H \mathbf{w}) = 0. \end{aligned} \tag{57}$$

Then, its augmented Lagrangian is constructed as

$$\begin{aligned} \mathcal{L}_3(\mathbf{w}, p, q, \bar{p}, \bar{q}, \mathbf{c}, \mathbf{d}, \lambda, \kappa, \boldsymbol{\xi}, \boldsymbol{\zeta}) = & \frac{\bar{q}}{\bar{p}} \\ + \lambda(q - \bar{q}) + & \frac{\rho}{2}|q - \bar{q}|^2 + \kappa(p - \bar{p}) + \frac{\rho}{2}|p - \bar{p}|^2 \\ + \Re\{\boldsymbol{\xi}^H(\mathbf{c} - & \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w})\} + \frac{\rho}{2}\|\mathbf{c} - \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}\|^2 \\ + \Re\{\boldsymbol{\zeta}^H(\mathbf{d} - & \mathbf{w})\} + \frac{\rho}{2}\|\mathbf{d} - \mathbf{w}\|^2 \\ \text{s.t.} \quad & \|\mathbf{c}\|^2 \leq q \\ & \|\mathbf{d}\|^2 \geq p \\ & \|\mathbf{A}^H \mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H \mathbf{w}) - 1 \\ & \Im(\bar{\mathbf{a}}^H \mathbf{w}) = 0, \end{aligned} \tag{58}$$

where $\{\lambda \in \Re, \kappa \in \Re, \boldsymbol{\xi} \in \mathbb{C}^{M \times 1}, \boldsymbol{\zeta} \in \mathbb{C}^{M \times 1}\}$ are the Lagrange multipliers corresponding to the constraints $q = \bar{q}$, $p = \bar{p}$, $\mathbf{c} = \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}$, and $\mathbf{d} = \mathbf{w}$, respectively, and $\rho > 0$ is the user-defined penalty parameter.

Based on the ADMM [50], [51], we solve (58) via the following iterative steps.

Step C1: Update $\{\mathbf{w}(t+1), \bar{p}(t+1), \bar{q}(t+1)\}$ with given

$\{p(t), q(t), \mathbf{c}(t), \mathbf{d}(t), \lambda(t), \kappa(t), \boldsymbol{\xi}(t), \boldsymbol{\zeta}(t)\}$ by solving:

$$\begin{aligned} \min_{\mathbf{w}, \bar{p}, \bar{q}} & \mathcal{L}_3(\mathbf{w}, p(t), q(t), \bar{p}, \bar{q}, \mathbf{c}(t), \mathbf{d}(t), \lambda(t), \kappa(t), \boldsymbol{\xi}(t), \boldsymbol{\zeta}(t)) \\ \text{s.t.} \quad & \|\mathbf{A}^H \mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H \mathbf{w}) - 1 \\ & \Im(\bar{\mathbf{a}}^H \mathbf{w}) = 0, \end{aligned} \tag{59}$$

which can be divided into two subproblems. One only involves $\{\bar{p}, \bar{q}\}$:

$$\min_{\bar{p}, \bar{q}} \frac{\bar{q}}{\bar{p}} + \frac{\rho}{2}(\check{q}(t+1) - \bar{q})^2 + \frac{\rho}{2}(\check{p}(t+1) - \bar{p})^2, \tag{60}$$

and another is related to only $\mathbf{w}$:

$$\begin{aligned} \min_{\mathbf{w}} & \|\bar{\mathbf{c}}(t+1) - \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}\|^2 + \|\bar{\mathbf{d}}(t+1) - \mathbf{w}\|^2 \\ \text{s.t.} \quad & \|\mathbf{A}^H \mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H \mathbf{w}) - 1 \\ & \Im(\bar{\mathbf{a}}^H \mathbf{w}) = 0, \end{aligned} \tag{61}$$

where the squares are completed and the constant term is ignored, and

$$\check{q}(t+1) = q(t) + \frac{\lambda(t)}{\rho} \in \Re, \tag{62}$$

$$\check{p}(t+1) = p(t) + \frac{\kappa(t)}{\rho} \in \Re, \tag{63}$$

$$\bar{\mathbf{c}}(t+1) = \mathbf{c}(t) + \frac{\boldsymbol{\xi}(t)}{\rho} \in \mathbb{C}^{M \times 1}, \tag{64}$$

$$\bar{\mathbf{d}}(t+1) = \mathbf{d}(t) + \frac{\boldsymbol{\zeta}(t)}{\rho} \in \mathbb{C}^{M \times 1}. \tag{65}$$

Here (61) is solved via the CVX toolbox [54]; whereas we solve (60) via the following procedure. Partial differentiating (60) with respect to $\bar{q}$ and $\bar{p}$ yields

$$\frac{1}{\bar{p}} + \rho(\bar{q} - \check{q}(t+1)) = 0 \Longrightarrow \bar{q} = \check{q}(t+1) - \frac{1}{\rho\bar{p}}, \tag{66}$$

$$-\frac{\bar{q}}{\bar{p}^2} + \rho(\bar{p} - \check{p}(t+1)) = 0. \tag{67}$$

Inserting (66) into (67), we obtain:

$$\frac{\rho^2 \bar{p}^4 - \rho^2 \check{p}(t+1)\bar{p}^3 - \rho\check{q}(t+1)\bar{p} + 1}{\rho\bar{p}^3} = 0, \tag{68}$$

where its four roots are given by $\{\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4\}$ (via MATLAB function "roots"). Inserting them into $\bar{q} = \check{q}(t+1) - \frac{1}{\rho\bar{p}}$ yields $\{\bar{q}_1, \bar{q}_2, \bar{q}_3, \bar{q}_4\}$. From the four pairs $\{\bar{p}_i, \bar{q}_i\}_{i=1}^4$, we pick the real pair with the smallest objective function value of (60), denoted as $\{\bar{p}(t+1), \bar{q}(t+1)\}$.

Step C2: Update $\{p(t+1), q(t+1), \mathbf{c}(t+1), \mathbf{d}(t+1)\}$ with given $\{\mathbf{w}(t+1), \bar{p}(t+1), \bar{q}(t+1), \lambda(t), \kappa(t), \boldsymbol{\xi}(t), \boldsymbol{\zeta}(t)\}$ by solving (see Appendix D for detailed derivation)

$$\begin{aligned} \min_{p, q, \mathbf{c}, \mathbf{d}} & \mathcal{L}_3(\mathbf{w}(t+1), p, q, \bar{p}(t+1), \bar{q}(t+1), \mathbf{c}, \mathbf{d}, \\ & \lambda(t), \kappa(t), \boldsymbol{\xi}(t), \boldsymbol{\zeta}(t)) \\ \text{s.t.} \quad & \|\mathbf{c}\|^2 \leq q \\ & \|\mathbf{d}\|^2 \geq p. \end{aligned} \tag{69}$$

Step C3: Update $\{\lambda(t+1), \kappa(t+1), \boldsymbol{\xi}(t+1), \boldsymbol{\zeta}(t+1)\}$ with given $\{\mathbf{w}(t+1), p(t+1), q(t+1), \bar{p}(t+1), \bar{q}(t+1), \mathbf{c}(t+1), \mathbf{d}(t+1)\}$:

$$\lambda(t+1) = \lambda(t) + \rho(q(t+1) - \bar{q}(t+1)), \tag{70}$$

$$\kappa(t+1) = \kappa(t) + \rho(p(t+1) - \bar{p}(t+1)), \tag{71}$$

$$\boldsymbol{\xi}(t+1) = \boldsymbol{\xi}(t) + \rho(\mathbf{c}(t+1) - \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}(t+1)), \tag{72}$$

$$\boldsymbol{\zeta}(t+1) = \boldsymbol{\zeta}(t) + \rho(\mathbf{d}(t+1) - \mathbf{w}(t+1)). \tag{73}$$

Steps C1) to C3) are repeated until a given SC is reached. We summarize the aforementioned procedure in **Stage 1 of WEIS-R** in Fig. 2.

Similar to (14), (51) helps us to determine the INNR bound $\hat{\sigma}^2$ accurately, but it may not suppress the output noise power sufficiently. Therefore, we construct the following model to further reduce the output noise power with $\hat{\sigma}^2$:

$$
\begin{aligned}
\min_{\mathbf{w},\mathbf{s}_0} \quad & \mathbf{w}^H\mathbf{w} \\
\text{s.t.} \quad & \frac{\mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H\mathbf{w}}{\mathbf{w}^H\mathbf{w}} \leq \hat{\sigma}^2 \\
& \|\mathbf{A}^H\mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H\mathbf{w}) - 1 \\
& \Im(\bar{\mathbf{a}}^H\mathbf{w}) = 0 \\
& |s_0(n)| = 1, \quad n = 1,\cdots,N,
\end{aligned}
\tag{74}
$$

which may be solved via replacing the constraint $\mathbf{w}^H\mathbf{a}(\theta_0) = 1$ in (16) of the WEIS with $\|\mathbf{A}^H\mathbf{w}\| \leq \Re(\bar{\mathbf{a}}^H\mathbf{w}) - 1$ and $\Im(\bar{\mathbf{a}}^H\mathbf{w}) = 0$ (whereas other steps are the same as **Stage 2 of WEIS**). Thus, (51) and (74) form the complete WEIS-R method. Here we give a block diagram in Fig. 2 for the complete procedure to show the inputs and outputs of each block in all stages of the WEIS-R method (see detailed SC5-SC8 discussion in Exp. 3).

V. SIMULATION RESULTS

Computer simulations are conducted to assess the performance of the proposed methods. We consider a uniform linear antenna array with half-wavelength inter-element spacing. In addition, we set the transmit waveform $\mathbf{s}_0$ and the interference $\{\mathbf{s}_i\}$ as constant modular signals with random phases, and the additive noise is assumed to be a zero-mean circularly symmetric complex Gaussian random process with variance δ^2 at each antenna. To compare the proposed methods with their counterparts, we compute the corresponding normalized mean square error (NMSE) based on $L = 100$ Monte Carlo runs as:

$$\text{NMSE} = \frac{1}{L}\sum_{l=1}^{L}\frac{\|\hat{\mathbf{s}}_0(l) - \mathbf{s}_0(l)\|^2}{\|\mathbf{s}_0(l)\|^2}, \tag{75}$$

where $\mathbf{s}_0(l)$ and $\hat{\mathbf{s}}_0(l)$ are the true and estimated waveforms in the lth Monte Carlo run, respectively.

A. Performance of WEIS Algorithm

Experiment 1. Convergence Performance and SC Setting: In this subsection, we explore the convergence performance and SCs for the WEIS algorithm. Here we set the direction of SOI as $\theta_0 = 0°$, and the DOAs of two interferences to be $\pm 60°$. In
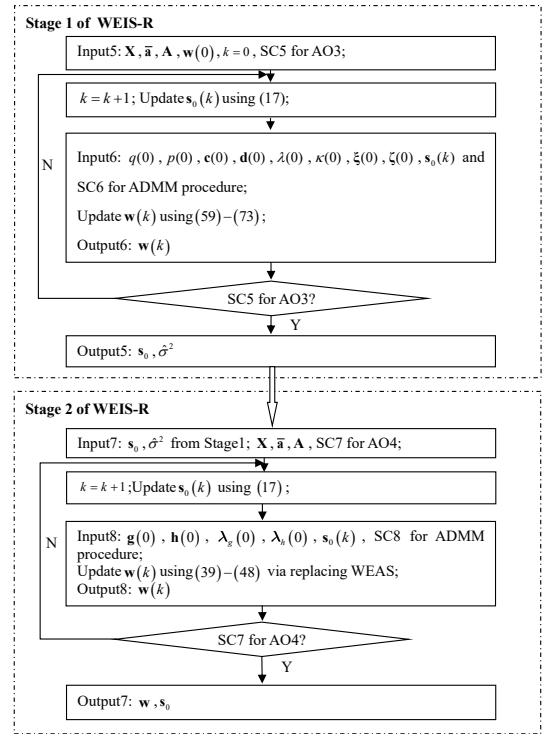

Fig. 2: Block diagram of WEIS-R algorithm

(a) OFVs versus IN (AO).

(b) DOFVs versus IN (AO).

(c) PEs of $\mathbf{s}_0(k)$ versus IN (AO).

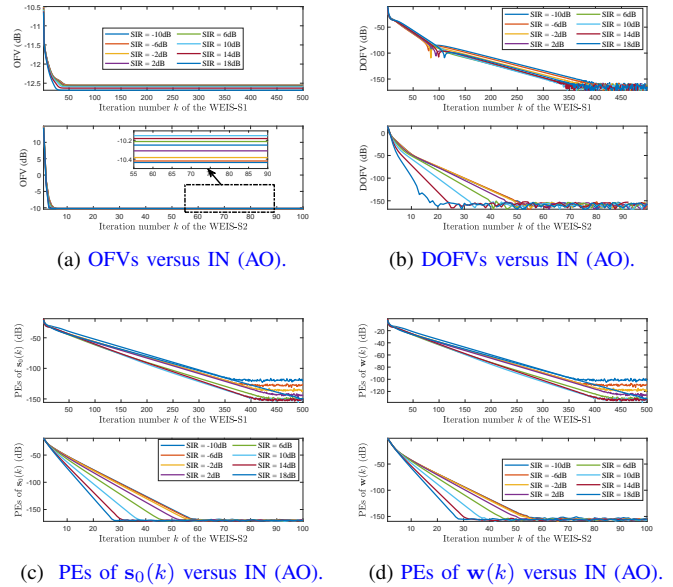(d) PEs of $\mathbf{w}(k)$ versus IN (AO).

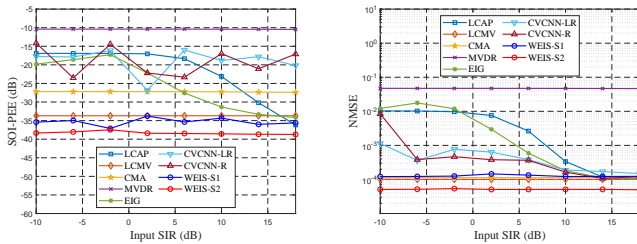Fig. 3: Convergence performance of WEIS-AO in Exp. 1.

addition, the sensor number is $M = 10$, and snapshot number is $N = 200$. Meanwhile, the input SNR is fixed at 15dB but the SIR varies from -10dB to 18dB. For WEIS-S1, the penalty parameter is set as $\rho = 0.05$, and $\mathbf{w}$ is initialized using the MVDR beamformer; whereas the output $\{\mathbf{s}_0, \hat{\sigma}^2\}$ of WEIS-S1 and $\rho = 0.1$ are used for WEIS-S2. Other parameters, such as auxiliary variables and Lagrange multipliers, are initialized randomly for the two stages.

Note that here we set sufficiently large numbers of iter-
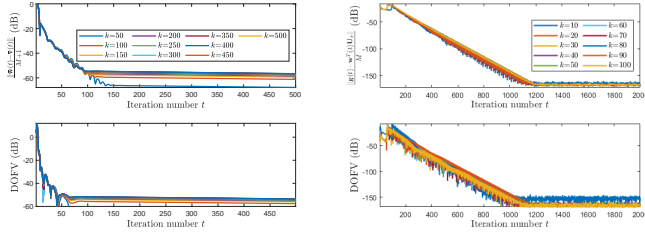
(a) CEs and DOFVs versus IN (WEIS-S1-ADMM).

(b) CEs and DOFVs versus IN (WEIS-S2-ADMM).
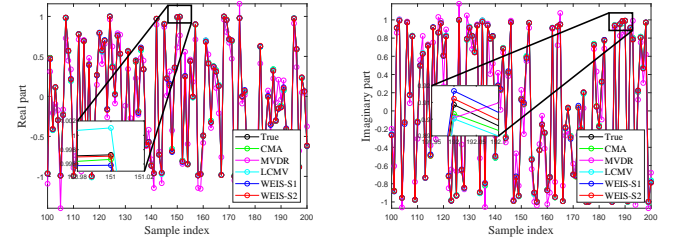
Fig. 4: Convergence performance of WEIS-ADMM in Exp.1.

(a) Real part.

(b) Imaginary part.

Fig. 6: SOI waveform estimation results in Exp. 1.

(a) SOI-PEE versus input SIR.

(b) NMSE versus input SIR.

Fig. 5: Power estimation and waveform NMSE in Exp. 1.

ations, namely $K_{\text{AO}}^{S1} = 500$ and $T_{\text{ADMM}}^{S1} = 500$ **(they are not the SC)** for AO and ADMM procedures of WEIS-S1, respectively, to observe the convergence behaviour; whereas $K_{\text{AO}}^{S2} = 100$ and $T_{\text{ADMM}}^{S2} = 2000$ **(they are not the SC)** are employed for WEIS-S2.

For AO, we first compute the objective function value (OFV) $\frac{\mathbf{w}^H(k)\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*(k)\mathbf{s}_0^T(k))\mathbf{X}^H\mathbf{w}(k)}{\mathbf{w}^H(k)\mathbf{w}(k)}$, the difference of successive OFVs (DOFVs) $|\frac{\mathbf{w}^H(k)\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*(k)\mathbf{s}_0^T(k))\mathbf{X}^H\mathbf{w}(k)}{\mathbf{w}^H(k)\mathbf{w}(k)} - \frac{\mathbf{w}^H(k-1)\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*(k-1)\mathbf{s}_0^T(k-1))\mathbf{X}^H\mathbf{w}(k-1)}{\mathbf{w}^H(k-1)\mathbf{w}(k-1)}|$ for WEIS-S1, and OFVs $\mathbf{w}^H(k)\mathbf{w}(k)$ and DOFVs $|\mathbf{w}^H(k)\mathbf{w}(k) - \mathbf{w}^H(k-1)\mathbf{w}(k-1)|$ for WEIS-S2, respectively. Then, we calculate the progress errors (PEs) $\{\frac{1}{M}\|\mathbf{w}(k) - \mathbf{w}(k-1)\|, \frac{1}{N}\|\mathbf{s}_0(k) - \mathbf{s}_0(k-1)\|\}$ of $\{\mathbf{w}(k), \mathbf{s}_0(k)\}$ in successive AO iterations for WEIS-S1 and WEIS-S2. In addition, we compute the CEs $\frac{1}{M+1}\|\bar{\mathbf{w}}(t) - \bar{\mathbf{v}}(t)\|$ and DOFVs $|\bar{\mathbf{w}}^H(t)\bar{\mathbf{R}}_1(k)\bar{\mathbf{w}}(t) - \bar{\mathbf{w}}^H(t-1)\bar{\mathbf{R}}_1(k)\bar{\mathbf{w}}(t-1)|$ for the ADMM procedure of the WEIS-S1 (WEIS-S1-ADMM) with different $k \in \{50, 100, \cdots, 500\}$ since $\bar{\mathbf{R}}_1(k)$ is related to k; whereas the CEs $\frac{1}{M+1}\|\mathbf{g}(t) - \mathbf{w}^H(t)\mathbf{U}_k\|$ and DOFVs $|\mathbf{w}^H(t)\mathbf{w}(t) - \mathbf{w}^H(t-1)\mathbf{w}(t-1)|$ are computed for the ADMM procedure of WEIS-S2 (WEIS-S2-ADMM) with different $k \in \{10, 20, \cdots, 100\}$.

We plot the values of OFVs, DOFVs, PEs of $\{\mathbf{s}_0, \mathbf{w}\}$ in AO procedure versus iteration number (IN) in Figs. 3(a) to (d) under different SIRs, respectively. For SIR=10dB, we plot the CEs and DOFVs of ADMM procedure versus IN in Figs. 4(a) and (b) under different values of k (IN of AO) for WEIS-S1 and WEIS-S2, respectively. We observe that: i) all the OFVs, DOFVs, PEs of $\{\mathbf{w}(k), \mathbf{s}_0(k)\}$ basically decrease with the increase of INs under different input SIRs; ii) for WEIS-S1, upon convergence, all the DOFVs are less than $10^{-15.68}$

, and the PEs of $\mathbf{w}(k)$ and $\mathbf{s}_0(k)$ remain below $10^{-9.84}$ and $10^{-11.65}$, respectively; iii) for WEIS-S2, all the DOFVs upon convergence are less than $10^{-15.05}$, and the PEs of $\mathbf{w}(k)$ and $\mathbf{s}_0(k)$ remain below $10^{-15.06}$ and $10^{-16.79}$, respectively. While from the ADMM results in Fig. 4, we see that: i) both the CEs and DOFVs exhibit the decreasing and converging trend at different outer AO iterations; ii) in ADMM procedure of WEIS-S1, the maximum CEs and DOFVs upon convergence are not larger than $10^{-5.6}$ and $10^{-5.3}$, respectively; and iii) in ADMM procedure of WEIS-S2, the maximum CEs and DOFVs upon convergence are not larger than $10^{-15.93}$ and $10^{-14.48}$, respectively.

Based on the aforementioned discussion, we consider the following SC settings for the four cycles of Fig. 1: i) for AO1 of WEIS-S1, SC1 is: the values of DOFVs, PEs of $\mathbf{s}_0$, and PEs of $\mathbf{w}$ being below 10^{-15}, 10^{-11}, and 10^{-9}, respectively; otherwise, reaching the preset maximum iteration number $K_{\text{AO}}^{S1} = 500$; ii) for AO2 of WEIS-S2, SC3 is: the values of DOFVs, PEs of $\mathbf{s}_0$ and PEs of $\mathbf{w}$ being below 10^{-15}, 10^{-16}, and 10^{-15}, respectively; otherwise, reaching the preset maximum iteration number $K_{\text{AO}}^{S2} = 100$. Whereas for the ADMM procedure in WEIS-S1, SC2 is: both the values of DOFVs and CEs being less than 10^{-5}; otherwise, reaching the preset maximum iteration number $K_{\text{ADMM}}^{S1} = 500$. Finally for the ADMM procedure in WEIS-S2, SC4 is: the values of DOFVs and CEs being below 10^{-14} and 10^{-15}, respectively; otherwise, reaching the preset maximum iteration number $K_{\text{ADMM}}^{S2} = 2000$.

Furthermore, to ascertain the convergence and its impact on the precision of power and waveform estimation, we calculate the corresponding SOI-power estimation error (SOI-PEE) $||\hat{\alpha}_0|^2 - |\alpha_0|^2|$ and NMSE upon convergence under different SIRs, where $\hat{\alpha}_0 = \frac{\mathbf{w}^H\mathbf{X}\mathbf{s}_0^*}{N}$. For comparison purposes, we implement LCAP (loading Capon beamformer) with loading factor $10\delta^2$ [7], MVDR [11], CMA (constant modulus beamforming with accurate steering vector) in Chapter 6 of [1], EIG (eigenspace-based) [56], and CVCNN-LR (CVCNN-the lower triangular elements of the sample covariance matrix (L$\mathbf{R}$)) and CVCNN-R (CVCNN-the whole sample covariance matrix ($\mathbf{R}$)) in [38]. From Figs. 5(a) and (b), we observe that: i) Both WEIS-S1 and WEIS-S2 outperform the competing beamformers in terms of SOI-PEE, that is, they provide more accurate SOI power estimation; ii) for NMSE, WEIS-S1 performs comparably with CMA and LCMV; whereas the refinement scheme from (16) enables WEIS-S2 to have an

(a) NMSE versus input SNR.   (b) NMSE versus input SIR.   (c) NMSE versus snapshot number.
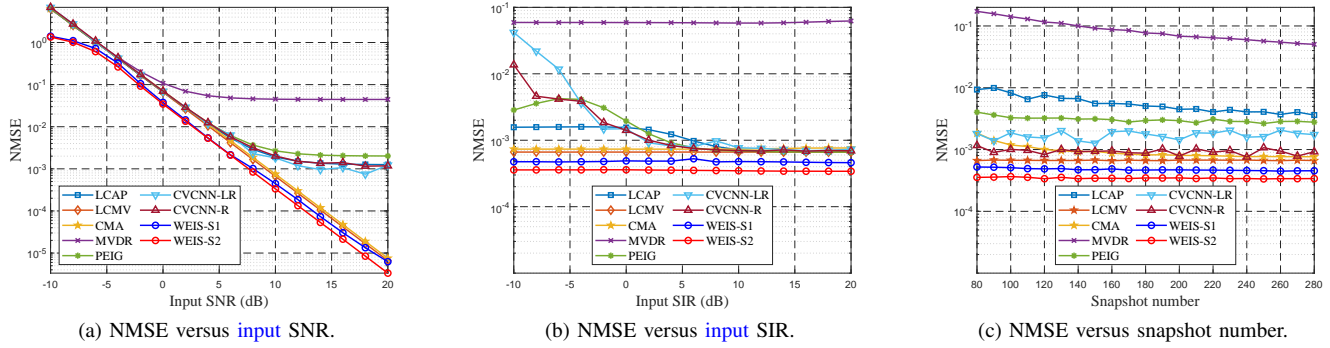
Fig. 7: NMSE versus input SNR, input SIR and snapshot number.

obvious NMSE improvement, achieving the highest waveform estimation accuracy. As a direct exhibit of accurate waveform estimate, in Fig. 6, we plot the real and imaginary components of the waveform estimate extracted from a single Monte Carlo run at SIR = 6dB.

Experiment 2: Output and Beampattern Performance in WEIS Algorithm: In this experiment, we set the sensor number as $M = 15$, and the directions of SOI and two interferences are $\theta_0 = 0°, \theta_1 = -50°$ and $\theta_2 = 70°$, respectively. Here we adopt the similar SC settings and maximum iteration numbers as Exp. 1, with $\rho = 0.05$ and $\rho = 0.1$ being applied for the two stages, respectively. In addition, similar initializations as Exp. 1 are used.

First, we investigate the effects of input SNR, input SIR, and snapshot number on the NMSE performance of WEIS algorithm. In Fig. 7, we plot the NMSE computed from different input SNR, input SIR, and snapshot number configurations. Especially, in Fig. 7(a), we set snapshot number $N = 300$, input SIR= 0dB, and the input SNR varies from -10dB to 20dB; whereas in Fig. 7(b), we set snapshot number $N = 300$, input SNR= 10dB, but the input SIR varies from -10dB to 20dB. In Fig. 7(c), input SNR= 10dB and input SIR= 0dB are set, but the snapshot number varies from 80 to 280. We can see that both WEIS-S1 and WEIS-S2 produce lower NMSE than the competing schemes; whereas WEIS-S2 outperforms WEIS-S1 with the use of the result of WEIS-S1, which is consistent with Exp. 1.

In addition to NMSE, we also explore the relationship between the output SINR, output SNR and input SNR, which are plotted in Figs. 8(a) and (b). It is observed that the WEIS-S1 occasionally performs worse than the CMA, and the output SNR of WEIS-S2 is slightly lower than that of the EIG, CVCNN-LR, CVCNN-R and LCMV. However, in high input SNR region, the output SINR of WEIS-S2 is closest to the optimal output SINR among all beamformers except for the LCMV, which requires the IDs prior. Furthermore, Fig. 8(c) depicts the beampattern comparisons at input SNR= 10dB, showing that WEIS-S2 not only produces the deep notches in the unknown IDs, i.e., -140.25dB in $-50°$ and -209.59dB in $70°$, but also has good sidelobe suppression effect.

B. Performance of WEIS-R Algorithm

In this subsection, the direction of SOI is $\theta_0 = 0°$, and we define the inaccurate DOA set $\Omega_{\hat{\theta}} = \{\hat{\theta}_n\}$, where $\hat{\theta} \in [\theta_0 - \Delta, \theta_0 + \Delta]$. In addition, we consider 200 angle values uniformly sampled from the region $\hat{\theta} \in [\theta_0 - \Delta, \theta_0 + \Delta]$ [1][15]. As a result, the center $\bar{\mathbf{a}}$ and the configuration matrix $\mathbf{A}$ of the ellipsoid used in (50) are given by [1][15]:

$$\bar{\mathbf{a}} = \frac{1}{\text{Card}(\Omega_{\hat{\theta}})} \sum_{\theta_n \in \Omega_{\hat{\theta}}} \mathbf{a}(\theta_n) \tag{76}$$

$$\mathbf{A} = \frac{1}{\epsilon \, \text{Card}(\Omega_{\hat{\theta}})} \sum_{\theta_n \in \Omega_{\hat{\theta}}} (\mathbf{a}(\theta_n) - \bar{\mathbf{a}})(\mathbf{a}(\theta_n) - \bar{\mathbf{a}})^H \tag{77}$$

where $\text{Card}(\Omega_{\hat{\theta}})$ represents the element number (sample size) of $\Omega_{\hat{\theta}}$, and the normalization factor ϵ is denoted by [1][15]:

$$\epsilon = \sup(\mathbf{a}(\theta_n) - \bar{\mathbf{a}})^H \mathbf{A}^{-1} \sup(\mathbf{a}(\theta_n) - \bar{\mathbf{a}}), \theta_n \in \Omega_{\hat{\theta}}. \tag{78}$$

In order to evaluate the performance of the WEIS-R algorithm, we compare with a number of beamformers, including the CMA beamformer based on inaccurate steering vector in Chapter 6 of [1], the least-squares (LS) beamformer [47], the principal eigenvector (PEIG) beamformer [57], the robust beamformer (ROB) [29], and the WCPO [16], in which diagonal loading parameters of LS, PEIG, ROB are set as 30, 50 and 50, respectively.

Experiment 3. Convergence Performance and SC Setting: In this experiment, the input SNR and input SIR are fixed while the snapshot number varies from 110 to 200. Besides, $\Delta = 1$ and we sample $\hat{\theta}$ uniformly. The sensor number is $M = 15$, the DOAs of two interferences are $\pm 45°$ while input SNR= 15dB and input SIR= 0dB. For both WEIS-R-S1 and WEIS-R-S2, we set $\rho = 0.1, K_{AO} = 50, T_{ADMM} = 200$, and adopt the same initializations as before.

From the AO results plotted in Fig. 9, we observe that: i) all the OFVs, DOFVs, and PEs, including WEIS-R-S1 and WEIS-R-S2, basically decrease and then converge with the increase of INs under different snapshots ($N \in \{110, 120, \cdots, 200\}$); ii) for WEIS-R-S1, all the DOFVs, and the PEs of $\mathbf{s}_0(k)$ and $\mathbf{w}(k)$ upon convergence are less than $10^{-7.81}$, $10^{-9.99}$, and $10^{-9.14}$, respectively; iii) for WEIS-R-S2, all the DOFVs, and the PEs of $\mathbf{s}_0(k)$ and $\mathbf{w}(k)$ upon convergence are less than $10^{-8.30}$, $10^{-10.14}$, and $10^{-9.44}$, respectively. While the AD-MM results versus AO iteration ($k \in \{1, 10, 20, 30, 40, 50\}$)
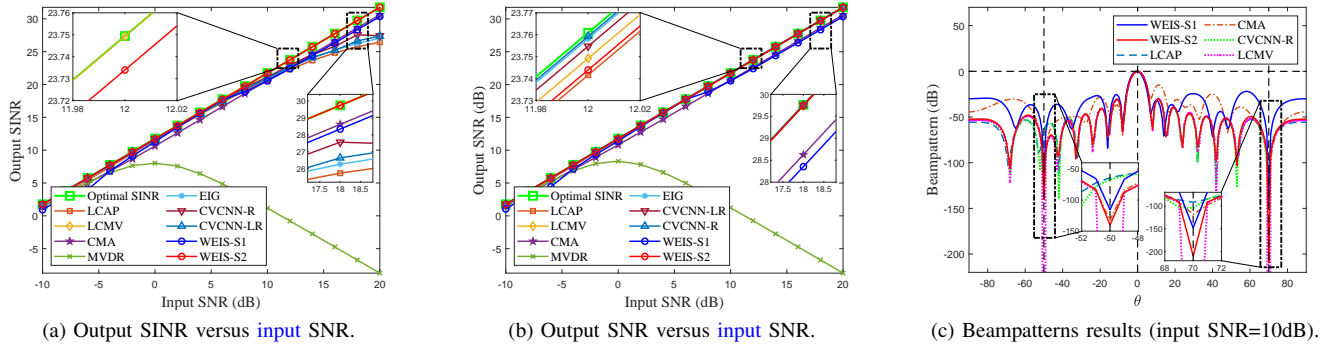
(a) Output SINR versus input SNR.

(b) Output SNR versus input SNR.

(c) Beampatterns results (input SNR=10dB).

Fig. 8: Comparison of output SINR, output SNR and beampatterns.

(a) OFV versus IN (AO).

(b) DOFV versus IN (AO).

(c) PEs of $\mathbf{s}_0(k)$ versus IN (AO).

(d) PEs of $\mathbf{w}(k)$ versus IN (AO).
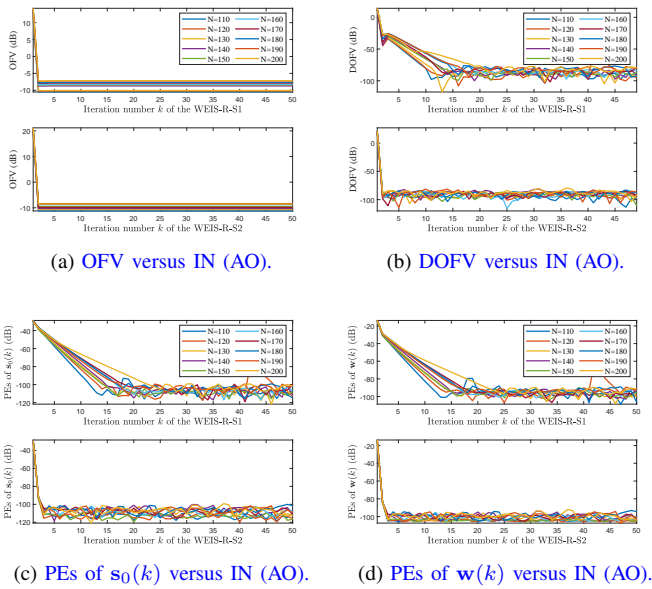
Fig. 9: Convergence performance of WEIS-R-AO in Exp. 3

(a) CEs and DOFVs versus IN (WEIS-R-S1-ADMM).

(b) CEs and DOFVs versus IN (WEIS-R-S2-ADMM).
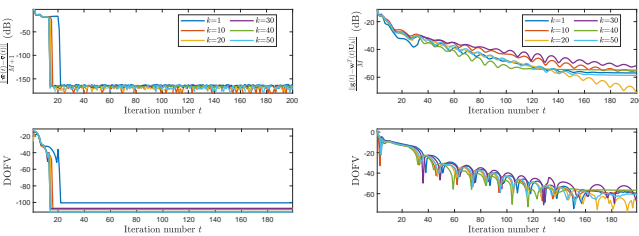
Fig. 10: Convergence performance of WEIS-R-ADMM in Exp.3.

in Fig. 10 indicate that: i) all the CEs and DOFVs, including WEIS-R-S1 and WEIS-R-S2, exhibit the descending and converging trends; ii) in ADMM procedure of WEIS-R-S1, all the CEs and DOFVs upon convergence are not larger than $10^{-16.28}$ and $10^{-10.04}$, respectively; and iii) in

ADMM procedure of WEIS-R-S2, all the CEs and DOFVs upon convergence are not larger than $10^{-5.16}$ and $10^{-5.67}$, respectively.

Based on the aforementioned discussion, we consider the following SC settings for the four cycles (SC5-SC8) of Fig. 2: i) for AO3 of WEIS-R-S1, SC5 is: the values of DOFVs, PEs of $\mathbf{s}_0$, and PEs of $\mathbf{w}$ being below 10^{-7}, 10^{-9}, and 10^{-9}, respectively, otherwise reaching the preset maximum iteration number $K_{\text{AO}}^{S1} = 50$; ii) for AO4 of WEIS-R-S2, SC7 is: the values of DOFVs, PEs of $\mathbf{s}_0$ and PEs of $\mathbf{w}$ being below 10^{-8}, 10^{-10}, and 10^{-9}, respectively, otherwise reaching the preset maximum iteration number $K_{\text{AO}}^{S2} = 50$. Whereas for the ADMM procedure in WEIS-R-S1, SC6 is: the values of DOFVs and CEs being below 10^{-10} and 10^{-16}, respectively, otherwise reaching the preset maximum iteration number $K_{\text{ADMM}}^{S1} = 200$. Finally for the ADMM procedure in WEIS-R-S2, SC8 is: both the values of DOFVs and CEs being less than 10^{-5}, otherwise reaching the preset maximum iteration number $K_{\text{ADMM}}^{S2} = 200$.

Experiment 4. Output and Beampattern Performance of WEIS-R: The results of Figs. 11 to 13 are obtained with the settings: the sensor number $M = 20$, snapshot number $N = 200$, input SIR = 0dB, and the DOAs of two interferences are $-40°$ and $70°$, respectively. Here we adopt the similar SC settings and maximum iteration numbers as Exp. 3. To evaluate the performance of WEIS-R-S1 and WEIS-R-S2, we consider two scenarios, i.e., $\Delta = 3$ and $\Delta = 6$, where $\rho = 10$ in WEIS-R-S1 and $\rho = 20$ in WEIS-R-S2 are set for the former ($\Delta = 3$); whereas $\rho = 10$ in WEIS-R-S1 and $\rho = 1$ in WEIS-R-S2 are adopted for the latter ($\Delta = 6$).

(a) $\Delta = 3$

(b) $\Delta = 6$
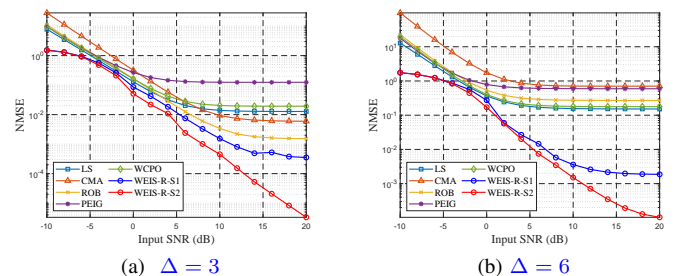
Fig. 11: NMSE versus input SNR for different values of Δ

As shown in Fig. 11, the NMSEs obtained by the WEIS-R-S2 are lower than those of other counterparts, particularly at moderate and high SNR regions, which implies that our proposed WEIS-R method estimates constant modulus waveform with higher accuracy.

Figs. 12(a) and (b) and Figs. 13(a) and (b) plot the results of output SINRs and output SNRs versus input SNR at $\Delta = 3$ and $\Delta = 6$, respectively, from which it is easily found that the WEIS-R-S2 has the higher output SINR and output SNR than those of the WEIS-R-S1 as well as the competing schemes in moderate and high input SNR ranges. The observations indicate again that the introduced WEIS-R-S2 is helpful in improving the output SINR and output SNR.

Moreover, for $\Delta = 3$, we use the obtained weight vectors from WEIS-R-S1 and WEIS-R-S2 to generate the beampatterns, as shown in Fig. 12(c). It is seen that the WEIS-R-S2 has the deepest notch levels compared with the other ones, i.e., -150.19dB in $-40°$ and -141.23dB in $70°$.

In addition, Fig. 13(c) shows the beampattern results at $\Delta = 6$, which indicates that the notches generated by the WEIS-R-S2 are more deeper (-85.44dB in $-40°$ and -90.68dB in $70°$) than others. The wide mainlobe in Fig. 13(c) costs more degree-of-freedom than 12(c) and thus the resultant notches are inferior to that of $\Delta = 3$.

VI. CONCLUSION

In this paper, we have proposed two-stage algorithms for constant modular waveform estimation and interference suppression via fractional programming, where one is developed for the known steering vector case; while the other is designed for the inaccurate steering vector case. Simulation results demonstrate that the proposed methods are superior to existing algorithms in terms of both the NMSE and the achieved notch level. As a future work, we will extend our studies to handle desired signals without modulus or energy constraints.

APPENDIX A: POSITIVE SEMIDEFINITE PROPERTY OF $\mathbf{R}_1(k)$

In (12), the denominator has the form of $\|\mathbf{w}^H\mathbf{X} - \alpha_0\mathbf{s}_0^T\|^2$. Thus, we apply the standard least-squares method to represent α_0 as $\alpha_0 = \frac{\mathbf{w}^H\mathbf{X}\mathbf{s}_0^*}{N}$ (as shown in (13)), where the unimodularity property of $\mathbf{s}_0$ is utilized (i.e., $\mathbf{s}_0^T\mathbf{s}_0^* = N$).

Then, inserting $\alpha_0 = \frac{\mathbf{w}^H\mathbf{X}\mathbf{s}_0^*}{N}$ into $\|\mathbf{w}^H\mathbf{X} - \alpha_0\mathbf{s}_0^T\|^2$ yields:

$$
\begin{aligned}
&\|\mathbf{w}^H\mathbf{X} - \alpha_0\mathbf{s}_0^T\|^2 \\
&= \|\mathbf{w}^H\mathbf{X} - \frac{1}{N}\mathbf{w}^H\mathbf{X}\mathbf{s}_0^*\mathbf{s}_0^T\|^2 = \|\mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\|^2 \\
&= \mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)^H\mathbf{X}^H\mathbf{w} \\
&= \mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T + \frac{1}{N^2}\mathbf{s}_0^*\mathbf{s}_0^T\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H\mathbf{w} \\
&= \mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T + \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H\mathbf{w} \\
&= \mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H\mathbf{w}. \quad (79)
\end{aligned}
$$

Employing the definition of $\mathbf{R}_1(k) = \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H$ in (20), we have:

$$
\mathbf{R}_1(k) = \left(\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\right)\left(\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\right)^H. \quad (80)
$$

According to (80), for any vector $\mathbf{w}$, we have:

$$
\mathbf{w}^H\mathbf{R}_1(k)\mathbf{w} = \|\mathbf{w}^H\mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\|^2 \geq 0, \quad (81)
$$

Therefore, $\mathbf{R}_1(k)$ is positive semidefinite.

APPENDIX B: ON STEP A2 SOLUTION OF STAGE 1 IN WEIS

Completing the square and ignoring the irrelevant terms in (34), we have:

$$
\begin{aligned}
\min_{\bar{\mathbf{w}}} \quad & \bar{\mathbf{w}}^H\bar{\mathbf{R}}_1(k)\bar{\mathbf{w}} + \frac{\rho}{2}\|\bar{\mathbf{w}} - \tilde{\mathbf{w}}(t+1)\|^2 \\
\text{s.t.} \quad & \bar{\mathbf{w}}^H\bar{\mathbf{I}}_{M+1}\bar{\mathbf{w}} = 1. \quad (82)
\end{aligned}
$$

Similar to (32), $\tilde{\mathbf{w}}(t+1) \in \mathbb{C}^{(M+1)\times 1}$ is introduced:

$$
\tilde{\mathbf{w}}(t+1) = \bar{\mathbf{v}}(t+1) - \frac{\boldsymbol{\lambda}(t)}{\rho}. \quad (83)
$$

The Lagrangian of (82) is easily constructed as:

$$
l_1(\bar{\mathbf{w}}, \lambda_w) = \bar{\mathbf{w}}^H\bar{\mathbf{R}}_2\bar{\mathbf{w}} - \frac{\rho}{2}\bar{\mathbf{w}}^H\tilde{\mathbf{w}} - \frac{\rho}{2}\tilde{\mathbf{w}}^H\bar{\mathbf{w}} + \lambda_w(\bar{\mathbf{w}}^H\bar{\mathbf{I}}_{M+1}\bar{\mathbf{w}} - 1), \quad (84)
$$

where the $t+1$ in $\tilde{\mathbf{w}}(t+1)$ and k in $\bar{\mathbf{R}}_1(k)$ are omitted,

$$
\bar{\mathbf{R}}_2 = \bar{\mathbf{R}}_1(k) + \frac{\rho}{2}\mathbf{I}_{M+1}, \quad (85)
$$

and λ_w is the Lagrange multiplier corresponding to the constraint $\bar{\mathbf{w}}^H\bar{\mathbf{I}}_{M+1}\bar{\mathbf{w}} = 1$.

Differentiating (84) with respect to $\bar{\mathbf{w}}$ yields:

$$
2(\bar{\mathbf{R}}_2 + \lambda_w\bar{\mathbf{I}}_{M+1})\bar{\mathbf{w}}(t+1) - \rho\tilde{\mathbf{w}} = \mathbf{0}, \quad (86)
$$

from which $\bar{\mathbf{w}}(t+1)$ is easily determined as:

$$
\bar{\mathbf{w}}(t+1) = \frac{\rho}{2}(\bar{\mathbf{R}}_2 + \lambda_w\bar{\mathbf{I}}_{M+1})^{-1}\tilde{\mathbf{w}}. \quad (87)
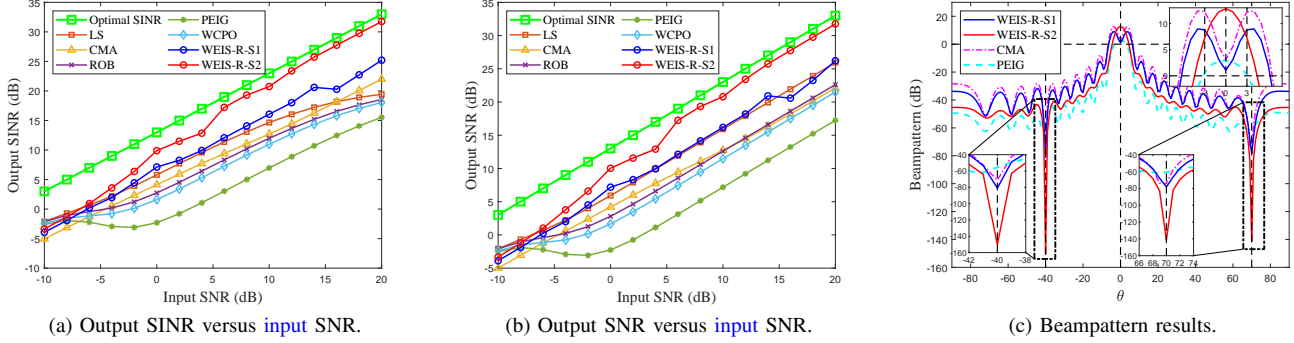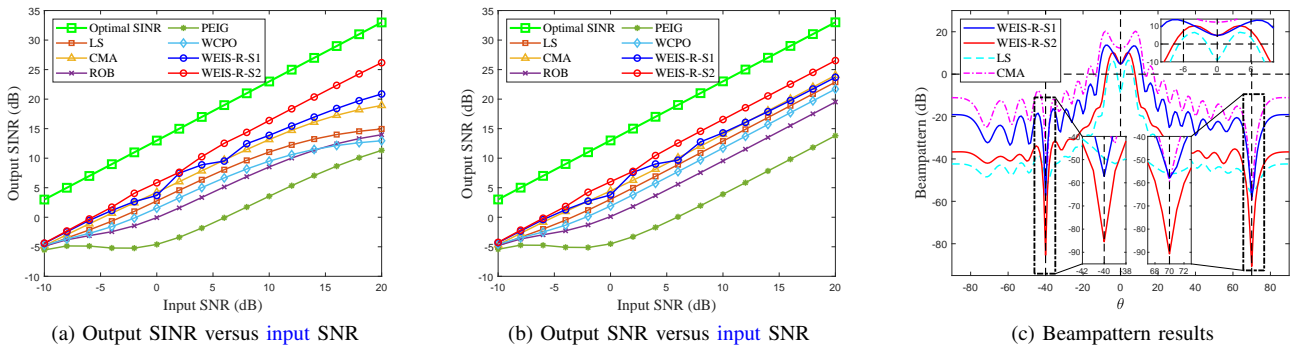$$

Inserting (87) into $\bar{\mathbf{w}}^H\bar{\mathbf{I}}_{M+1}\bar{\mathbf{w}} = 1$ leads to:

$$
\frac{\rho^2}{4}\tilde{\mathbf{w}}^H(\bar{\mathbf{R}}_2 + \lambda_w\bar{\mathbf{I}}_{M+1})^{-1}\bar{\mathbf{I}}_{M+1}(\bar{\mathbf{R}}_2 + \lambda_w\bar{\mathbf{I}}_{M+1})^{-1}\tilde{\mathbf{w}} = 1. \quad (88)
$$

Since the matrix $\bar{\mathbf{R}}_2$ is positive definite (see (85) or see [15] for details), we implement the eigenvalue decomposition (EVD) of $(\bar{\mathbf{R}}_2^{\frac{1}{2}})^{-1}\bar{\mathbf{I}}_{M+1}(\bar{\mathbf{R}}_2^{\frac{1}{2}})^{-1}$ to yield the eigenvector matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_{M+1}]$, where $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ and $\mathbf{U}\mathbf{U}^H = \mathbf{I}$. Next, the **generalized eigenvalue** pair of $\{\bar{\mathbf{R}}_2, \bar{\mathbf{I}}_{M+1}\}$ are given by:

$$
\sigma_m = \mathbf{u}_m^H\bar{\mathbf{R}}_2\mathbf{u}_m, \quad \bar{\sigma}_m = \mathbf{u}_m^H\bar{\mathbf{I}}_{M+1}\mathbf{u}_m, \quad (89)
$$

which are corresponding to the eigenvector $\mathbf{u}_m$ for $m = 1, \cdots, M$. Then, we have $\bar{\mathbf{R}}_2 = \mathbf{U}\text{diag}\{\sigma_1, \cdots, \sigma_M\}\mathbf{U}^H$ and $\bar{\mathbf{I}}_{M+1} = \mathbf{U}\text{diag}\{\bar{\sigma}_1, \cdots, \bar{\sigma}_M\}\mathbf{U}^H$. Thus, inserting $\bar{\mathbf{R}}_2 = \mathbf{U}\text{diag}\{\sigma_1, \cdots, \sigma_M\}\mathbf{U}^H$ and $\bar{\mathbf{I}}_{M+1} =$

(a) Output SINR versus input SNR. (b) Output SNR versus input SNR. (c) Beampattern results.

Fig. 12: Comparison of output SINR, output SNR and beampatterns ($\Delta = 3$).

(a) Output SINR versus input SNR (b) Output SNR versus input SNR (c) Beampattern results

Fig. 13: Comparison of output SINR, output SNR and beampatterns ($\Delta = 6$)

$\mathbf{U}\text{diag}\{\bar{\sigma}_1, \cdots, \bar{\sigma}_M\}\mathbf{U}^H$ into the left side of (88) (denoted as a function of λ_w, i.e., $g(\lambda_w)$) yields :

$$g(\lambda_w) = \frac{\rho^2}{4} \sum_{m=1}^{M+1} \frac{\bar{\sigma}_m |\tilde{\mathbf{w}}^H \mathbf{u}_m|^2}{(\sigma_m + \lambda_w \bar{\sigma}_m)^2} = 1. \qquad (90)$$

Similar to the analysis of [13], [15], $g(\lambda_w)$ is a monotonically decreasing function with respect to λ_w in the region $[\max_m - \frac{\sigma_m}{\bar{\sigma}_m}, +\infty)$. Since $\lim_{\lambda_w \to +\infty} g(\lambda_w) = 0$ and $\lim_{\lambda_w \to \max_m - \frac{\sigma_m}{\bar{\sigma}_m}} g(\lambda_w) = +\infty$, there exists a solution in the region $[\max_m - \frac{\sigma_m}{\bar{\sigma}_m}, +\infty)$, which can be determined via the bisection method [13].

APPENDIX C: ON STEP B2 SOLUTION OF STAGE 2 IN WEIS

Completing the squares and ignoring the irrelevant term, we rewrite (47) as:

$$\min_{\mathbf{g}, \mathbf{h}} \quad \|\mathbf{g} - \hat{\mathbf{g}}(t+1)\|^2 + \|\mathbf{h} - \hat{\mathbf{h}}(t+1)\|^2$$

$$\text{s.t.} \quad \frac{\|\mathbf{g}\|^2}{\|\mathbf{h}\|^2} \leq \hat{\sigma}^2, \qquad (91)$$

where

$$\hat{\mathbf{g}}(t+1) = \mathbf{w}(t+1)^H \mathbf{U}_k - \frac{\boldsymbol{\lambda}_g(t)}{\rho}, \qquad (92)$$

$$\hat{\mathbf{h}}(t+1) = \mathbf{w}(t+1) - \frac{\boldsymbol{\lambda}_h(t)}{\rho}. \qquad (93)$$

Then, we construct the Lagrangian for (91):

$$l_2(\mathbf{g}, \mathbf{h}, \mu) = \|\mathbf{g} - \hat{\mathbf{g}}(t+1)\|^2 + \|\mathbf{h} - \hat{\mathbf{h}}(t+1)\|^2 \\ - \mu(\|\mathbf{h}\|^2 \hat{\sigma}^2 - \|\mathbf{g}\|^2), \qquad (94)$$

where μ is the Lagrange multiplier. According to the Karush-Kuhn-Tucker (KKT) theory, we have

$$\frac{\partial l_2(\mathbf{g}, \mathbf{h}, \mu)}{\partial \mathbf{g}} = 2(\mathbf{g} - \hat{\mathbf{g}}(t+1) + \mu \mathbf{g}) = 0, \qquad (95)$$

$$\frac{\partial l_2(\mathbf{g}, \mathbf{h}, \mu)}{\partial \mathbf{h}} = 2(\mathbf{h} - \hat{\mathbf{h}}(t+1) - \mu \hat{\sigma}^2 \mathbf{h}) = 0, \qquad (96)$$

$$\mu(\|\mathbf{h}\|^2 \hat{\sigma}^2 - \|\mathbf{g}\|^2) = 0, \qquad (97)$$

$$\|\mathbf{g}\|^2 \leq \hat{\sigma}^2 \|\mathbf{h}\|^2. \qquad (98)$$

Thus, we have $\mathbf{g} = \frac{\hat{\mathbf{g}}(t+1)}{1+\mu}$ and $\mathbf{h} = \frac{\hat{\mathbf{h}}(t+1)}{1-\mu\hat{\sigma}^2}$. Then, we discuss μ from two cases: i) if $\mu = 0$ and $\|\mathbf{g}\|^2 < \hat{\sigma}^2 \|\mathbf{h}\|^2$, we have $\mathbf{g} = \hat{\mathbf{g}}(t+1)$ and $\mathbf{h} = \hat{\mathbf{h}}(t+1)$; and ii) if $\mu \neq 0$ and $\|\mathbf{g}\|^2 = \hat{\sigma}^2 \|\mathbf{h}\|^2$, μ should satisfy:

$$\frac{\|\hat{\mathbf{g}}(t+1)\|^2}{(1+\mu)^2} = \hat{\sigma}^2 \frac{\|\hat{\mathbf{h}}(t+1)\|^2}{(1-\mu\hat{\sigma}^2)^2}, \qquad (99)$$

the solution to which is given by:

$$\mu = \frac{-(2A + 2\hat{\sigma}^2) + \sqrt{8A\hat{\sigma}^2 + 4\hat{\sigma}^4 A + 4A}}{2(A - \hat{\sigma}^4)}, \qquad (100)$$

where $A = \frac{\|\hat{\mathbf{h}}(t+1)\|^2 \hat{\sigma}^2}{\|\hat{\mathbf{g}}(t+1)\|^2}$.

APPENDIX D: ON STEP C2 SOLUTION OF STAGE 1 IN WEIS-R

In Step C2, we may divide (69) into two independent subproblems, i.e.,

$$\min_{\mathbf{c},q} \ (q - \hat{q}(t+1))^2 + \|\mathbf{c} - \hat{\mathbf{c}}(t+1)\|^2$$
$$\text{s.t.} \ \|\mathbf{c}\|^2 \leq q, \tag{101}$$

and

$$\min_{\mathbf{d},p} \ (p - \hat{p}(t+1))^2 + \|\mathbf{d} - \hat{\mathbf{d}}(t+1)\|^2$$
$$\text{s.t.} \ \|\mathbf{d}\|^2 \geq p, \tag{102}$$

where the squares are completed and the constant terms are ignored, and $\{\hat{\mathbf{c}}(t+1) \in \mathbb{C}^{M \times 1}, \hat{\mathbf{d}}(t+1) \in \mathbb{C}^{M \times 1}, \hat{q}(t+1) \in \mathbb{C}^{1 \times 1}, \hat{p}(t+1) \in \mathbb{C}^{1 \times 1}\}$ are introduced for the presentation simplicity, i.e.,

$$\hat{\mathbf{c}}(t+1) = \mathbf{R}_1^{\frac{1}{2}}(k)\mathbf{w}(t+1) - \frac{\boldsymbol{\xi}(t)}{\rho}, \tag{103}$$

$$\hat{\mathbf{d}}(t+1) = \mathbf{w}(t+1) - \frac{\boldsymbol{\zeta}(t)}{\rho}. \tag{104}$$

$$\hat{p}(t+1) = \bar{p}(t+1) - \frac{\kappa(t)}{\rho}, \tag{105}$$

$$\hat{q}(t+1) = \bar{q}(t+1) - \frac{\lambda(t)}{\rho}, \tag{106}$$

The convex problem in (101) may be solved via the CVX toolbox. In fact, the nonconvex problem in (102) can still be solved (via the semidefinite relaxation trick [61]) by the CVX toolbox, but it requires introducing a high-dimensional auxiliary matrix variable. Here we derive the solution for the nonconvex problem in (102) according to two cases:

Case 1: if $\|\hat{\mathbf{d}}(t+1)\|^2 \geq \hat{p}(t+1)$, we have

$$\begin{cases} \mathbf{d}(t+1) = \hat{\mathbf{d}}(t+1) \\ p(t+1) = \hat{p}(t+1), \end{cases} \tag{107}$$

Case 2: if $\|\hat{\mathbf{d}}(t+1)\|^2 \leq \hat{p}(t+1)$, $\mathbf{d}(t+1)$ is given by

$$\mathbf{d}(t+1) = \frac{\hat{\mathbf{d}}(t+1)}{\|\hat{\mathbf{d}}(t+1)\|}\sqrt{p}, \tag{108}$$

with the assumption that p is known. Thus, inserting (108) into (102) yields an optimization problem on a single-variable p:

$$\min_p \ (p - \hat{p}(t+1))^2 + \|\frac{\hat{\mathbf{d}}(t+1)}{\|\hat{\mathbf{d}}(t+1)\|}\sqrt{p} - \hat{\mathbf{d}}(t+1)\|^2. \tag{109}$$

Let $\gamma = \sqrt{p} \geq 0$, and we transform (109) into an optimization problem on γ without the square root operation:

$$\min_\gamma \ \gamma^4 + (1 - 2\hat{p}(t+1))\gamma^2 - 2\gamma\|\hat{\mathbf{d}}(t+1)\|$$
$$\text{s.t.} \ \gamma \geq 0, \tag{110}$$

the derivative of which (denoted by $g(\gamma)$) with regard to γ is set as 0, i.e.,

$$\frac{\mathrm{d}g(\gamma)}{\mathrm{d}\gamma} = 4\gamma^3 + 2(1 - 2\hat{p}(t+1))\gamma - 2\|\hat{\mathbf{d}}(t+1)\| = 0. \tag{111}$$

With the Cardano's formula, we can obtain three roots to (111). Then, we select the real nonnegative one $\gamma(t+1)$ with the smallest $g(\gamma)$ to generate $p(t+1)$.

APPENDIX E: COMPUTATIONAL COMPLEXITY AND LOCAL CONVERGENCE ANALYSIS

We first evaluate the computational complexity of multiplication operations of the presented methods.

For WEIS-S1, the complexity of determining $\mathbf{s}_0$ is $\mathcal{O}(NM)$, and in the inner cycle employing the ADMM, the complexities of updating $\bar{\mathbf{v}}$, $\bar{\mathbf{w}}$ and $\boldsymbol{\lambda}$ are $\mathcal{O}(2M)$, $\mathcal{O}(M^3)$ [62], $\mathcal{O}(M)$, respectively, in which the middle term is due to the inversion operation of an $M \times M$ matrix [62]. Similarly, the complexity of updating $\{\mathbf{w}, \mathbf{g}, \mathbf{h}, \boldsymbol{\lambda}_g, \boldsymbol{\lambda}_h\}$ in the single inner cycle of WEIS-S2 is $\mathcal{O}(2M^2 + 2M)$. As a result, the WEIS's overall computational complexity is $\mathcal{O}(K_1(NM + T_1(M^3 + 3M))) + K_2(NM + T_2(2M^2 + 2M))))$, where K_1 and T_1 correspond to the total iteration numbers of outer (AO) and inner (ADMM) cycles of WEIS-S1, respectively, and K_2 and T_2 are similarly defined.

For WEIS-R-S1, the primal-dual interior point method with complexity $\mathcal{O}(M^3 \log(\varsigma))$ may be applied to determine $\mathbf{w}$ [54], where ς is the acceptable duality gap. Additionally, the complexity of remaining steps to obtain $\{p, q, \mathbf{c}, \mathbf{d}, \lambda, \kappa, \boldsymbol{\xi}, \boldsymbol{\zeta}\}$ is $\mathcal{O}(2M^2 + 3M)$. Then, the total computational complexity of WEIS-R is $\mathcal{O}(K_1(NM + T_1(M^3 \log(\varsigma) + 2M^2 + 3M))) + K_2(NM + T_2(M^3 \log(\varsigma) + 2M^2 + 2M))))$.

Now, we discuss the convergence of the proposed methods. Note that all the provided methods, including WEIS-S1, WEIS-S2, WEIS-R-S1 and WEIS-R-S2, are realized with AO procedure [58]. Here we take the AO procedure of WEIS-S1 (shown in (17)-(18)) as an example to discuss the convergence with the definition (i.e., the objective function of (14)):

$$f_1(\mathbf{s}_0, \mathbf{w}) = \frac{\mathbf{w}^H \mathbf{X}(\mathbf{I}_N - \frac{1}{N}\mathbf{s}_0^*\mathbf{s}_0^T)\mathbf{X}^H\mathbf{w}}{\mathbf{w}^H\mathbf{w}}. \tag{112}$$

Since the $\mathbf{s}_0(k)$ computed from (17) satisfies

$$\mathbf{s}_0(k) = \arg \ \min_{\mathbf{s}_0} \ f_1(\mathbf{s}_0, \mathbf{w}(k-1))$$
$$\text{s.t.} \ |s_0(n)| = 1, n = 1, \cdots, N, \tag{113}$$

$f_1(\mathbf{s}_0, \mathbf{w}(k-1))$ surely attains its minimum at $\mathbf{s}_0 = \mathbf{s}_0(k)$ under unimodular constraints. Therefore, we have:

$$f_1(\mathbf{s}_0(k), \mathbf{w}(k-1)) \leq f_1(\mathbf{s}_0(k-1), \mathbf{w}(k-1)). \tag{114}$$

In addition, the ADMM procedure shown in (19)-(36) solves the problem $\min_{\mathbf{w}} \ f_1(\mathbf{s}_0(k), \mathbf{w})$ s.t. $\mathbf{w}^H\mathbf{a}(\theta_0) = 1$ to obtain $\mathbf{w}(k)$. Therefore, when the obtained $\mathbf{w}(k)$ provided by the ADMM procedure is a local solution to (18), $f_1(\mathbf{s}_0(k), \mathbf{w})$ will attain one of its locally minima at $\mathbf{w} = \mathbf{w}(k)$ under the constraints $\mathbf{w}^H\mathbf{a}(\theta_0) = 1$. That is, we have:

$$f_1(\mathbf{s}_0(k), \mathbf{w}(k)) \leq f_1(\mathbf{s}_0(k), \mathbf{w}(k-1)). \tag{115}$$

Combining (114) and (115) yields:

$$f_1(\mathbf{s}_0(k), \mathbf{w}(k)) \leq f_1(\mathbf{s}_0(k), \mathbf{w}(k-1))$$
$$\leq f_1(\mathbf{s}_0(k-1), \mathbf{w}(k-1)), \tag{116}$$

which shows that the objective function value is non-increasing along with the iterations, and thus the AO procedure is convergent (or see [58] for details). Similar procedure derivations and analyses also hold for WEIS-S2, WEIS-R-S1, and WEIS-R-S2.

Note that the ADMM convergence analysis on nonconvex problems still remains an open problem [50], [59], [60], [63], [64], [65]. Here we borrow the idea from them to discuss the procedure shown in (19)-(36).

To illustrate the convergence of (19)-(36) employing the ADMM [50], we denote $\{\bar{\mathbf{w}}(t), \bar{\mathbf{v}}(t), \boldsymbol{\lambda}(t)\}$ as the obtained results at the t-th iteration with $\rho > 0$. According to (35), we introduce the condition [50], [59], [60], [63], [64], [65]:

$$\lim_{t \to \infty} \boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}(t) = \mathbf{0}. \tag{117}$$

Therefore, combining (35) and (117) yields

$$\lim_{t \to \infty} \bar{\mathbf{w}}(t+1) - \bar{\mathbf{v}}(t+1) = \mathbf{0}. \tag{118}$$

Moreover, since (31) is convex, there exists a stationary point $\bar{\mathbf{v}}^\diamond$ satisfying $\lim_{t \to \infty} \bar{\mathbf{v}}(t) = \bar{\mathbf{v}}^\diamond$. Then we construct the inequality:

$$\|\bar{\mathbf{w}}(t)\| \le \|\bar{\mathbf{w}}(t) - \bar{\mathbf{v}}(t)\| + \|\bar{\mathbf{v}}(t)\| \tag{119}$$

where the vector $\bar{\mathbf{w}}(t)$ is also bounded, and thus there exists a stationary point $\bar{\mathbf{w}}^\diamond$ satisfying:

$$\lim_{t \to \infty} \bar{\mathbf{w}}(t) = \bar{\mathbf{w}}^\diamond. \tag{120}$$

Therefore, we have the consensus convergence result [59], [60], [63], [64], [65]:

$$\lim_{t \to \infty} \bar{\mathbf{w}}(t+1) - \bar{\mathbf{v}}(t+1) = \bar{\mathbf{w}}^\diamond - \bar{\mathbf{v}}^\diamond = \mathbf{0}. \tag{121}$$

Additionally, we have [50], [59], [60], [63], [64], [65]:

$$(\bar{\mathbf{w}}^\diamond)^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}^\diamond \le \bar{\mathbf{w}}^H(t+1) \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}(t+1) \\ + \Re\{(\boldsymbol{\lambda}^\diamond)^H \mathbf{r}(t+1)\}, \tag{122}$$

and

$$\bar{\mathbf{w}}^H(t+1) \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}(t+1) - (\bar{\mathbf{w}}^\diamond)^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}^\diamond \le \\ - \rho \Re\{(\bar{\mathbf{v}}(t+1) - \bar{\mathbf{v}}(t))^H (-\mathbf{r}(t+1) + \bar{\mathbf{v}}(t+1) - \bar{\mathbf{v}}^\diamond)\} \\ - \Re\{\boldsymbol{\lambda}^H(t+1) \mathbf{r}(t+1)\}, \tag{123}$$

where $\boldsymbol{\lambda}^\diamond$ is a stationary point, and $\mathbf{r}(t+1) = \bar{\mathbf{w}}(t+1) - \bar{\mathbf{v}}(t+1)$.

The right hand side in (122) goes to zero as $t \to \infty$ since the consensus error $\mathbf{r}(t+1)$ goes to $\mathbf{0}$ [50]. On the other hand, the right hand side in (123) goes to $\mathbf{0}$ as $t \to \infty$, because $\bar{\mathbf{v}}(t+1) - \bar{\mathbf{v}}^\diamond$ is bounded and both $\mathbf{r}(t+1)$ and $\bar{\mathbf{v}}(t+1) - \bar{\mathbf{v}}(t)$ go to $\mathbf{0}$. Thus, we have $\lim_{t \to +\infty} \bar{\mathbf{w}}^H(t+1) \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}(t+1) = (\bar{\mathbf{w}}^\diamond)^H \bar{\mathbf{R}}_1(k) \bar{\mathbf{w}}^\diamond$, i.e., the objective is convergent [50].

Similar results hold for other ADMM procedures, including those of WEIS-S1, WEIS-S2, WEIS-R-S1, and WEIS-R-S2.

REFERENCES

[1] J. Li, and P. Stoica, *Robust Adaptive Beamforming*, Eds. Hoboken, NJ: Wiley, 2006.

[2] A. M. Elbir, K. V. Mishra, S. A. Vorobyov and R. W. Heath, "Twenty-five years of advances in beamforming: From convex and nonconvex optimization to learning techniques," *IEEE Signal Process. Magazine*, vol. 40, no. 4, pp. 118-131, Jun. 2023.

[3] B. Ottersten, R. Roy and T. Kailath, "Signal waveform estimation in sensor array processing.," *Twenty-Third Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA,USA, pp. 787-791,1989.

[4] B.D. van Veen, and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Process. Magazine*, vol. 5, pp. 4-24, 1998.

[5] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 13655-1376, 1987.

[6] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, vol. 30, no. 1, pp. 27-34, Jan. 1982.

[7] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 397-401, Jul. 1988.

[8] L. Tong, Y. Inouye and R. -W. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2461-2470, 1993.

[9] H. Liu, and G. Xu, "A subspace method for signature waveform estimation in synchronous CDMA systems," *IEEE Trans. Communications*, vol. 44, no. 10, pp. 1346-1354, Oct. 1996.

[10] K. Zarifi and A. B. Gershman, "Blind subspace-based signature waveform estimation in BPSK-modulated DS-CDMA systems with circular noise," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3592-3602, 2006.

[11] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408-1418, 1969.

[12] A. Elnashar, S. M. Elnoubi and H. A. El-Mikati, "Further study on robust adaptive beamforming with optimum diagonal loading," *IEEE Trans. Antennas and Propagation*, vol. 54, no. 12, pp. 3647-3658, Dec. 2006.

[13] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol.51, pp. 1702-1715, 2003.

[14] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust Capon beamformer," *IEEE Trans. Signal Process.*, vol.52, pp. 240-2423, 2004.

[15] R. G. Lorenz, and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, pp.1684-1696, 2005.

[16] S. A. Vorobyov, A. B. Gershman, and Z. -Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, pp. 313-324, 2003.

[17] Y. Gu, and A. Leshem, "Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3881-3885, Jul. 2012.

[18] Z. L. Yu, W. Ser, M. H. Er, Z. Gu and Y. Li, "Robust adaptive beamformers based on worst-case optimization and constraints on magnitude response," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2615-2628, 2009.

[19] Z. L. Yu, Z. Gu, J. Zhou, Y. Li, W. Ser and M. H. Er, "A robust adaptive beamformer based on worst-case semi-definite programming," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5914-5919, 2010.

[20] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664-678, Feb. 2010.

[21] B. D. van Veen, "Minimum variance beamforming with soft response constraints," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 1964-1972, Sep. 1991.

[22] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Magazine*, vol. 27, no. 3, pp. 62-75, May 2010.

[23] P. Stoica, Z. Wang, and J. Li, "Robust Capon beamforming," *IEEE Signal Process. Letters*, vol. 10, no. 6, pp. 172-175, Jun. 2003.

[24] J. Liu, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Adaptive beamforming with sidelobe control: A second-order cone programming approach," *IEEE Signal Process. Letters*, vol. 10, no. 11, pp. 331-334, Nov. 2003.

[25] X. Jiang, W. -J.Zeng, A. Yasotharan, H. C. So and T. Kirubarajan, "Robust beamforming by linear programming," *IEEE Trans. Signal Process.*, vol.62, no.7, pp.1834-1849, Apr. 2014.

[26] C. -Y. Chen and P. P. Vaidyanathan, "Quadratically constrained beamforming robust against direction-of-arrival mismatch," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4139-4150, Aug. 2007.

[27] K. L. Bell, Y. Ephraim, and H. L. van Trees, "A Bayesian approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 48, pp. 386-398, 2000.

[28] A. Hassanien, S. A. Vorobyov, and K. M. Wong, "Robust adaptive beamforming using sequential quadratic programming: An iterative solution to the mismatch problem," *IEEE Signal Process. Letters*, vol. 15, pp. 733-736, 2008.

[29] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257-2269, Sep. 2003.

[30] X. Jiang, W.-J. Zeng, A. Yasotharan, H. C. So and T. Kirubarajan, "Minimum dispersion beamforming for non-Gaussian signals," *IEEE Trans. Signal Process.*, vol.62, no.7, pp.1879-1893, Apr. 2014.

[31] X. Jiang, W.-J. Zeng, A. Yasotharan, H. C. So and T. Kirubarajan, "Quadratically constrained minimum dispersion beamforming via gradient projection," *IEEE Trans. Signal Process.*, vol.63, no.1, pp.192-205, Jan. 2015.

[32] X. Jiang, J. Chen, X. Liu, A. M. Zoubir and Z. Zhou, "Phase-only robust minimum dispersion beamforming," *IEEE Trans. Signal Process.*, vol. 68, pp. 5664-5679, 2020.

[33] Y. Huang, H. Fu, S. A. Vorobyov and Z. -Q. Luo, "Robust adaptive beamforming via worst-case SINR maximization with nonconvex uncertainty sets," *IEEE Trans. Signal Process.*, vol. 71, pp. 218-232, 2023.

[34] Y. Huang, S. A. Vorobyov and Z. -Q. Luo, "Quadratic matrix inequality approach to robust adaptive beamforming for general-rank signal model," *IEEE Trans. Signal Process.*, vol. 68, pp. 2244-2255, 2020.

[35] Z. Tian, K. L. Bell, and H. L. Van Trees, "A recursive least squares implementation for LCMP beamforming under quadratic constraint," *IEEE Trans. Signal Process.*, vol. 49, no. 6, pp. 1138-1145, Jun. 2001.

[36] Y.-H. Chen, and C.-T. Chiang, "Adaptive beamforming using the constrained Kalman filter," *IEEE Trans. Antennas and Propagation*, vol. 41, no. 11, pp. 1576-1580, Nov. 1993.

[37] A. El-Keyi, T. Kirubarajan and A. B. Gershman, "Robust adaptive beamforming based on the Kalman filter," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3032-3041, Aug. 2005.

[38] S. Mohammadzadeh, V. H. Nascimento, R. C. de Lamare and N. Hajarolasvadi, "Robust beamforming based on complex-valued convolutional neural networks for sensor arrays," *IEEE Signal Process. Lett.*, vol. 29, pp. 2108-2112, 2022.

[39] P. Stoica, H. He, and J. Li, "New algorithms for designing unimodular sequences with good correlation properties," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1415-1525, Apr. 2009.

[40] H. He, P. Stoica, and J. Li, "Designing unimodular sequence sets with good correlations – including an application to MIMO radar," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4391-4405, Nov. 2009.

[41] W. Fan, J. Liang and J. Li, "Constant modulus MIMO radar waveform design with minimum peak sidelobe transmit beampattern," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4207-4222, Aug. 2018.

[42] J. Liang, H. C. So, J. Li and A. Farina, "Unimodular sequence design based on alternating direction method of multipliers," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5367-5381, 2016.

[43] A. van der Veen, and A. Paulraj, "An analytical constant modulus algorithms," *IEEE Trans. Signal Process.*, vol. 44, pp. 1136-1155, 1996.

[44] A. van der Veen, "Statistical performance analysis of the algebraic constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 50, pp. 3083-3097, 2002.

[45] A. van der Veen, "Asymptotic properties of the algebraic constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 49, pp. 1796-1807, 2001.

[46] O. Besson, A. A. Monakov, and C. Chalus, "Signal waveform estimation in the presence of uncertainties about the steering vector," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2432-2440, 2004.

[47] Y. C. Eldar, A. Nehorai, and P. S. La Rosa, "An expected least-squares beamforming approach to signal estimation with steering vector uncertainties," vol.13, no. 5, pp. 288-291, May 2006.

[48] Y. C. Eldar, A. Nehorai, and P. S. La Rosa, "A competitive mean-squared error approach to beamforming," *IEEE Trans. Signal Process.*, vol. 55, no.11, pp. 5143-5154, 2007.

[49] Y. C. Eldar and A. Nehorai, "Mean-squared error beamforming for signal estimation: A competitive approach, in *Robust Adaptive Beamforming*, E, J. Li and P. Stoica, Eds. Hoboken, NJ: Wiley, 2006.

[50] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol.3, no.1, pp.1-122, 2011.

[51] D. Gabay, "Applications of the method of multipliers to variational inequalities," *in Augmented Lagrangian Methods: Applications to the Solution of Boundary-Value Problems*, North-Holland: Amsterdam,1983.

[52] J. Liang, H. C. So, J. Li, A. Farina, D. Zhou, "On optimizations with magnitude constraints on frequency or angular responses," *Signal Process.*, vol. 145, pp.214-224, 2018.

[53] W. Fan, J. Liang, G. Yu, H. C. So and J. Li, "Robust Capon beamforming via ADMM," *2019 IEEE International Conference on Acoustics, Speech and Signal Process.*, pp. 4345-4349, Brighton, UK, 2019.

[54] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014.

[55] Y. C. Eldar, A. Nehorai and P. S. La Rosa, "A competitive mean-squared error approach to beamforming," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5143-5154, Nov. 2007.

[56] D. D. Feldman and L. J. Griffiths,"A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, pp. 867-876, Apr. 1994.

[57] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, Vol. 53, pp. 305-324, Dec. 1999.

[58] J. C. Bezdek, R. J. Hathaway, "Convergence of alternating optimization," *Neural, Parallel and Scientific Computations*, vol. 11, no. 4, pp. 351-368, Dec. 2003.

[59] C. G. Tsinos and B. Ottersten, "An efficient algorithm for unit-modulus quadratic programs with application in beamforming for wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 25, no. 2, pp. 169-173, Feb. 2018

[60] Z.Wen, C.Yang, X. Liu, and S. Marchesini, "Alternating direction methods for classical and ptychographic phase retrieval," *Inverse Problems*, vol. 28, no. 11, Nov. 2012, Art. no. 115010.

[61] Z.-Q. Luo, W.-K. Ma, A. M. -C. So, Y. Ye and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Magazine*, vol. 27, no. 3, pp. 20-34, May 2010.

[62] G. Strang, *Linear Algebra and Its Applications (4th Edition)*, Brooks, 2004.

[63] M. Hong, and Z.-Q. Luo, "On the linear convergence of the alternating direction method of multipliers," *Math. Program.*, vol. 162, no. 1, pp. 165-199, Mar. 2017.

[64] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM J. Optim.*, vol. 26, no. 1, pp. 337-364, 2016.

[65] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *J. Scientific Comput.*, vol. 78, no.1, pp. 29-63, 2019.